International Journal of Research in Engineering and Science (IJRES)
ISSN (Online): 2320-9364, ISSN (Print): 2320-9356
www.ijres.org Volume 11 Issue 9 ǁ September 2023 ǁ PP. 110-114

Examination of the factors affecting apple yield in Benalux countries with Panel Data Analysis

Şenol Çelik
Biometry and Genetic Unit, Department of Animal Science, Faculty of Agriculture, Bingol University, Bingol 12000, Turkey
Correspondence: senolcelik@bingol.edu.tr

Abstract: *The aim of this study is to test the relationship between harvest area and production amount and apple yield using panel data analysis. For this purpose, Belgium, Luxembourg and the Netherlands countries were discussed and the period of 2000-2021 was analyzed with panel data method. The empirical findings revealed that the harvested area affected the apple yield negatively and the production amount positively affected the yield.*

Keywords: *Apple, Benalux countries, panel data analysis*

Date of Submission: 07-09.2023 Date of acceptance: 20-09-2023

I. INTRODUCTION

The Benelux countries are three neighboring countries in Northwestern Europe: Belgium, the Netherlands, and Luxembourg. The official languages of the Benelux and its institutions are Dutch and French. 82% live in the Netherlands or Belgium's Flemish region, where Dutch is the official language. The French-speaking Walloon Region has 12,5%, while the Brussels Capital Region (3,8%; legally bilingual) also has a Francophone majority. Luxembourg (1,7% of the population) is officially trilingual, with Luxembourgish (a German dialect) serving as the national language. When combined with Belgium's German-speaking community, this amounts to an estimated 2% of native German speakers. The Benelux Economic Union Treaty (signed in 1958) promotes the free movement of workers, capital, services, and goods in the region. The major goals of a new treaty (2010) are to continue and expand the three member nations' political and official collaboration within a larger European environment [1].

According to [2], the policy-making procedures of this group diverge from those of the rest of the EU. Their policies are strongly pro-export, and their economies are primarily focused on producing export commodities, in which these countries achieve either absolute or relative competitive advantage and are thus constantly capable of positioning their production in the global market. [3] identify similar distinctions. The authors discovered that the Visegrad countries prioritize mutual support, solidarity, and assistance in their collaboration, whereas the BENELUX countries prioritize collective competitiveness in the global market. In order to conduct a more in-depth examination, [4] examined the growth of the gross domestic product.

The economics of the Benelux countries evolved substantially between the 1960s and the 1990s. In Belgium, the wealth balance shifted from south to north, with the Flemish-speaking provinces becoming the new political and economic core. Meanwhile, the Walloon South's mature industries stagnated, and coal mining ceased entirely. Luxembourg grew as a financial hub, attracting a significant immigrant population, and consolidating its position as the home of several large European institutions. The Netherlands experienced a brief era of manufacturing prosperity before returning to its roots as a primarily service-sector economy. Simultaneously, Dutch agriculture grew and consolidated, becoming not only one of Europe's most productive food exporters, but also one of its most heavily capitalized [5].

The aim of this study is to examine and evaluate the effect of factors affecting apple yield in Belgium, Netherlands and Luxembourg by panel data analysis.

II. MATERIAL AND METHOD

The data belonging to the study are apples belonging to 3 Benelux countries, Belgium, Netherlands and Luxembourg, for the period 2000-2021 under the title of “Crops and livestock products” on the www.fao.org website of FAO (Food and Agriculture Organization of the United Nations) yield (100 g/ha), harvested area (ha) and production amount (tons) [6].

A panel data analysis with the fixed-effect (FE) model was conducted using ordinary least squares (OLS) estimation, with standard errors for clustering at the country level as described in Equation (1) [7].

$$y_{it} = a_i + \beta_1 X_{it} + v_i + \mu_{it} \quad (1)$$

where y_{it} indicates the apple yield (outcome) in the i-th country in the t-th year; xit denotes explanatory variables for the i-th country and all covariates in the t-th year; β stands for the coefficient of each variable; νi indicates the fixed effects of the i-th country; αre presents the intercept; and μ_{it} is the error term of the i-th country in the t-th year. The regression models included all time-varying confounders. e. g.,

$$yield_{it} = a_i + \beta_1 harvest_{1it} + \beta_2 product_{2it} + v_i + \mu_{it} \quad (2)$$

Unlike the fixed effects model, the variation between entities in the random effects (RE) model is believed to be random and uncorrelated with the predictor or independent variables included in the model: The key distinction between fixed and random effects is whether the unobserved individual effect contains elements that are correlated with the model's regressors, not whether these effects are stochastic or not" [8].

