

Ljiljana KEČA and Milica MARČETA¹

IMPLEMENTATION OF PANEL DATA IN MARKETING RESEARCHES IN FORESTRY

SUMMARY

Series of panel data are a combination of comparative data and time series. One of the most significant advantages of using an analysis based on a series of panel data refers to obtaining maximum information from a limited number of observations per unit of observation in a given period, which is very frequent in forestry. The aim of the research is to analyse the individual market factors, examine their interdependence, and establish strength of linkage. In this case study the goal is to examine the influence of independent variables (placement on the domestic market, export and sales prices of non-timber forest products - NTFPs) on the dependent variable - the purchase of raw materials.

Variables that are taken into consideration were: buying, selling on the domestic market, export and selling prices of certain types of NTFPs. In the analysis were used R Program and the special package, which is dealing with panel data. Based on Pooling OLS estimators determined that the variable export and sales on the domestic market are important to purchase of NTFPs. By analyzing the variations between objects of observation (between) it is established, as the previous estimator that the variables export and prices are important to the purchase of NTFPs. According to First difference estimator export and sales price are the factors that most influence the purchase of raw NTFPs. According to all tree estimators: Pooled OLS; between and First differences, it was found that the placement on the domestic market variable which, compared with other, largely determines the purchase of raw NTFPs (in terms of the analyzed data). This is evident based on the greatest values that range in interval 0.80 to 1.04 for the domestic market.

Keywords: panel data, market, forestry, estimator, strategy.

INTRODUCTION

Series of panel data are a combination of comparative data and time series (Jovičić, 2010). As such, they are very suitable for different types of analysis, especially applicable to the analysis of market. These series are mainly oriented to the analysis of heterogeneity among observation units, as well as the analysis of changes in the structure over time (Greene, 2005). Heterogeneity means that the impact of independent variables on the dependent varies from unit to the

¹ Ljiljana Keča, (corresponding author: ljiljana.keca@sfb.bg.ac.rs, University of Belgrade, Faculty of Forestry, SERBIA, Milica Marčeta, University of Belgrade, Faculty of Forestry, SERBIA
Notes: The authors declare that they have no conflicts of interest. Authorship Form signed online.

object of observation and / or over time (Dragutinović – Mitrović, 2001, Baltagi, 2008). These models allow examining nature of the variability of parameters that are important for the analysis. In addition, as one of the most significant advantages of using an analysis based on a series of panel data refers to obtaining maximum information from a limited number of observations per unit of observation in a given period (Dragutinović – Mitrović, 2001). Assumptions about the homogeneity of observation units, which are common in classical statistical analysis, often in the economy are not met. In such situations, it is necessary to use models that combine time series and comparative data, because their use does not require the usual assumptions about the constancy of parameters.

Testing of variations through panel data series can be directed in many directions, thus are distinguished:

- Overall (variations in time and within the data to be analysed, namely in the specific case of the types of products),
- Between (variation between species analysed products) (Kennedy, 2008),
- Within variation (among species of analysed products over time) (Račić 2013).

In panel researches are most often used the following regression models (Katchova, 2011):

- model with constant parameters (pooled OLS model) - the assumption is that free members and coefficients of slope are constant in space and in time. A errors in model include individual and the time differences (Gujarati, 2004).

- model with fixed effects (fixed-effects model) - coefficients of slope are constant, but the free members vary according to observation units and / or in time (Gujarati, 2004).

- model with stochastic effects (random-effects model) - free members and the coefficients of slope vary according to observation units and through time.

When testing the variability of regression parameters determines the existence of individual or timed effects, raises the question of the selection of model of panel data. In the case where the data relate to a few geographic regions, industries etc. suggest the fixed effects model. When the units are randomly selected from large basic set, then is used model of stochastic effects. When concluding is conditioned only by the individual characteristics of the sample, than is used model of fixed effects. Before selecting of a linear model in econometric panel analysis, it is necessary to test the heterogeneity of regression parameters, or the significance of individual and time effects. Specifically, the first model is estimated with the restriction (pooling OLS), and then model without limitation (fixed effect, random effect).

The aim of the research is to analyse the individual market factors, examine their interdependence, and establish strength of linkage. In this case study the goal is to examine the influence of independent variables (placement on the domestic market, export and sales prices of non-timber forest products - NTFPs) on the dependent variable - the purchase of raw material.

