Zagreb International Review of Economics & Business, Vol. 24, Special Conference Issue, pp. 27-47, 2021

ISSN 1331-5609; UDC: 33+65
DOI: 10.2478/zireb-2021-0018
CONFERENCE PAPER

Performance Evaluation of the Fruit and Vegetable Subsectors in the Azerbaijani Economy: a Combinatorial Analysis Using Regression and Principal Component Analysis

Ibrahim Niftiyev *

Abstract: *Azerbaijan has an oil-led economy, which according to the well-known resource curse and Dutch disease hypotheses decreases the role of non-oil tradable sectors. Nevertheless, the government has actively fostered the growth of non-oil tradable sectors as the export orientation of Azerbaijan is being leveraged by the recently adopted economic policies. However, performance evaluations at the subsectoral level remain rare. The present paper evaluates the performance of the fruit and vegetable subsectors in Azerbaijan from 1995 to 2020 based on multiple key indicators, such as production, profitability, and productivity via principle component analysis (PCA). The purpose of the study was to provide a comparison of two key subsectors in Azerbaijan that are strong candidates for non-oil tradable exports. The results revealed that the vegetable subsector outperformed the fruit subsector in terms of production and profitability from 1999 to 2014; however, it experienced a sharp decline from 2014 to 2015 (the period of the rapid commodity price downturns), which gives rise to the question of whether the extractive industry negatively affected the subsector. Compared to the vegetable subsector, production and profitability in the fruit subsector demonstrated a more stable upward trend. In addition, labor input in both subsectors decreased over time, indicating efficiency gains via new technology transfers and productivity enhancements. Ordinary Least Squares (OLS) results demonstrated a strong and statistically significant negative relationship between the performance of the vegetable subsector with the oil revenue boom period (2008–2015).*

Keywords: Azerbaijan economy; agriculture; subsectoral performance; vegetable production; fruit production

JEL Classification: E01, C38, O13, Q11, Q18

* Ibrahim Niftiyev is a PhD candidate at the University of Szeged, Doctoral School in Economics, Institute of Finance and International Economic Relations (6722, Kálvária sgt. 1) and a researcher at the Centre for Studies on European Economy (AIM) - Azerbaijan State University of Economics (UNEC, AZ 1001, Istiglaliyyet str. 6), email: niftiyev@eco.u-szeged.hu, phone: +994507415989 and +36203975435

Introduction

As a small, oil-driven, and developing country, Azerbaijan's long-term sustainable economic development depends on the development of non-resource tradable sectors. However, since 1994, Azerbaijan's economic growth has consistently depended on the oil sector. Therefore, the risk of oil dependence has become more apparent. After the popularization of the "resource curse" (a term coined by Auty in 1993) and the Dutch disease phenomenon (The Economist, 1976), many resource-rich countries became testing grounds to assess whether or not these theories are appropriate to explain the structural changes. Various studies have examined the relevance of the resource curse and Dutch disease models in Azerbaijan, and there has been an upward trend in the literature over the past few years (Bayramov and Conway, 2010; Ibadoghlu, 2012; Ibadoghlu et al., 2013; Ibadoglu, 2008; Niftiyev, 2020a). An accurate diagnosis of the resource curse and the Dutch disease is important, because appropriate policies can correct for policy failures in the presence of these phenomena. However, the economic structure of Azerbaijan and the increasing number of studies on the country's economy necessitates a narrower focus on, for example, sectors or subsectors scope in order to evaluate their performance and understand whether or not their efficiency or productivity levels allow for the diversification of the national economy.

An extractive, industry-led economy poses several challenges if they are not addressed in a timely manner. One of them is diversification. Compared to neighboring post-Soviet countries such as Russia and Kazakhstan, Azerbaijan shows low levels of diversification (Hamidova, 2020). Thus, Azerbaijan is considered one of the most oil-dependent countries in the world (Czech, 2018). Bayramov and Abbas (2017) reported low diversification of exports in Azerbaijan, Russia, and Kazakhstan; apart from the mineral sector, the rest of the economy is heavily dependent on government spending. Other alarming characteristics of an undiversified economy include windfall-financed government expenditure and subsidy dependence in non-oil sectors. In fact, after oil prices collapse, the country's gross domestic product (GDP) declines alongside the output of non-oil sectors (Zulfigarov and Neuenkirch, 2020). This demonstrates the cyclical nature of the economy with regard to oil variables (Niftiyev and Namazova, 2020), which also can be observed among other resource-rich post-Soviet economies (Niftiyev, 2020b).

The economic potential and long-term viability of non-oil sectors in Azerbaijan raise multiple questions in terms of future development prospects. However, it is also useful to evaluate past performance in order to predict the future economic viability of non-resource sectors. A recent analysis by Pashayev and Aliyev (2020) stated that the ability of state expenditures to cover the budgetary costs of non-oil sectors is decreasing. This finding highlights the need to understand the capacity, efficiency, and productivity of non-oil tradable sectors in order to design adaptive economic policies. Therefore, the main focus of this study is the fruit and vegetable subsectors of Azerbaijan, which has gained a large share of agricultural exports. Recent studies have also supported the increasing role of fruits and vegetables (i.e., horticulture) as strong non-oil and agricultural exports (Kerimova, 2014; Shalbuzov et al., 2020).

Recently, the literature has also started to capture the growing role of agriculture in the diversification process (see Eldarli, 2018). Agricultural reforms have been an integral part of the transition process. Aslund (2013) evaluated Azerbaijan as a "star performer" in the transformation from a command economy, indicating that the country has successfully applied agricultural policies. However, Azerbaijan has also experienced difficulties. According to Spoor and Visser (2001), shortcomings within agricultural policies were also part of the reforms that aimed to support a smooth transition from a command economy to a market economy. They argued that recommendations and suggestions from international organizations were ill-considered, because such organizations used China and Vietnam as benchmarks. In other words, Azerbaijan-specific policies were lacking.

Nevertheless, agriculture in Azerbaijan is of immense importance, as the sector has the largest share of employment; however, it has low output (Cornell, 2014). Despite the fact that the value added per worker in constant 2010 US dollars has risen in agriculture, forestry, and fishing in recent years, employment in agriculture and its share in GDP has declined since the early 2000s (Niftiyev, 2020c). This may also indicate productivity gains through the application of modern technology in agricultural production and farming, but conclusive evidence on the role of agricultural performance is lacking. In addition, Niftiyev and Czech (2020) argued for the presence of Dutch disease in vegetable exports, meaning its low levels during the oil boom.

