|  |
| --- |
|  |
| Factors Affecting Arable Land Price |
| (Econometric analysis) |

**Prepared by**

**Sergey Mikhaylov**

**Nikita Tymchik**

**Nevezhin Maksim**

**Introduction.**

In this paper, we will try to identify the factors that explain the price of arable land among the EU countries in 2016.

To form a set of explanatory variables, it is necessary to provide a list of economic and social indicators that can somehow affect the formation of the value of arable land.

As we know, the price is formed from the equality of supply and demand. Accordingly, it is necessary to find indicators that could explain to some extent the supply and demand for arable land. However, it is important not only to identify such indicators, but also to find the necessary data.

Demand factors may include the following: GDP per person, logistics accessibility, crop output per hectare of sown area, percentage of urban population.

The supply factor in the land market is traditionally availability of arable land per person.

**Relevance of the selected factors.**

We chose GDP per capita as the first demand factor. The variables related to the income of the population are statistically significant in some studies, which indicates that economic income, which should be GDP per capita, plays an important role in determining the price of agricultural land [4]. This factor shows how much wealth a person has in a particular country. Wealth gives him the opportunity to acquire the necessary goods, which increases aggregate demand, which affects the rise in prices of all economic goods, and also increases access to investment in land assets. This indicator is broad enough to study specific issues, however, it is accessible and accurate.

According to some authors, an important role in determining the level of prices for arable land is played by infrastructure, which includes: the development of industry, urban infrastructure, improved logistics supply network, etc. The more developed these factors are, the higher the level of viability of agriculture. Non-agricultural opportunities offered to potential investors make the development of farming more profitable [2]. Some studies have concluded that the distance to the city center has a big impact on land prices. There are empirical estimates based on a spatial index that shows the spatial relationship of each location with all urban centers across the country [3]. Based on these studies the next demand factor is the logistics efficiency index. The availability of the necessary infrastructure, the quality of road networks have a positive effect on the attractiveness of agriculture for several reasons: they increase accessibility for establishing relationships with counterparties, and reduce the costs of transportation and storage of finished products.

The third demand factor is the output of crop production per hectare of arable lands. This indicator shows the productivity of land holdings, reflecting together geographical, institutional, social characteristics and competitive advantages, thus potentially increasing their attractiveness to investors. However, it should be borne in mind that the indicator may be distorted due to different domestic prices for products. OECD states that crop production depends on the availability of arable land and is affected in particular by yields, macroeconomic uncertainty, as well as consumption patterns . This is also in line with Ricardo's theory of surplus rent, which allows one unit of land to generate more income

The final demand factor in our analysis is the percentage of urban population. The variable of the urbanization index is quite statistically significant. Also, the change in the price of agricultural land is strongly associated with the expected capital gain [4]. Urbanization contributes to rising prices for real estate and land plots within cities, which can lead to an increase in prices of the entire land market. In addition, this indicator reflects the availability of labor in rural areas, which can reduce production costs in agriculture. It is not surprising then that the existing literature suggests urban influence is a dominant factor in the determination of agricultural land values [1].



Table a. Final Data (Source: World Bank, Eurostat)

**Descriptive statistics.**

The average value of the price for arable land (table 1) differs greatly from the minimum and maximum values. 50% of lands fall short of the average, with the maximum value significantly different from the values of the top 25% of lands. A similar picture is observed when comparing these indicators for GDP per capita and crop output per hectare of arable land. The remaining indicators already have a not so large standard deviation.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Arable land price, euro per ha | GDP per person, euro | Logistics performance index score | Arable land availability, hectares per person | Urban population, per cent | Crop output per arable land hectares, euro per ha |
| count | 23,00 | 23,00 | 23,00 | 23,00 | 23,00 | 23,00 |
| mean | 15279,57 | 28019,39 | 3,62 | 0,29 | 0,74 | 2337,99 |
| std | 14940,27 | 20301,11 | 0,41 | 0,19 | 0,13 | 2669,76 |
| min | 1958,00 | 6697,49 | 2,81 | 0,06 | 0,54 | 456,48 |
| 25% | 4156,50 | 13029,80 | 3,33 | 0,11 | 0,66 | 977,77 |
| 50% | 9083,00 | 19472,06 | 3,67 | 0,26 | 0,74 | 1357,44 |
| 75% | 21935,50 | 38307,75 | 3,91 | 0,42 | 0,85 | 3055,07 |
| max | 62972,00 | 92124,24 | 4,22 | 0,75 | 0,98 | 13138,94 |