MATERIAL AND METHODS

The research covered an area of Central Serbia and all related data the purchase, placement on the domestic market, export and selling prices of certain types of NTFPs. Data were collected in the 30 companies that are aimed at marketing of NTFPs, and operate in the specified area. The majority of companies (63) are situated in statistical region of Central Serbia (Keča, 2016). All of the companies are private owned, having more than 10 permanent workers, and conduct placement on domestic and foreign market. The period 2006-2013 was considered.

In this case analysis based on the panel series, compared to other methods has been applied, because of the advantages that are reflected in the fact that it does not require long time series, as is the case for e.g. regression analysis, that the results would be authentic and precise. We did not have the time series longer than 5 years. Therefore, the analysis based panel series is estimated to be adequate in the applied study. On the other hand, this methodological approach is very prominent when, generally, comes to market research (Hsiao, 2003, Wooldridge, 2010).

Variables that are taken into consideration (buying, selling on the domestic market, export and selling prices of certain types of NTFPs) are significant, if we analyze the market of this product category. The starting assumption of the research is that, compared to other, placement sales on the domestic market has the greatest impact on the amount of the purchase of raw materials. This is reflected in the assumption that the increase in placements on the domestic market leads to increased demand for raw materials companies, and therefore a quantitative increase in the purchase of raw materials (in order to ensure higher production and sales at higher market prices).

In the analysis were used R Program (Dessau and Pippner, 2008) and the special package, which is dealing with panel data (Chen, 2013).

Advantages of panel analysis (Klevmarken, 1989, Ripley, 2001) in relation to the others are reflected in the following:

- Panel data contains more data than the corresponding spatial data and time series,
- Dependent variables are changing through the units of observation and time, and obtained estimations derived through panel analysis are more precise,
- Panel data reduced the bias of parameters which occurs due to lack of data,
- Panel data allow defining and testing the complicated econometric models,
- Panel data reduced the problem of multicollinearity (Graham, 2003),
- Allow measuring diversity within the observation unit.

Series of panel data are a combination of comparative data and time series. Each panel has two dimensions structure and time, and where in the econometric analysis of panel data most used are linear models (model of panel data).

For the analysis of panel were identified through variables such as:

- amount of the purchase of raw material per year (t)
- amounts placed on the domestic market (t)
- quantities exported (t)
- the prices at which the products realized on the market (RSD).

Products that have been analyzed for different categories of NTFPs covered the period 2006-2013. Thus, the surveys involve forest fruit (blueberries, juniper, wild rose, raspberry, blackberry and wild strawberries), herbs and mushrooms (boletus and chanterelles). Since the analysis provided data types of products in all units of time, it is a "balanced" or complete panel series. In case the missing data for a variable, it would be the so-called "unbalanced" panel series (Croissant and Millo, 2008).

RESULTS AND DISCUSSION

The first part shows the results of descriptive statistics relating to the dependent variable (the purchase of raw materials) and analyzed independent variables, such as placement on the domestic market, exports and prices at which the products are realized in the market (Table 1). The reason for the separation of purchase as the dependent variable is reflected in the fact that sought to determine the strength of the influence of factors such as market prices, amount of investments into the domestic market and exports (independent variable) on the amount of purchased raw materials in analysed time interval. Raw material is taken as a starting unit of observation, because it preceded, as sales in the domestic market and the export of final products. In subsequent stages, it has been substituted the variables, so that every independent variable would be analyzed as a dependent, with the aim of establishing their interdependence in all relations.

Table 1. Descriptive statistics for analyzed variables

Purchase	Placement	Export	Prices
Min.: 5.0	Min.: 2.0	Min.: 1.0	Min.: 40.0
1st Qu.: 92.5	1st Qu.: 5.0	1st Qu.: 42.5	1st Qu.: 200.0
Median: 200.0	Median: 32.0	Median: 141.0	Median: 400.0
Mean: 311.9	Mean: 93.7	Mean: 220.5	Mean: 629.9
3rd Qu.: 283.0	3rd Qu.: 99.5	3rd Qu.: 234.5	3rd Qu.: 542.5
Max.: 1200.0	Max.: 600.0	Max.: 858.0	Max.: 2600.0

Results of analysis according to individual estimators are given as follows:

Table 2. Results according to Pooling OLS estimators

Coefficients :	Estimate	Std. Error	t-value	Pr(> t)	Sign.
(Intercept)	3.5521434	6.4358014	0.5519	0.5837	
Xplacement	1.0408989	0.0455480	22.8528	<2e-16	***
Xexport	0.9589093	0.0353677	27.1125	<2e-16	***
Xprices	0.0010241	0.0108347	-0.0945	0.9251	
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					
Total Sum of Squares:			6062500		
Residual Sum of Squares:			37042		
R-Squared			0.99389		
Adj. R-Squared :			0.91438		
F-statistic: 2494.23			2494.23		
p-value:			< 2.22e-16		
Residuals :					
Min.	1st Qu.	Median	3rd Qu.	Max.	
-125.00	-3.29	-1.47	4.69	98.40	

Based on Pooling OLS estimators determined that the variable export and sales on the domestic market are important to purchase of NTFPs (***), while the variable price of this estimator, there is no impact on the amount of purchase NTFPs. There was a strong correlation (Adj., R-squared = 0.914). The coefficient of correlation was statistically significant (F = 2494.2), as well as parameters (p value <2.22e-16) (Table 2).

Table 3. Results according "between" estimators

Table 8: Results according to each estimator					
Coefficients :	Estimate	Std. Error	t-value	Pr(> t)	Sign.
(Intercept)	2.02826298	1.80546636	1.1234	0.3042	
Xplacement	1.03013182	0.01545503	66.6535	7.671e-10	***
Xexport	0.96856851	0.01227245	78.9222	2.786e-10	***
Xprices	-0.00038405	0.00357620	-0.1074	0.9180	
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					
Total Sum of Squares:			1190500		
Residual Sum of Squares:			70.337		
R-Squared			0.99994		
Adj. R-Squared :			0.59996		
F-statistic:			33850		
p-value:			4.511e-13		
Residuals :					
Min.	1st Qu.	Median	3rd Qu.	Max.	
-3.030	-1.920	-0.919	1.500	4.540	

By analyzing the variations between objects of observation (between) it is established, as the previous estimator that the variables export and prices are important to the purchase of NTFPs (***), and a variable price, according to the estimator, there is no impact on the amount bathe NTFPs. There was a strong correlation (Adj., R-squared = 0.999). The coefficient of correlation was

statistically significant ($F = 3.3850$), as well as parameters (p -value $4.511e-13$) (Table 3).

Table 4. Results according to First differences estimators

Coefficients :	Estimate	Std. Error	t-value	Pr(> t)	Sign.
(Intercept)	1.874016	5.027649	0.3727	0.71153	
Xplacement	0.804902	0.166348	4.8387	2.457e-05	***
Xexport	0.773047	0.079922	9.6726	1.500e-11	***
Xprices	-0.065969	0.033299	-1.9811	0.05526	
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					
Total Sum of Squares:			129170		
Residual Sum of Squares:			35277		
R-Squared			0.72688		
Adj. R-Squared :			0.65419		
F-statistic: 2494.23			31.9371		
p-value:			3.0077e-10		
Residuals:					
Min.	1st Qu.	Median	3rd Qu.	Max.	
-141.000	-5.810	0.105	7.980	78.800	

According to First difference estimator export and sales price are the factors that most influence the purchase of raw NTFPs (***). Also, the price and in the case of previous estimators indicate that they do not affect the amount of the purchase NTFPs (Table 4).

The relevance based on the F test suggests that in choosing the estimator should take into account the model "fixed effect

Table 5. Results according to F test for individual effects

Products	Year	Byuing	Placement	Export	Prices
Length:50	Min. :2006	Min. : 5.0	Min. : 2.0	Min. : 1.0	Min. : 40.0
Class :character	1st Qu.:2007	1st Qu.: 92.5	1st Qu.: 5.0	1st Qu.: 42.5	1st Qu.: 200.0
Mode :character	Median :2011	Median : 200.0	Median : 32.0	Median :141.0	Median : 400.0
	Mean :2010	Mean : 311.9	Mean : 93.7	Mean :220.5	Mean : 629.9
	3rd Qu.:2012	3rd Qu.: 283.0	3rd Qu.: 99.5	3rd Qu.:234.5	3rd Qu.: 542.5
	Max. :2013	Max. :1200.0	Max. :600.0	Max. :858.0	Max. :2600.0
Byuing of raw NTFPs	Pooled OLS regression	Between		Within or fixed effects	First differences
Placement on domestic market	1.0408989	1.03013182			0.804902
Export	0.9589093	0.96856851			0.773047
Prices	0.0010241	-0.00038405			-0.065969

According to estimators: pooled OLS; between; first differences, it was found that the placement on the domestic market variable, which, compared with other, largely determines the purchase of raw NTFPs (in terms of the analyzed data). This is evident based on the greatest values that range in interval 0.80 to 1.04 for the domestic market (table 5) ($f = 126.16$, $df1 = 11$, $df2 = 37$, $p\text{-value} < 2.2e-16$ alternative hypothesis: significant effects). The second by relevance is a variable of export of NTFPs, measured by all three estimators. Prices were evaluated as the least important factor as measured by the three estimators.