Thus, the present study evaluates the performance of the fruit and vegetable subsectors in Azerbaijan based on principal component analysis (PCA). The latter provides valuable information based on large data sets, and relevant studies should also be conducted in the Azerbaijani context in order to build a sophisticated body of literature on subsectoral dynamics. By aggregating the seven main indicators (overall 14, for the details, see "Data and Methodology" section) related to the fruit and vegetable sectors provided in official statistics, it is possible to conduct a comprehensive analysis.

As a multivariate statistical tool, PCA allows the construction of an index to explain the overall performance of certain phenomena as a dimension reduction tool. In this study, the constructed index for the vegetable subsector was used in ordinary least squares (OLS) regression to estimate the association between its performance and that of the extractive industry. The collected data covers the years from 1995 to 2020 for PCA and 1996 to 2020 for OLS. Hence, this study extends the understanding of subsectoral performance evaluation in Azerbaijan, which could guide decision-makers and policymakers through the proposal of a practical methodological strategy for future policy considerations.

Data and Methodology

Economic performance evaluation requires the combination of multiple indicators. In the current paper, subsectors of agriculture, such as fruits and vegetables, were analyzed through the principal components produced from seven variables: profitability, production, production per capita, yield, sown, or cultivated area, labor input, and production cost. This approach is experimental in nature; Niftiyev (2021) reported the initial results and slightly different PCA methods regarding the fruit and vegetable sectors.

The data source is the agriculture section of the State Statistical Committee of the Republic of Azerbaijan's (SSCRA) official web resource (SSCRA - The State Statistical Committee of the Republic of Azerbaijan, 2021). The data set did not contain any missing values. However, in order to obtain a wider and more up-to-date time period, values for 2020 were predicted using the TREND function of the Google Sheets online application. Using linear least squares, the TREND function predicts a data point based on partial data. Values for 2020 are based on data from the main time range in the data set (1995–2019).

Before the current study, the working paper by Niftiyev (2021) identified several outliers among variables related to production, production per capita, labor input, and sown and cultivated area in the fruit and vegetable subsectors. Since PCA is highly sensitive to outliers in the data set (Jolliffe and Cadima, 2016), the outliers were replaced using the Winsorization method suggested by Kwak and Kim (2017) (for more details, see Kwak and Kim (2017, p. 410)). It should be noted that Winsorized data was only analyzed within the PCA. Descriptive and correlation analysis included the outliers.

The working paper by Niftiyev (2021) provided the Gaussian and box plot distributions for the collected data to visually identify whether or not the data was normally distributed. However, the current paper presents the results of a formal normality test, namely the Shapiro-Wilk normality test. Table 1 provides variable explanations, their levels of measurement, and the results of the Shapiro-Wilk test. Two variable categories—fruits and vegetables—contained seven identical variables of interest. According to the results of the Shapiro-Wilk test, five out of 14 overall variables showed a normal distribution.

Table 1: Variable Names, Descriptions, Measurements, and Shapiro-Wilk Test Results (1995–2020)

			Shapiro-Wilk test	
Variable	**Indicator**	**Measurement**	**Statistic**	**Sig.**
VEG_PROF	Profitability of vegetable products	%	0.976	0.781
VEG_PROD	Production of vegetable products	thousand tons	0.930	0.077
VEG_PROD_PC	Per capita production of vegetable products	kg	0.879	0.005
VEG_YIELD	Yield of vegetable products	100 kg/ha	0.907	0.022
VEG_SOWN	Sown area of vegetable products	ha	0.785	0.000
VEG_LI	Labor input of vegetable products	person/hour per 100 kg of vegetables in agricultural enterprises	0.812	0.000
VEG_PCOST	Production cost of vegetable products	in AZN, per 100 kg of vegetables in agricultural enterprises	0.908	0.024
FR_PROF	Profitability of fruit products	%	0.955	0.299
FR_PROD	Production of fruit products	thousand tons	0.959	0.381
FR_PROD_PC	Per capita production of fruit products	kg	0.955	0.308
FR_YIELD	Yield of fruit products	100 kg/ha	0.892	0.010
FR_CULT	Cultivated area of fruit products	ha	0.825	0.000
FR_LI	Labor input of fruit products	person/hour per 100 kg of vegetables in agricultural enterprises	0.716	0.000
FR_PCOST	Production cost of fruit products	in AZN, per 100 kg of vegetables in agricultural enterprises	0.837	0.001

Note. Degrees of freedom (Df) for all variables tested via the Shapiro-Wilk test was 26. The null hypothesis for the Shapiro-Wilk test was that the data was non-normally distributed. The profitability of the fruit and vegetable subsectors was expressed in terms of year-on-year percentages. Data from SSCRA (2021).

Table 2 provides descriptive statistics for the variables of interest. As shown in the table, vegetables had a higher value for all indicators, excluding the mean value of production cost (VEG_PCOST). Half of the variables were fairly symmetrical in terms of distribution around the mean, but variables such as vegetable production cost (VEG_PCOST), labor input in vegetable production (VEG_LI), and labor input in fruit production (FR_LI) had moderate and high skewness values.

Table 2: Descriptive Statistics (1995–2020)

	N	Range	Min.	Max.	Mean	Std. dev.	Var.	Skew.	Kurt.
VEG_PROF	26	138.70	-44.40	94.30	19.70	36.61	1340.56	0.19	-0.37
VEG_PROD	26	1290.60	424.10	1714.70	1086.24	338.62	114666.36	-0.45	-0.22
VEG_PROD_PC	26	117.00	56.00	173.00	122.69	30.79	948.07	-0.89	0.16
VEG_YIELD	26	45.00	133.00	178.00	148.53	12.10	146.37	0.89	0.38
VEG_SOWN	26	145.98	49.40	195.38	149.56	41.92	1757.37	-1.44	0.89
VEG_LI	26	11.20	20.20	31.40	23.53	2.08	4.32	2.15	7.81
VEG_PCOST	26	11.02	4.28	15.30	8.68	3.53	12.45	0.55	-1.03
FR_PROF	26	131.80	-64.10	67.70	15.39	29.11	847.24	-0.83	1.06
FR_PROD	26	778.50	321.20	1099.70	676.52	234.58	55026.42	0.07	-1.04
FR_PROD_PC	26	69.00	42.00	111.00	76.07	20.54	422.03	-0.19	-0.95
FR_YIELD	26	130.80	79.60	210.40	123.67	38.72	1499.07	0.84	-0.22
FR_CULT	26	48.33	31.20	79.53	64.92	12.20	148.91	-1.50	1.66
FR_LI	26	11.90	13.20	25.10	22.58	2.51	6.31	-2.60	7.89
FR_PCOST	26	37.89	3.75	41.64	18.81	14.41	207.67	0.48	-1.44

Note. Descriptive statistics include outliers. Data from SSCRA (2021); author performed calculations in SPSS.