Table . Summary statistic table

A visual representation of the results obtained is shown below in Figure 1

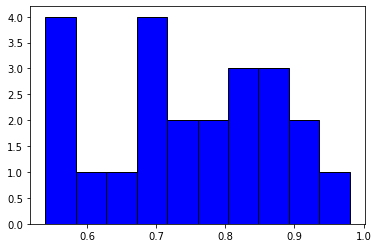
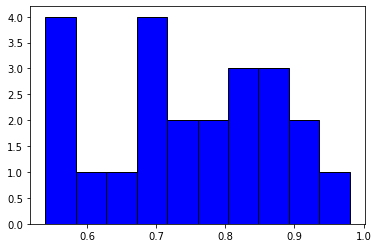
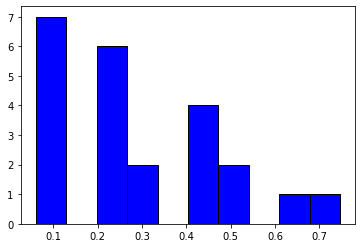
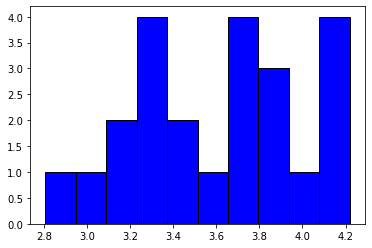
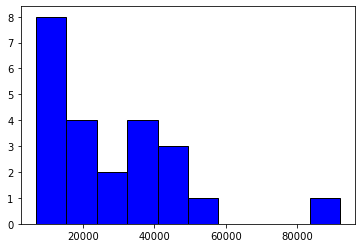


Figure 1. Distribution of GDP per person, Logistics Performance Index estimate, Availability of arable land, Urban population, Crop production per hectare of arable land.

Thus, the average price of arable land by country is 15,280 euros, with a per capita income of 28,019 euros. In Europe, logistics is well developed, as evidenced by the index of 3.62, and there is also a high level of urbanization (0.74). On average, there are 0.29 hectares per person, while one hectare produces crops worth 2338 euros.

**The results of the analysis.**

To account for the high standard errors of arable land price and GDP per capita and abnormality of their distribution, we take the logarithm of these factors to correct the model.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Source** | **SS** | **df** | **MS** | Number of obs = 23  F( 5, 17) = 9.19  Prob > F = 0.0002  R-squared = 0.7299  Adj R-squared = 0.6505  Root MSE =.58268 |
| **Model** | 15.5974048 | 5 | 3.11948095 |
| **Residual** | 5.7717459 | 17 | .339514465 |
| **Total** | 21.3691507 | 22 | .97132503 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Inp** | **Coef.** | **Std. Err.** | **t** | **P>|t|** | **[95% Conf. Interval]** | |
| **lngdp\_person** | .3644474 | .4237263 | 0.86 | 0.402 | -.529537 | 1.258432 |
| **logistic\_perf** | .1277504 | .7505826 | 0.17 | 0.867 | -1.45584 | 1.711341 |
| **land\_ha\_person** | -2.202032 | .9029664 | -2.44 | 0.026 | -4.107124 | -.2969391 |
| **urban\_pop** | .3470695 | 1.464307 | 0.24 | 0.815 | -2.742347 | 3.436486 |
| **crop\_output\_ha** | .0001025 | .0000592 | 1.73 | 0.102 | -.0000225 | .0002275 |
| **\_cons** | 5.227283 | 2.645357 | 1.98 | 0.065 | -.3539318 | 10.8085 |

Table 2. Results of the first regression

Using the Stata software, we carry out the necessary calculations. We get the following results:

- the factor reflecting the amount of arable land per person in the country turned out to be significant, while, as expected, with a negative coefficient in front of it, since this parameter reflects the supply of land.