The following is a basic equation for the RE model [9]:

$$y_{it} = a_i + \beta_1 X_{it} + v_i + \mu_{it}$$

Where, y_{it} is the dependent variable, X_{it} is the independent variable. It must be exogenous; the coefficient for the independent variable is β_1. The unobserved entity-specific time-constant error term is denoted by v_i. In contrast to the fixed effect, it is believed to be uncorrelated with X_{it}. The incorrect phrase is μ_{it}. This, like the fixed effect, is thought to be unrelated to X_{it}.

The choice between FE and FE is determined by whether v_i is associated with any of the model's other explanatory variables [10]. When there is such a correlation, the fixed effect strategy is superior. Otherwise, the random effect is more frugal and yields more accurate estimates [10]. The [11] specification test is a formal technique for analyzing the relationship between unobserved heterogeneity and other explanatory variables. Tests of cross-sectional dependence consist of Pesaran's CD test, Friedman's test and Frees' test [12].

Robust Estimation of Linear Fixed Effects Panel Data Models

Three types of outliers can lead least squares to fail in cross-sectional regression analysis. Vertical outliers, bad leverage points, and good leverage points are defined by [13]. Vertical outliers are observations that deviate from the mean in the y-dimension but not from the mean in the space of the explanatory factors (x-variables). Their presence has an effect on both the estimation of the intercept and the estimation of the regression coefficients, but the latter has a lesser effect. Bad leverage points are observations that are both out of the explanatory variable space and far away from the regression line. They have a significant impact on the calculation of both the intercept and slope coefficients. Finally, good leverage points are observations that are far from the explanatory variable space but are close to the regression line. Their presence has a minor impact on the estimate of both the intercept and the regression coefficients, but it does have an impact on inference. When working with Panel data, a fourth type of outlier should be considered: block concentrated outliers, which occur when the majority of outlying observations are concentrated in a small number of time series [14]. [14] suggest two estimators, the Within Groups Generalized M-estimator (WGM) and the Within Groups MS-estimator (WMS), to deal with the presence of any of these categories of outliers.

The equation estimate

$$yield_{it} = \alpha_i + \beta_1 harvest_{1it} + \beta_2 product_{2it} + \varepsilon_{it}$$

is in the form.

III. RESULTS AND DISCUSSION

The findings obtained as a result of the fixed effects model are summarized in Table 1 below.

Table 1. Fixed effects model regression results

yield	Coefficient	St. Err.	t-value	p-value	[95% Conf	Interval]	Sig
harvest	-0.017	0.001	-16.06	0.001	-0.019	-0.015	***
product	0.025	0.001	24.99	0.001	0.023	0.027	***
Constant	5305.055	242.276	21.90	0.001	4820.595	5789.516	***

Mean dependent var	5537.939	SD dependent var	3848.142
R-squared	0.911	Number of obs	66
F-test	312.259	Prob > F	0.000
Akaike crit. (AIC)	975.010	Bayesian crit. (BIC)	981.579

**** p<0.001*

The relationship between the variables discussed in Table 1 was estimated with the fixed effects model. According to the results obtained, the F test, which expresses the general significance of the model, shows that the model is significant. The R^2 value, which shows the explanatory power of the model, was calculated as approximately 91.1%. The remaining disclosure rate is explained by the variables not included in the model. When the significance levels of the variables are examined, the harvest and product variables and the constant term are statistically significant (p<0.001).

After the fixed effects model was estimated, the relationship between the variables was estimated with the help of the random effects model. The random effects model estimation results are summarized in Table 2 below.

Table 2. Random effects model

yield	Coefficient	St. Err.	t-value	p-value	[95% Conf	Interval]	Sig
harvest	-0.008	0.002	-3.64	0.001	-0.019	-0.015	***
product	0.030	0.002	13.85	0.001	0.023	0.027	***
Constant	1521.402	295.977	5.14	0.001	4820.595	5789.516	***

R-squared	0.722	Number of obs	66
Wald-test	985.05	Prob > chi2	0.000

*** *p<0.001*

The relationship between the variables discussed in Table 2 was estimated by the random effects model, and the Wald test, which expresses the general significance of the model, showed that the model was significant. The R^2 value, which shows the explanatory power of the model, was calculated as 72.2%. In other words, the independent variables together explain about 72.2% of the total change in apple yield. The remaining disclosure rate is explained by the variables not included in the model. When the significance levels of the variables are examined, the harvest, product variables and the constant term are statistically significant ($p<0.001$).

After estimating the fixed effects and random effects model, it is necessary to choose between these two estimators. With the Hausman test, a choice is made between two estimators and the Hausman test results are given in Table 3.