Since 10 out of 14 variables were not normally distributed, Spearman's Rho rank correlations were applied instead of Pearson's R correlation analysis to document correlations between variables. Tables 3 and 4 report the results for vegetables and fruits, respectively.

Table 3: Correlation Between Vegetable Subsector Variables Using Spearman's Rho (1995–2020)

		1	2	3	4	5	6	7
1	VEG_PROF	1						
2	VEG_PROD	0.79**	1					
3	VEG_PROD_PC	0.64**	0.88**	1				
4	VEG_YIELD	0.18	0.43*	0.24	1			
5	VEG_SOWN	0.61**	0.52**	0.65**	-0.22	1		
6	VEG_LI	-0.37	-0.64**	-0.51**	-0.78**	-0.12	1	
7	VEG_PCOST	0.62**	0.71**	0.55**	0.20	0.46*	-0.42*	1
** Correlation is significant at the 0.01 level (two-tailed).								
* Correlation is significant at the 0.05 level (two-tailed).								

Note. Calculations were performed in SPSS.

In Table 3, it is worth noting the statistically significant and negative correlations between labor input in the vegetable subsector (VEG_LI), vegetable production (VEG_PROD), vegetable production per capita (VEG_PROD_PC), and vegetable yield (VEG_YIELD). This is because as labor input decreases, vegetable production and yield increase (see Figure 1A in Appendix). This may indicate efficiency and productivity gains in the vegetable sector. Similarly, a significant negative correlation between vegetable production cost and labor input could mean that production cost in the vegetable subsector has increased due to non-labor factors.

Table 4: Correlation Between Fruit Subsector Variables Using Spearman's Rho (1995–2020)

		1	2	3	4	5	6	7
1	FR_PROF	1						
2	FR_PROD	0.61**	1					
3	FR_PROD_PC	0.62**	1.00**	1				
4	FR_YIELD	0.45*	0.89**	0.89**	1			
5	FR_CULT	0.68**	0.65**	0.66**	0.48*	1		
6	FR_LI	0.17	-0.28	-0.28	-0.43*	0.04	1	
7	FR_PCOST	0.47*	0.92**	0.92**	0.89**	0.56**	-0.28	1
** Correlation is significant at the 0.01 level (two-tailed).								
* Correlation is significant at the 0.05 level (two-tailed).								

Note. Calculations were performed in SPSS.

Compared to the correlation analysis of vegetable variables, all statistically significant correlation coefficients were positive for fruit sector variables.

In the presence of significant and high correlations, PCA is considered to be appropriate. This study employed direct oblimin as an oblique rotation method. Niftiyev (2021) reported the results of rotations such as Varimax, Equamax, and Quartimax (orthogonal rotations). In this study, PCA was conducted with IBM SPSS Statistics software (version 23), while OLS regressions were performed in Eviews (version 10).

Moreover, regression analysis was used to understand the underlying reasons for performance volatility in the vegetable subsector. Because the first component of PCA output accounted for the most variance within the variables and provided information about profitability and production, time series data generated through PCA for each year based on factor loadings was used in the regression analysis. Table 3A reports the variable list used in the regression and correlation analysis, while Table 4A illustrates results from the Augmented Dickey-Fuller (ADF) test of stationarity. The latter documented non-stationarity in all variables, except for the variable OIL_EXPORTS/GDP. The first difference of data (which proved to be stationary) was employed in the OLS to avoid spurious estimates.

The vegetable subsector's performance was regressed against extractive industry-related variables in time *t* using the OLS technique. The main estimation model is described below:

$$VEG\_Per_{i,t} = \beta_0 + \beta_1 REER_t + \beta_2 Mining/GDP_t + \beta_3 Oil\_rents_t + \\ + \beta_4 Dummy_t(Rev\_booming_t; Extr\_booming_t) + \varepsilon_t \quad (1) \qquad (1)$$

where β_0 is the intercept and ε_t is the error term. For more explanation on the variables, data sources, and correlation analysis, see Table 3A in Appendix.

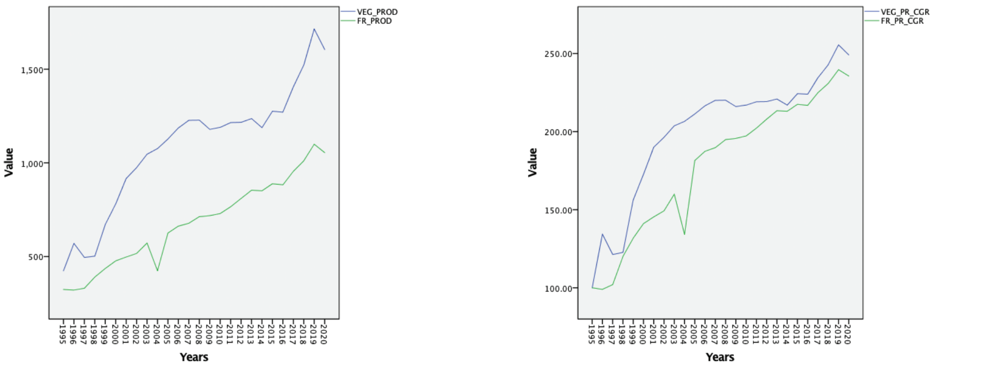
Results

From 1995 to the 2007–2008 period, the vegetable subsector's output (measured in thousands of tons) and cumulative growth rates rapidly increased (see Figure 1). The fruit subsector also exhibited an upward trend, with a one-year drop in 2004. However, the period from 2008 to 2015 saw more of a leveling off in vegetables, while the fruit subsector's performance was relatively stable. Thus, the abovementioned periodic changes in agricultural subsectoral performance must be comprehensively checked and analyzed against the main oil-related indicators in order to outline the oil sector's impacts on their production and profitability.

Figure 1: Overall Performance of Fruit and Vegetable Production in Azerbaijan (1991–2020)

a. Production of vegetables and fruits (in thousands of tons)

b. Cumulative growth rates of fruit and vegetable production (in %), 1995 = 100%



Note. VEG_PR_CGR and FR_PR_CGR represent the cumulative growth rates of the fruit and vegetable subsectors, respectively. Data from SSCRA (2021).