- significant (at a 15% significance level) was the factor of the volume of crop output per hectare of arable land, which confirms the hypothesis of soil productivity. It should be noted that the high level of significance assumed by us in this work is due to the small number of observations.

- the accuracy of the model is quite high (R-squared = 0.73), while the model is adequate.

The Ramsey test reflects the fact that there are no missing variables in the model.

|  |
| --- |
| . ovtest  Ramsey RESET test using powers of the fitted values of lnp  Ho: model has no omitted variables  F (3, 14) = 0.19  Prob > F = 0.8992  Table 3. Ramsey RESET test |

**Multicollinearity**

However, slightly overestimated VIFs for GDP and logistics may indicate multicollinearity. The high correlation of these indicators (0.88) also reflects the presence of a connection between them. In addition, the low standard errors of the logistics performance index indicates that it will not be able to explain price fluctuations. Thus, we have the opportunity to remove this indicator in order to test the significance of GDP per capita in the new model.

|  |  |  |
| --- | --- | --- |
| **Variable** | **VIF** | **1/VIF** |
| **logistic\_p~f** | 6.04 | 0.165596 |
| **Ingdp\_person** | 5.46 | 0.183288 |
| **urban\_pop** | 2.39 | 0.418920 |
| **land\_ha\_pe~n** | 1.96 | 0.509387 |
| **crop\_outpu~a** | 1.62 | 0.616831 |
| **Mean VIF** | 3.49 | |

Table 4. VIFs of the first regression

|  |  |  |
| --- | --- | --- |
|  | **1ngdp\_~n** | **logist\_~f** |
| **Ingdp\_person** | 1.0000 | |
| **logistic\_p~f** | 0.8854 | 1.0000 |

Table 5. Сorrelation between GDP per person and logistic perfomance index

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Source** | **SS** | **df** | **MS** | Number of obs = 23  F ( 4, 18). = 12.13  Prob > F = 0.0001  R-squared = 0.7294  Adj R-squared = 0.6693  Root MSE = .56674 |
| **Model** | 15.5875695 | 4 | 3.89689238 |
| **Residual** | 5.78158115 | 18 | .321198953 |
| **Total** | 21.3691507 | 22 | .97132503 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Inp** | **Coef.** | **Std. Err.** | **t** | **P>|t|** | **[95% Conf. Interval]** | |
| **lngdp\_person** | . 4183023 | .2741149 | 1.53 | 0.144 | -. 1575917 | .9941962 |
| **land\_ha\_person** | -2.189521 | .8753581 | -2.50 | 0.022 | -4.02858 | -. 3504618 |
| **urban\_pop** | . 4557763 | 1.281639 | 0.36 | 0.726 | -2.236847 | 3.1484 |
| **crop\_output\_ha** | .0001036 | .0000573 | 1.81 | 0.087 | -.0000168 | .000224 |
| **\_cons** | 5.063485 | 2.396706 | 2.11 | 0.049 | .0281925 | 10.09878 |

The new model, at a 15% significance level, additionally reflects the significance of GDP

Table 6. Results of the second regression

per capita.

|  |
| --- |
| . ovtest  Ramsey RESET test using powers of the fitted values of lnp  Ho: model has no omitted variables  F (3, 15) = 0.23  Prob > F = 0.8753 |

|  |  |  |
| --- | --- | --- |
| **Variable** | **VIF** | **1/VIF** |
| **Ingdp\_person** | 2.41 | 0.414338 |
| **land\_ha\_pe~n** | 1.95 | 0.512785 |
| **urban\_pop** | 1.93 | 0.517345 |
| **crop\_outpu~a** | 1.60 | 0.623455 |
| **Mean VIF** | 1.98 | |

Table 7. VIFs of the second regression

At the same time, the situation with multicollinearity has greatly improved with the unchanged results of the test for missing variables.