Table 3: Hausman Test

Dependent variable: yield	
Chi2(2)	1174.20
Prob>chi2	0.0001

According to the Hausman test result given in Table 3, the fixed effects estimator gives effective and consistent results. So, as a result, the fixed effects model will be preferred to the random effects model.
H_0: Random effects estimator enabled
H_1: Fixed effects estimator enabled
As a result of the Hausman test, it was determined that the fixed effects model gave more effective and consistent results. However, in the fixed effects model, basic assumption tests must be made and deviations from the assumption must be checked. Changing variance, autocorrelation and inter-unit correlation problems also negatively affect the fixed effects model results. In case of deviation from these three assumptions, the results to be obtained will be biased and inconsistent. For this reason, in the fixed effects model, it is necessary to control the deviations from the assumption, and if there are deviations from the assumption, resistant estimators should be used, which give resistant results against the deviations from the assumption. First, the variable variance situation in the fixed effects model will be examined by Levene, Brown and Forsythe's varying variance test. The test results obtained are given in the table below (Table 4).

Table 4. Levene, Brown, and Forsythe Variance Test of Variance

Statistics	Value	Prob>F
W0	4.1369643	0.02050739
W50	4.1070777	0.02105659
W10	4.1829838	0.01969046

df(2,63)

As can be seen in Table 4, three test statistics were given as a result of Levene, Brown and Forsythe's test for the fixed effects model. The main hypothesis here is that there is no varying variance in the fixed effects model. According to the test results obtained, the H_0 basic hypothesis is rejected. In other words, there is a problem of varying variance in the fixed effects model.

In the fixed effects model, the autocorrelation status was examined with the Serrial Correlation test, and the test results obtained are summarized in Table 5.

Table 5. Serial Correlation Test

Statistics	Value	P>chi2(1)
LM (Lambda=0)	7.14	0.0076
ALM (Lambda=0)	4.90	0.0269

In Table 5, the autocorrelation status in the fixed effects estimator was examined with the Serrial Correlation test statistical values. According to the results obtained, since $p<0.05$, there is an autocorrelation problem in the fixed effects model.

Inter-unit correlation was checked with the help of Pesaran, Friedman, and Frees's tests. The results obtained are summarized in Table 6.

Table 6. Correlation between units

Tests	Value	$p>\chi^2$
Pesaran	-0.946	1.6557
Friedman	18.257	0.0001
Frees Alpha=0.01 Alpha=0.05 Alpha=0.10	0.338 0.1174 0.1537 0.2225	

In Table 6, the inter-unit correlation in the fixed effects estimator was examined with the tests of Pesaran, Friedman, and Frees. Except for the Pesaran test, other test results support each other. For the tests, the H_0 basic hypothesis was established as "There is no correlation between units". According to the test results, the H_0 basic hypothesis was rejected for the Friedman and Frees test. In other words, there is a correlation between units in the fixed effects model.

As a result, there were deviations from the assumption of varying variance, autocorrelation and correlation between units in the fixed effects model made for the econometric model. Deviations from this assumption cause the results of the fixed effects model to be biased. For this reason, robust estimators that take into account deviations from the assumption should be applied. In case of varying variance, autocorrelation and correlation between units in panel data analysis, it is appropriate to use the Parks-Kmenta estimator. Resistive estimator results are presented in Table 7.