PCA of the Fruit and Vegetable Subsectors

If the Kaiser-Meyer-Olkin measure of sampling adequacy (KMO value) is higher than 0.400, then PCA is recommended and considered to be expedient for the given dataset. Table 5 shows that the KMO value was 0.621 for the vegetable subsector and 0.707 for the fruit subsector, which means that it is appropriate to apply PCA. Moreover, Bartlett's test of sphericity is highly significant for both subsectors, indicating that at least one pair of the correlation is statistically significant among the variables.

The relevance of PCA is highly based on communalities. A communality value for a variable in PCA indicates the level of variance that the component explains. The higher the communality value, the better. Excluding production costs in the vegetable subsector, all variables showed a high level of communality, as demonstrated by Table 6. This is defined as surpassing 0.700; the rule of thumb value is 0.200.

Table 5: Kaiser-Meyer-Olkin Values and Bartlett's Test Results for Fruit and Vegetable Variables

Vegetable subsector		
Kaiser-Meyer-Olkin measure of sampling adequacy		0.621
Bartlett's test of sphericity	Approx. chi-square	205.774
	df	21
	Sig.	0.000
Fruit subsector		
Kaiser-Meyer-Olkin measure of sampling adequacy		0.707
Bartlett's test of sphericity	Approx. chi-square	304.341
	df	21
	Sig.	0.000

Note. Calculations were performed in SPSS.

Table 6: Communalities of Variables Related to the Fruit and Vegetable Subsectors

		Communalities				**Communalities**	
		Initial	**Extraction**			**Initial**	**Extraction**
Vegetable subsector	VEG_PROF	1	0.718	**Fruit subsector**	FR_PROF	1	0.854
	VEG_PROD	1	0.926		FR_PROD	1	0.987
	VEG_PROD_PC	1	0.896		FR_PROD_PC	1	0.987
	VEG_YIELD	1	0.951		FR_YIELD	1	0.908
	VEG_SOWN	1	0.882		FR_CULT	1	0.813
	VEG_LI	1	0.904		FR_LI	1	0.820
	VEG_PCOST	1	0.575		FR_PCOST	1	0.876

Note. Calculations were performed in SPSS.

The first constructed component for the vegetable sector accounted for 59.6% of the variance; after the 24.0% contribution from the second component, the cumulative explanation of the variance was 83.6%. In other words, the value of six variables was protected in the PCA for the vegetable subsector (see Table 7). Meanwhile, the first component of the fruit subsector possessed higher explanatory power (67.6%), and the cumulative percentage was 89.2%. The PCA for the fruit subsector also saved the information of six variables via the two components, as illustrated by the extraction sums of squared loadings in Table 7.

Table 7: Total Variance Explained in the Fruit and Vegetable Subsectors

Vegetable subsector							
Comp.	**Initial eigenvalues**			**Extraction sums of squared loadings**			**Extraction sums of squared loadings**
	Total	% of var.	Cum. %	Total	% of var.	Cum. %	Total
1	4.171	59.583	59.583	4.171	59.583	59.583	4.066
2	1.682	24.023	83.606	1.682	24.023	83.606	2.051
Fruit subsector							
Comp.	**Initial eigenvalues**			**Extraction sums of squared loadings**			**Extraction sums of squared loadings**
	Total	% of var.	Cum. %	Total	% of var.	Cum. %	Total
1	4.731	67.582	67.582	4.731	67.582	67.582	4.717
2	1.515	21.636	89.218	1.515	21.636	89.218	1.569

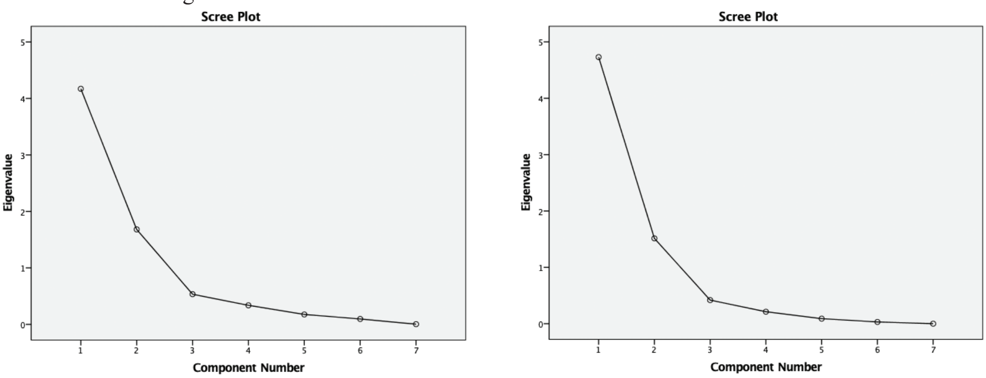
Note. All components greater than two had eigenvalues value lower than 1.000. Calculations were performed in SPSS.

In Figure 2, the scree plots described in panels *a* and *b* suggest two components at best based on the eigenvalues. From the beginning of the research the quantity of the components was fixed, being two, to explain production and profitability and labor input aspects of the fruit and vegetable subsectors. Scree plots supported this approach giving two components as an optimal number to reduce the original seven variables.

Table 8 is the component matrix for both the fruit and vegetable subsectors. In the vegetable subsector, labor input was found to load negatively and yield was found to load weakly on Component 1, while the other variables loaded highly and positively, indicating their strong interconnection. Component 2 had higher and negative loadings from yield compared to the low and positive values of Component 1, but high and positive loadings from labor input. Therefore, the first component mainly reflected the production and profitability dimensions of the variables, and the second component mainly covered productivity.

Figure 2: Scree Plots of the Fruit and Vegetable Subsectors

a. Vegetable subsector b. Fruit subsector



Note. Calculations were performed in SPSS.

Table 8: Component Matrix of PCA.

		Component				Component	
		1	2			1	2
Vegetable subsector	VEG_PROF	0.820	0.212	**Fruit subsector**	FR_PROF	0.724	0.574
	VEG_PROD	0.962	-0.007		FR_PROD	0.991	-0.071
	VEG_PROD_PC	0.939	0.116		FR_PROD_PC	0.994	0.016
	VEG_YIELD	0.244	-0.944		FR_YIELD	0.873	-0.383
	VEG_SOWN	0.806	0.482		FR_CULT	0.826	0.360
	VEG_LI	-0.637	0.706		FR_LI	-0.068	0.903
	VEG_PCOST	0.758	0.020		FR_PCOST	0.887	-0.298

Note. Calculations were performed in SPSS.