**Outliers**

Let's check the data for atypical observations: vertical outliers using studentized residuals, unbalance points using Cook's distance. The test results showed the presence of one vertical outlier (Slovakia) and the presence of two points of imbalance (Slovakia and the Netherlands). If we talk about Slovakia, then there is indeed an unreasonably high cost of land with relatively low explainable parameters. If we talk about the Netherlands, then the point of imbalance is due to the logarithm of the price, but not the logarithm of the crop output per ha, since in the Netherlands they are high. If there is no logarithm of the dependent variable, there is no imbalance point in the model.

When the outlier (Slovakia) is removed, the model improves somewhat, however, the removal must be justified, so we will leave it unchanged.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Source** | **SS** | **df** | **MS** | Number of obs = 22  F( 4, 17) = 20.04  Prob > F = 0.0000  R-squared = 0.8250  Adj R-squared = 0.7839  Root MSE = .45604 |
| **Model** | 16.6723432 | 4 | 4.16808579 |
| **Residual** | 3.53551543 | 17 | .207971496 |
| **Total** | 20.2078586 | 21 | .962278981 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Inp** | **Coef.** | **Std. Err.** | **t** | **P>|t|** | **[95% Conf. Interval]** | |
| **lngdp\_person** | .3719499 | .221021 | 1.68 | 0.111 | -.0943637 | .8382634 |
| **urban\_pop** | 1.518383 | 1.080791 | 1.40 | 0.178 | -.7618857 | 3.798652 |
| **crop\_output\_ha** | .0001156 | .0000463 | 2.50 | 0.023 | .000018 | .0002132 |
| **land\_ha\_person** | -1.922757 | .7090316 | -2.71 | 0.015 | -3.418683 | -.4268309 |
| **\_cons** | 4.558648 | 1.934653 | 2.36 | 0.031 | .4768871 | 8.640409 |

Table 8. Results regression without log of price

**Heteroskedasticity**

Since the statistics are small, the Breusch-Pagan test can be used to test the regression for heteroscedasticity. After it was carried out, we received F-value= 0.91 and p-value=0.48, from which it follows that we do not reject the null hypothesis, and there are no signs of heteroscedasticity in our model. White's test showed a similar result p-value=0.13. Thus, both tests give us the right to assert homoscedasticity in the model.

**Conclusions.**

Taking advantage of the fact that the logistics efficiency index across European countries does not differ much or has a low standard errors, as well as the fact that the logarithm of GDP per capita and the logistics efficiency index are highly correlated, VIF is inflated, we chose to remove the logistics efficiency index from the model, leaving only 4 explanatory parameter. At the same time, we took the logarithm of the price per hectare of arable land and GDP per capita, preferred not to remove emissions, since their removal is not only unreasonable, but also does not greatly improve the model, since it is already quite accurate and passes standard tests. The results were as follows:

- an increase in GDP per capita by 1 percent contributes to an increase in the price of arable land by 0.42%, assuming a 15% significance level

- an increase in the amount of arable land per person by 0.1 contributes to a decrease in the price of arable land by approximately 80%

- an increase in output per hectare per unit contributes to an increase in land prices by 0.01%

- the indicator of urbanization was insignificant.

**References**

1. Blank, S.C. (2007). "Farmland Values as an Indicator of Regional Economic Performance?" Agricultural and Resource Economics Update, University of California. <https://s.giannini.ucop.edu/uploads/giannini_public/33/df/33dfe89a-f7f0-4a6f-b6ee-066b7cf6226c/v10n6_1.pdf>
2. FACTORS AFFECTING FARMLAND PRICES IN FINLAND//Perttu Pyykkönen// Helsinki 2006
3. Location determinants of agricultural and prices Standort Determinanten von agrarwirtschaftlichen Landpreisen// Pia Nilsson, Sara Johansson // Jahrbuch für Regionalwissenschaft volume 33, pages 1–21 (2013) // https://link.springer.com/article/10.1007/s10037-012-0071-4
4. Agricultural Land Values under Urbanizing Influences // Yue Jin Shi, Timothy T. Phipps, Dale Colyer // Land Economics, Vol. 73, No. 1 (Feb., 1997), pp. 90-100 (11 pages) // <https://www.jstor.org/stable/3147079>