Table 7. PARKS-KMENTA Estimator

```
Random-effects GLS regression                   Number of obs      =        66
Group variable: id                              Number of groups   =         3

R-sq:                                           Obs per group:
     within  = 0.7218                                         min =         22
     between = 0.9989                                         avg =       22.0
     overall = 0.9399                                         max =         22

                                                Wald chi2(2)       =    280.23
corr(u_i, X)   = 0 (assumed)                    Prob > chi2        =    0.0000

                                    (Std. Err. adjusted for 3 clusters in id)
------------------------------------------------------------------------------
             |               Robust
       yield |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     harvest |  -.0082608   .0021884    -3.77   0.000    -.0125499   -.0039716
     product |   .0302029   .0018778    16.08   0.000     .0265225    .0338833
       _cons |   1521.402   787.2815     1.93   0.053    -21.64178    3064.445
-------------+----------------------------------------------------------------
     sigma_u |          0
     sigma_e |  388.16804
         rho |          0   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

In Table 7, the results of Parks-Kmenta resistant estimators that give resistant results in case of autocorrelation, varying variance and inter-unit correlation are summarized. Considering the results obtained, the model was estimated statistically significant according to the Wald test result ($p<0.001$). In the resistant panel data analysis, harvested area (ha) had a negative effect on apple yield, while production amount (tonnes) had a positive effect on apple yield.

[15] developed two models, a fixed-effect model and a dynamic panel model, to assess the determinants associated with overall health expenditure increase and its key components for 167 countries from 1993 to 2013. Overall, the findings revealed that total health expenditure per capita is rising in all nations over time as incomes rise. However, estimations showed that the income elasticity of health expenditure ranged from 0.75 to 0.96 in the fixed-effect static panel model, but it was lesser and ranged from 0.16 to 0.47 in the dynamic panel model.

[16] evaluated the competitiveness of national economies in the BENELUX (Belgium, the Netherlands, and Luxembourg) economic union. Because of their united economies, innovative potential, industrial maturity, technical preparedness, and a high-quality educational system, the BENELUX countries benefited from their competitiveness. The BENELUX countries consistently ranked well in terms of competitiveness, owing to systematic efforts to improve their own competitive qualities based on a well-thought-out economic plan, education, and institutional foundation.

IV. CONCLUSION

In this study, the relationship between apple yield, harvested area and production amount in 3 Benallux countries was investigated. Fixed-effects and random-effects models were tried and fixed-effects model was preferred. For the fixed-effects model, none of the assumptions of varying variance, autocorrelation and inter-unit correlation were met. For this reason, the Parks-Kmenta resistive estimator model was examined. As a result of the empirical findings obtained as a result of the resistant panel data analysis, the apple yield will decrease as the harvested area increases. As the harvested area increases by 1 unit, apple yield will decrease by 0.0083 units. As the production amount increases by 1 unit, apple yield will increase by 0.03 units. While the relationship between apple yield and harvested area was negative, the relationship between production amount and apple yield was positive, and these relationships were found to be statistically significant ($p<0.001$).

REFERENCES

[1]. Steffgen, G., Vandebosch, H., Völlink, T., Deboutte, G., Dehue, F. 2010. Cyberbullying in the Benelux-Countries: First findings and ways to address the problem. En J.A. Mora-Merchány T. Jäger (Eds.), Cyberbullying. A cross-national comparison (pp. 35-54). Landau: Verlag Empirische Pädagogik.

[2]. Kisluhin, V.A. 2012. Police of the Benelux countries: positive experience demands studying. Legal science and practice-Bulletin of Nizhniy Novgorod Academy of the Ministry if the Interior of Russia, 18(2): 79-82.

[3]. Drynochkin, A., Sergeev, E. 2016. Benelux and Visegrad countries: comparative analysis. Contemporary Europe-Sovremennaya Evropa, 6: 53-62. https://doi.org/10.15211/soveurope620165362

[4]. Cozanet, E. 2014. GDP of Denmark, the Benelux area. Biofutur, 351: 41

[5]. Jones, E. 2003. The Benelux Countries: Identity and Self-Interest. Simon Bulmer and Christian Lequesne, eds. Member States and the European Union Oxford: Oxford University Press.

[6]. FAO, 2023. Crops and livestock products. Food and Agriculture Organization of the United Nations. https://www.fao.org/faostat/en/#data/QCL (17 August 2023).

[7]. Abe K, Taniguchi Y, Kawachi I, Watanabe T, Tamiya N. 2021. Municipal long-term care work force supply and in-home deaths at the end of life: Panel data analysis with a fixed-effect model in Japan. Geriatr. Gerontol. Int., 21: 712–717. https://doi.org/10.1111/ggi.14200

[8]. Greene, W.H. 2008. Econometric Analysis, sixth edition, Upper Saddle River, NJ, Prentice-Hall.

[9]. Baltagi, B.H. 2001. Econometric Analysis of Panel Data, 2nd editions, New York, John Wiley.

[10]. Woolridge, J.M. 2002. Econometric Analysis of Cross Section and Panel Data. The MIT Press Cambridge, Massachusetts London, England.

[11]. Hausman, J.A. 1978. Specification Tests in Econometrics. Econometrica, 46(6): 1251–1271.

[12]. De Hoyos, R.E., Sarafidis, V. 2006. Testing for cross-sectional dependence in panel-data models. Stata Journal, 6(4): 482-496.

[13]. Rousseeuw, P.J., Leroy, A.M. 1987. Robust Regression and Outlier Detection. New York etc.: Wiley.

[14]. Bramati, M. C., Christophe, C. 2007. Robust Estimators for the Fixed Effects Panel Data Model. Econometrics Journal, 10(3): 521-540.

[15]. Younsi, M., Chakroun, M., Nafla, A. 2016. Robust analysis of the determinants of healthcare expenditure growth: Evidence from panel data for low, middle and high income countries. The International Journal of Health Planning and Management, 31: 580-601. DOI: 10.1002/hpm.2358.

[16]. Koraus, A., Mazak, M., Dobrovic, J. 2018. Quantitative Analysis of the Competitiveness of Benelux Countries. The International Journal Entrepreneurship and Sustainability Issues, 5(4): 1069-1083. http://doi.org/10.9770/jesi.2018.5.4(26).