The same factor loading patterns applied to the fruit subsector. Labor input loaded highly and positively on Component 2, while other variables mainly loaded on Component 1. However, profitability and cultivated area showed a slightly complex structure compared to the vegetable variable. The complex structure simply means loading highly in both components. Consequently, even if a simple structure is more desirable, the findings do not violate the PCA process.

Pattern matrix which also is provided if the direct oblimin method of the rotation in PCA is selected allows arguing about the significance of the loadings. The widely accepted rule of thumb is a value of 0.200; if the loadings exceed this value, the loading is considered to be significant. As Table 9 shows, all relevant variables for production and profitability dimensions had significant loadings for both the fruit and vegetable subsectors.

Oblique rotations enable loadings and correlations to be separately analyzed. This is why this paper also incorporates the structure matrix described in Table 10. The values of the structure matrix mirrored the correlation coefficients. In the vegetable subsector, the only negative correlation was between labor input and Component 1, while Component 2 had a positive correlation with labor input and sown area. The same pattern held for the fruit subsector; labor input was also negatively correlated with Component 1, but Component 2 had more positive correlations compared to the second component of the vegetable subsector.

Table 9: Pattern Matrix of the Fruit and Vegetable Subsectors

		Component				Component	
		1	**2**			**1**	**2**
Vegetable subsector	VEG_PROF	0.854	0.041	**Fruit subsector**	FR_PROF	0.645	0.621
	VEG_PROD	0.902	-0.208		FR_PROD	0.994	-0.006
	VEG_PROD_PC	0.928	-0.080		FR_PROD_PC	0.985	0.081
	VEG_YIELD	-0.141	-0.992		FR_YIELD	0.917	-0.326
	VEG_SOWN	0.947	0.312		FR_CULT	0.774	0.414
	VEG_LI	-0.321	0.837		FR_LI	-0.185	0.899
	VEG_PCOST	0.720	-0.138		FR_PCOST	0.921	-0.240

Note. Calculations were performed in SPSS.

Overall, the components as a whole were negatively correlated in the vegetable subsector but positively correlated in the fruit subsector (see Table 11). The main reason for this is the complex structure of several loadings in the fruit subsector (see Figure 3).

Table 10: Structure Matrix of the Fruit and Vegetable Subsectors

		Component				Component	
		1	2			1	2
Vegetable subsector	VEG_PROF	0.846	-0.120	Fruit subsector	FR_PROF	0.686	0.663
	VEG_PROD	0.941	-0.377		FR_PROD	0.993	0.060
	VEG_PROD_PC	0.943	-0.254		FR_PROD_PC	0.990	0.146
	VEG_YIELD	0.046	-0.965		FR_YIELD	0.896	-0.265
	VEG_SOWN	0.888	0.134		FR_CULT	0.801	0.465
	VEG_LI	-0.479	0.897		FR_LI	-0.126	0.886
	VEG_PCOST	0.746	-0.274		FR_PCOST	0.905	-0.179

Note. Calculations were performed in SPSS.

Nevertheless, the negative correlations between the labor input, profitability, and production dimensions suited expectations in both subsectors. Over time, production and profitability increased, which enabled reinvestment into more productive technologies and required fewer labor inputs, as measured in hours of work.

Table 11: Component Correlation Matrix

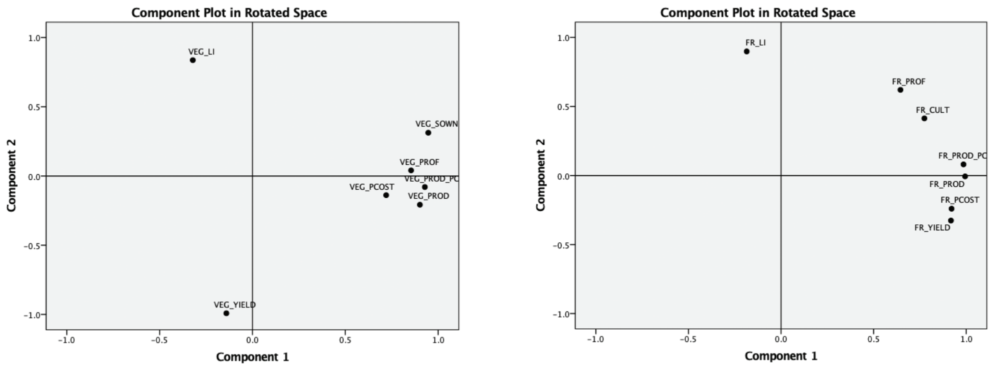
		Component			Component	
		1	2		1	2
Vegetable subsector	1	1.000	-0.188	Fruit subsector	1.000	0.066
	2	-0.188	1.000		0.066	1.000

Note. Calculations were performed in SPSS.

As panel *a* of Figure 3 shows, yield in the vegetable subsector is negatively correlated to the first component, but this correlation is rather weak. Meanwhile, a much stronger negative correlation was observed with the second component, which indicates an inverse relationship with labor input; labor input has decreased since 1999, but yield continued to rise (see also Figure 1A).

Figure 3: Component Plot in Rotated Space for the Fruit and Vegetable Subsectors

a. Vegetable subsector b. Fruit subsector



Note. Calculations were performed in SPSS.

The overall subsectoral performance of fruits and vegetables is represented by factor scores calculated for each year. Panel *a* of Figure 4 displays the fruit and vegetable subsector's performance as measured by the first component's factor loadings, while panel *b* presents the same information with the second component's factor loadings. In short, production and profitability within the fruit subsector exhibited stable and gradual improvement, while the performance of the vegetable subsector was volatile. Except for 2020, labor input cost demonstrated a downward trend in both subsectors.

Figure 4: Performance of the Fruit and Vegetable Subsectors Based on Factor Scores

a. Profitability and production b. Labor input

Note. Calculations were performed in SPSS.

OLS Results: Identifying the Extractive Industry's Impact on the Vegetable Subsector's Performance

Table 12 reports OLS results from the five estimated models in which the performance index for the vegetable subsector was the dependent variable. In all models, the intercept was statistically significant and positive. Other positive associations were found between the vegetable subsector's performance and Real Effective Exchange Rate (REER) and oil rents (models 1,3,5). More importantly, statistically significant and negative association emerged between the vegetable subsector's performance and the revenue boom period (2008–2015). Lastly, Table 5A shows that all models passed the Ramsey RESET test (with the number of fitted terms set at 1), which indicates the absence of model specification errors. Also, the Lagrange multiplier (LM) test of the residuals shows that there was no serial correlation up to two lags, and the cumulative sum (CUSUM) and cumulative sum of squares (CUSUMSQ) demonstrate the stability of the models at a 5% significance level. The heteroscedasticity test of the residuals also did not find any unstable residuals of the models. Lastly, Variance Inflation Factors (VIF) showed that there is not any multicollinearity issue in the estimated models.