CONCLUSIONS

Based on the research, which referred to NTFPs market, was determined the effectiveness of applying analysis based on panel series, particularly bearing in mind that it was not available long time series. In addition, in this way, we obtained two-dimensional, that is, structural and time comparability among the data within each panel observations. The subject of the analysis were quantitative market factors related to the quantity of raw materials purchase (the dependent variable), placement to the domestic market, exports and resilient price of final products (independent variables). Starting from the purchase of raw materials as the dependent variable, and on the basis of the level of influence of independent variables obtained according to estimators Pooling OLS, Between summary i First differences it was found that placement on the domestic market has the greatest impact on the amount of the purchase of raw NTFPs. Thus is achieved affirmativeness of starting hypothesis that precisely placement on the domestic market largely determines the amount of the purchase of raw materials. The establishment of this hypothetical starting point was that it sought to examine whether increase in placements in the local market causes and increased purchase of raw materials, while the prices and quantities of exports less important factors, if we consider the substantive effect of the independent variable on the dependent one.

ACKNOWLEDGMENTS

Gratitude for the realization of this study the authors suggest the Ministry of Education and Science of the Republic of Serbia which is financially supported this research within the project "Sustainable management of the total resources of forests in the Republic of Serbia" and "Forest plantations in order to increase afforestation of Serbia".

REFERENCES

- Baltagi B. (2008): *Econometric analysis of panel data*. John Wiley & Sons (388)
- Chen Y.M. (2013): *Panel Data Analysis*. Department of Finance. National Chung Hsing University
- Croissant Y, Milla G (2008); *Panel Data Econometrics in R: The plm Package*. "Journal of Statistical Software", 27(2), 1-43.
- Dessau R. B., Pipper C. B. (2008) : ["R"--project for statistical computing]. Ugeskrift for laeger, 170(5), (328-330)

- Dragutinović-Mitrović R. (2001): Analiza panel serija. Zadužbina Andrejević, Beograd
- Graham M. H. (2003): Confronting multicollinearity in ecological multiple regression. *Ecology*, 84(11), (2809-2815)
- Greene W. (2005): Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. *Journal of Econometrics*, 126(2), (269-303)
- Gujarati D.N. (2004): Basic Econometrics, The McGraw-Hill Companies, New York, 4th Edition.
- Hsiao C. (2003): Analysis of panel data. Cambridge: Cambridge University Press.
- Jović M. (2010): Ekonometrija 1D – Uvodni deo, dostupno na <http://ekonometrija.ekof.bg.ac.rs/doktorske%20studije/1D%20I%20deo/Ekonometrija%201D%20-%20Uvodni%20deo.pdf> (posećeno 15.06.2015)
- Katchova A. (2011): Repayment Capacity of Farmers: A Balanced Panel Data Approach, *The Journal of Applied Economics and Policy*, 30(1): 14-30
- Keča Lj. (2016): Value chain analysis and socio-economic aspects of non-wood forest products in Central Serbia, *Proceedings of International Symposium*, 9-11 May, 2016, Vienna, Austria (39-42)
- Kennedy P.A. (2008): Guide to Econometrics, 6th Edition, Blackwell Publishing, Carlton
- Klevmarken N.A. (1989): Panel studies: What can we learn from them? Introduction, *European Economic Review*, vol. 33, pp. 523–529.
- Račić Ž. (2013): Komparativna analiza uticaja veličine banaka na izloženost HPR-a kamatnom riziku - razvijeno finansijsko tržište vs. finansijsko tržište Republike Srbije. *Škola biznisa*, br. 1, (1-11).
- Ripley B. D. (2001): The R project in statistical computing. *MSOR Connections*. The newsletter of the LTSN Maths, Stats & OR Network, 1(1), (23-25)
- Wooldridge J. M. (2010): Econometric analysis of cross section and panel data. MIT press (754).