Table 12: Ordinary Least Squares (OLS) Estimates for the Performance of the Vegetable Subsector [(VEG_PER)–dependent variable], 1996–2020 and 1997–2020.

Equation name:	**1**	**2**	**3**	**4**	**5**
C	0.23***	0.26***	0.29***	0.27***	0.23***
	[3.66]	[4.25]	[4.57]	[4.44]	[3.66]
REER	0.0.1***	0.01***	0.02***	0.02***	0.01**
	[3.55]	[3.93]	[4.34]	[4.21]	[-1.92]
MINING/GDP	-1.95	-1.98	-1.11	-2.03	-1.92
	[-1.40]	[0.02]	[-1.28]	[-1.57]	[-1.34]
OIL_RENTS	0.02*	0.02	0.02*	0.02	0.02*
	[1.90]	[1.55]	[1.84]	[1.65]	[1.86]
REV_BOOMING	-0.26**	-0.31***	-0.37***	-0.33***	-0.25**
	[-2.31]	[-2.76]	[-3.19]	[-3.01]	[-2.13]
REER(-1)		0.01			
		[0.03]			
MINING/GDP(-1)			-1.12		
			[-1.28]		
OIL_RENTS(-1)				-0.01	
				[-0.92]	
EXTR_BOOMING					-0.01
					[-0.08]
Observations:	24	23	23	23	24
R-squared:	0.43	0.50	0.53	0.52	0.43
F-statistic	3.63	3.42	3.91	3.67	2.75
Prob(F-statistic)	0.02	0.03	0.02	0.02	0.05

Note. Calculations were performed in Eviews; the symbols *, **, and *** indicate statistical significance at 10%, 5%, and 1% levels, respectively; t statistics is in the brackets; the numbers were rounded to the second decimal point to be compact.

Controlling for additional variables which are mainly the lagged versions of the main variables of interest did not change the main model's (model 1) results in a significant manner. Furthermore, even if the Mining-to-GDP ratio was not statistically significant in all models, the sign of the coefficients was negative and the coefficients themselves were bigger than the others. Moreover, if we consider the fact that out of 5 models 4 are statistically significant [Prob(F-statistic<0.05)], then both Mining-to-GDP ratio and revenue booming period are jointly and negatively impacted the performance of the vegetable sector.

Conclusion

This paper evaluates the subsectoral performance of the agriculture sector by employing the fruit and vegetable subsectors as case studies. To date, there have been few or no studies that employed PCA to evaluate subsectoral performance in the Azerbaijani economy. According to PCA results, the vegetable subsector outperformed the fruit subsector in terms of production and profitability between 1999 and 2014, while labor input in both subsectors showed a downward trend. In addition, Azerbaijan has served as a case study for the Dutch disease and resource curse, which mainly assume that oil boom economies' non-resource tradable sectors lag behind due to specific drawbacks such as exchange rate overvaluation, low returns to capital, and rent-seeking behavior. Based on time series index values for the performance of the vegetable subsector, the current study used OLS models to identify whether there were statistically significant and theoretically meaningful associations between the extractive industry and subsectoral performance.

The study revealed partial evidence of a negative and statistically significant association between the performance of the vegetable subsector and oil-related economic indicators. Mining-to-GDP, revenue boom period, and the performance of the vegetable subsector were negatively associated. However, the performance of the vegetable subsector was found to have a statistically significant and positive relationship with oil rents and REER—the two most fundamental indicators for assessing the adverse effects of the extractive industry on non-resource tradable sectors. The last finding is surprising and unexpected but also might reflect the model-specific realities of the study which neglect additional factors. Hence, not every channel of the resource curse or Dutch disease theories can be relevant for the case of Azerbaijan. Conversely, the performance of the fruit subsector failed to demonstrate any statistically significant results based on the same models; thus, these results were not reported.

The study's limitations should also be considered. Firstly, some components in the PCA behaved in a complex manner, which may have slightly decreased reliability. Secondly, quantitative methods cannot capture underlying reasons for subsectoral performance in specific years. For instance, the sharp drop in the vegetable sector's performance in 2014–2015 could be explained by a degraded overall macroeconomic environment, which impacted local producers; however, the reasons behind the sharp drop in the performance of the fruit subsector in 2004 remain unknown. Therefore, qualitative methods such as expert interviews and focus group discussions could be beneficial to this end. Thirdly, in addition to OLS estimates, further research could focus on autoregressive distributed lags (ARDL) and error correction (EC) models to estimate short-run and long-run relationships in a more sophisticated and systematic way. Otherwise, OLS estimates alone cannot address the necessary patterns of the cause-and-effect relationship between subsectoral performance and the extractive industry's adverse effects in the case of the Azerbaijani economy. Lastly, OLS estimates provided mixed evidence of the extractive industry's negative impact on the vegetable susector. The type of data also matters in this case; annual data regarded REER usually is too aggregated to indicate comprehensiveness. Despite a certain degree of inconclusiveness, the current study provided the first examination of subsectoral performance based on PCA in the Azerbaijani economy.

REFERENCES

Aslund, A. (2013). How Capitalism was Built: The Transformation of Central and Eastern Europe, Russia, the Caucasus, and Central Asia. Cambridge University Press, New York. DOI: 10.1017/CBO9781139207850.

Auty, R. (1993). Sustaining Development in Mineral Economies: The Resource Curse Thesis. Routledge, London. DOI: 10.4324/9780203422595.

Bayramov, G., & Conway, P. (2010). The Azerbaijan Producers Survey: Dutch Disease and Financial Crisis. Economic Research Center, Baku, Azerbaijan. Retrieved from http://erc.az/files/reports/Bayramov_Conway_Azerbaijan.pdf.

Bayramov, V., & Abbas, G. (2017). Oil Shock in the Caspian Basin: Diversification Policy and Subsidized Economies. Resources Policy, 54, 149-156. DOI: 10.1016/j.resourpol.2017.10.006.

Cornell, S. E. (2014). Azerbaijan: Going It Alone. In Satrr, S.F. & Cornell, S. E. (Eds). Putin's Grand Strategy: The Eurasian Union and Its Discontents, 145-155.

Czech, K. (2018). Oil Dependence of post-Soviet Countries in the Caspian Sea Region: the Case of Azerbaijan and Kazakhstan. Acta Scientiarum Polonorum. Oeconomia, 17(3), 5-12.

Eldarli, J. (2018). The Role of Entrepreneurial Entities in Diversification of Economy. İpək Yolu, 8, 5-18, Retrived from http://au.edu.az/upload-files/menu/ipekyolu/2018_3/1EldarliJ.pdf. DOI 10.2139/ssrn.3132986.

Hamidova, L. (2020). Diversification of The Economy in the Resource-Rich Countries (By The Example of Azerbaijan, Kazakhstan, and Russia). Book of Proceedings of the 55th International Scientific Conference on Economic and Social Development, Vol. 2/4, 736-744.

Ibadoghlu, G., Alasgarov, K., & Bayramov, G. (2013). Oil and Gas Revenue Management in Azerbaijan. Policy Paper on Revenue Management in Azerbaijan. Retrieved from: Retrieved from: https://ssrn.com/abstract=3104106. DOI: 10.2139/ssrn.3104106.

Ibadoglu, G. (2008). Current State of Azerbaijan's Economy: Myths and Realities. Transition Studies Review, 15(2), 425-430. DOI: 10.1007/s11300-008-0015-4.

Jolliffe, I. T., & Cadima, J. (2016). Principal Component Analysis: A Review and Recent Developments. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 374(2065), 20150202. DOI: 10.1098/rsta.2015.0202.

Kerimova, T. I. (2014). Efficiency Indicators use for Estimating of the Azerbaijan Agrarian Sector Investment Potential in the Modern Period. American-Eurasian Journal of Agricultural & Environmental Sciences, 14(4): 352-357.

Kwak, S. K., & Kim, J. H. (2017). Statistical data Preparation: Management of Missing Values and Outliers. Korean Journal of Anesthesiology, 70(4), 407-411. DOI: 10.4097/kjae.2017.70.4.407.

Niftiyev, I. (2020a). The De-industrialization Process In Azerbaijan: Dutch Disease Syndrome Revisited. Proceedings of the 4th Central European PhD Workshop on Technological Change and Development, 357–396. Retrieved from: https://www.econstor.eu/handle/10419/227485.

Niftiyev, I. (2020b). Macroeconomic Effects of the Commodity Price Shocks on the post-Soviet Resource-rich Economies. Azərbaycanın Vergi Jurnalı, 2(139), 49-58.

Niftiyev, I. (2020c). The Role of Agriculture and Agrarian Sectors in the Azerbaijan Economy: Main Trends and Dynamics since 1991. Working Paper No. 1, SSRN Electronic Journal, Retrieved from: https://ssrn.com/abstract=3681896. DOI: 10.2139/ssrn.3681896.

Niftiyev, I. (2021). Understanding Principal Component Analysis (PCA) in the Azerbaijan Economy: Case Studies of Vegetable and Fruit Sectors. Working Paper No. 4, SSRN Electronic journal, Retrieved from: https://ssrn.com/abstract=3782426. DOI: 10.2139/ssrn.3782426.

Niftiyev, I., & Czech, K. (2020) Dutch Disease Perspective on Vegetable Exports in the Azerbaijan Economy. Journal of Applied Economic Sciences, 4(70), 813-827.

Niftiyev, I., & Namazova, N. (2020). Analysis of cyclicality in the Azerbaijan economy: Results of the Chi-Square test. Academic Journal of Economic Studies, 6(2), 122-134.

Pashayev, E., & Aliyev, U. (2020). Analysis of State Revenues and Expenditures to Ensure Sustainable Development in Azerbaijan and Directions for their Improvement. Expert Journal of Economics, 8(1), Retrieved from: http://economics.expertjournals.com/23597704-803/.

Shalbuzov, N., Fikretzade, F., & Huseyn, R. (2020). The International Competitiveness of Azerbaijani Fruit and Vegetable Products. Studies in Agricultural Economics, 122(1), 51-55.

Spoor, M., & Visser, O. (2001). The State of Agrarian Reform in the Former Soviet Union. Europe-Asia Studies, 53(6), 885-901, DOI: 10.1080/09668130120078540.

SSCRA – The State Statistical Committee of the Republic of Azerbaijan. (2021). Agriculture, Forestry and Fishing. Retrieved from: https://www.stat.gov.az/source/agriculture/?lang=en.

The Economist (1977). The Dutch disease. The Economist, 26, 82-83.

Zulfigarov, F., & Neuenkirch, M. (2020). The Impact of oil Price Changes on Selected Macroeconomic Indicators in Azerbaijan. Economic Systems, 44(4), DOI: 10.1016/j.ecosys.2020.100814.

Appendix

Table A1: Component Score Coefficient Matrix for the Fruit and Vegetable Subsector

		Component				Component	
		1	2			1	2
Vegetable sector	VEG_PROF	0.218	0.041	Fruit sector	FR_PROF	0.128	0.395
	VEG_PROD	0.225	-0.093		FR_PROD	0.212	-0.019
	VEG_PROD_PC	0.235	-0.023		FR_PROD_PC	0.209	0.038
	VEG_YIELD	-0.058	-0.541		FR_YIELD	0.200	-0.226
	VEG_SOWN	0.248	0.190		FR_SOWN	0.159	0.259
	VEG_LI	-0.063	0.446		FR_LI	-0.053	0.589
	VEG_PCOST	0.180	-0.059		FR_PCOST	0.200	-0.171

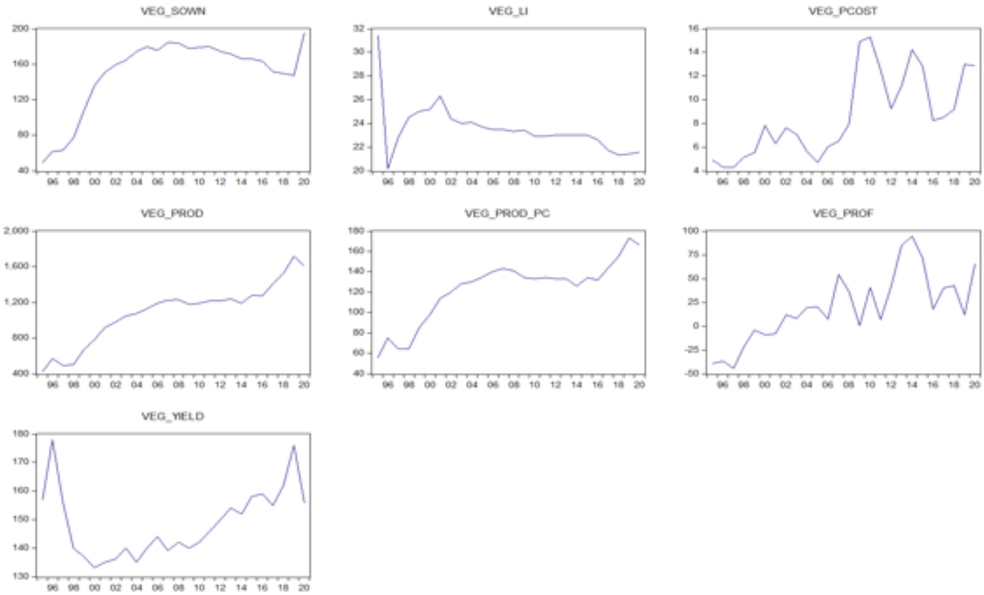
Note. Calculations were performed in SPSS.

Table A2: Component Score Covariance Matrix for the Fruit and Vegetable Subsector

		Component				Component	
		1	2			1	2
Vegetable sector	1	1.035	-0.376	Fruit sector	1	1.004	0.133
	2	-0.376	1.035		2	0.133	1.004

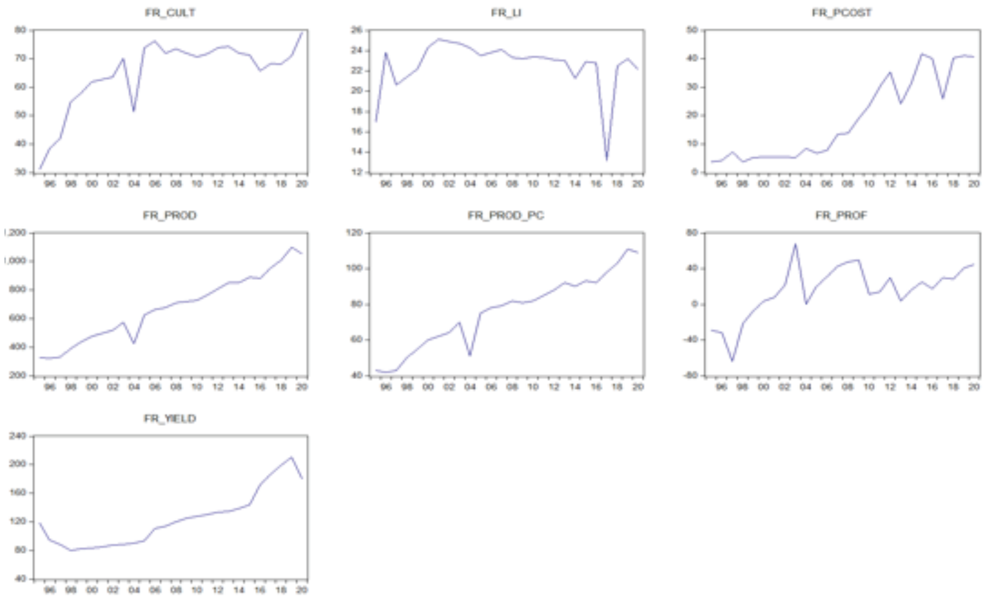
Note. Calculations were performed in SPSS.

Figure 1A: Vegetable variables.



Source. SSCRA (2021).

Figure 2A: Fruit Variables.



Source. SSCRA (2021).

Table 3A: The list of vaiables used in the regression analysis and correlation coefficients (Spearman's Rho) among them.

		1	2	3	4	Levels of measurement	Source
1	**Vegetable performance (Veg_Per)**	1				Index scores generated by PCA	SSCRA
2	**Oil Rents**	0.31	1			% of GDP	World Bank
3	**Mining/GDP**	0.64**	0.82**	1		Ratio based on mining industry's output and GDP in million AZN	SSCRA
4	**Real Effective Exhange Rate (REER)**	0.66**	0.05	0.44*	1	In %, 2007=100%	Bruegel Data sets
5	**Revenue Booming**					Dummy variable, 2008–2015	----
6	**Extraction Booming**					Dummy variable, 2006–2010	----

Note. Calculations were performed in SPSS; correlation analysis has been applied to the winsorized data; the symbols *, **, and *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

Table 4A: Stationarity test (ADF) results.

Null Hypothesis: the variable has a unit root					
	At Level				
		VEG_PER	REER	MINING_GDP	OIL_RENTS
With Constant	t-Statistic	-1.2934	-1.6446	-2.2905	-2.0029
	Prob.	***0.6155***	***0.4447***	***0.1831***	***0.2837***
		n0	n0	n0	n0
With Constant & Trend	t-Statistic	-1.7188	-2.5457	-2.2938	-1.9545
	Prob.	***0.7112***	***0.3055***	***0.4205***	***0.5954***
		n0	n0	n0	n0
Without Constant & Trend	t-Statistic	-1.124	0.3260	0.4932	-0.4722
	Prob.	***0.2296***	***0.7715***	***0.8145***	***0.5003***
		n0	n0	n0	n0
	At First Difference				
		d(VEG_PER)	d(REER)	d(MINING_GDP)	d(OIL_RENTS)
With Constant	t-Statistic	-3.7308	-3.1243	-3.2488	-4.3550
	Prob.	***0.0105***	***0.0387***	***0.0299***	***0.0026***
		**	**	**	***
With Constant & Trend	t-Statistic	-3.7161	-2.9549	-3.1986	-4.2935
	Prob.	***0.0417***	***0.1651***	***0.1092***	***0.0129***
		**	n0	n0	**
Without Constant & Trend	t-Statistic	-2.9032	-3.1771	-3.1870	-4.4296
	Prob.	***0.0057***	***0.0029***	***0.0028***	***0.0001***
		***	***	***	***

Note. Calculations were performed in Eviews; the symbols *, **, and *** indicate statistical significance at 10%, 5%, and 1% levels, respectively; Lag Length based on SIC; Probability based on MacKinnon (1996) one-sided p-values.

Table 5A: Stability Test Results of the OLS Models.

Test name	**Ramsey Reset Test**	**Variance Inflation Factors**	**LM Test**	**Heteroscedasticity Test**	**CUSUM**	**CUSUM SQ**
Models which passed the test	1,2,3,4,5	1, 2,3,4,5	1,2, 3,4,5	1, 2,3,4,5	1,2,3,4,5	1,2,3,4,5
Models which failed the test						
Total	5	5	5	5	5	5

Note. Calculations were performed in Eviews.