Ukrainian Catholic University

Faculty of Applied Sciences

Business Analytics & Computer Science Programmes

Ukrainian Pets Food Market

Econometrics final report

Authors:
Iryna Voitsitska
Yaryna Petruniv

1 May 2024

Contents

Abstract

This project investigates the impact of the full-scale invasion of Ukraine by Russia in 2022 on the Ukrainian internal and foreign markets for pet products. The main objectives are to determine whether significant changes occurred in the turnover of pet goods in Ukraine and if there was a decrease in export volume following the onset of war. The analysis employs econometric methods, suitable for panel data analysis. The code is available on *link*.

Keywords: Econometric Analysis, Pet Industry, Pet Market Analysis

1 Introduction

The topic of this project is the research of Ukrainian Pets Market between 2021 and 2023. The aim of this project is to explore and analyze Ukrainian internal and foreign markets for pets products. Determine whether it has changed drastically with the beginning of full-scale invasion of Ukraine launched by russia in 2022. This area is unique and interesting because there are no fundamental studies relating to pets products. Ukrainian companies that distribute such vendibles are the audience of this project, especially Suziria Distribution, which is a Ukrainian company that have been significantly influencing the market for the last 30 years. This is also the company that provided the data for this project. The main questions that are expected to be covered are the following:

- Was there a significant change in Ukrainian market for pets goods turnover?
- Did the export volume of goods for pets decrease with the beginning of full-scale war?

During the further research there might arise more questions that can be considered.

2 Literature Review

For discovering methods that are going to be used in this project the following literature is going to be researched:

- "Mostly Harmless Econometrics" by Joshua D. Angrist - It shows the basic tools of applied econometrics such as linear regression for statistical control, instrumental variables methods for the analysis of natural experiments, etc.
- "Econometric Analysis of Cross Section and Panel Data" by Jeffrey M. Wooldridge — The second edition of this textbook provides a comprehensive treatment of econometric research, covering both cross-sectional and panel data methods. The book offers a thorough framework for econometric analysis, including broader coverage of missing data and cluster problems, as well as new discussions on treatment effects with panel data;
- Video series *"A full course in econometrics - undergraduate level"* [1] by Ben Lambert - This selection of videos covers a full course in econometrics. It starts at the absolute beginning assuming no prior knowledge, and will eventually build up to more advanced topics in regression analysis;

- *Analysis of the Future Development Trend of the Pet Industry* [2] — This article addresses challenges through a analysis spanning political, economic, social, and technological fields. The findings indicate a projected upward trajectory for the market economy of the pet industry;

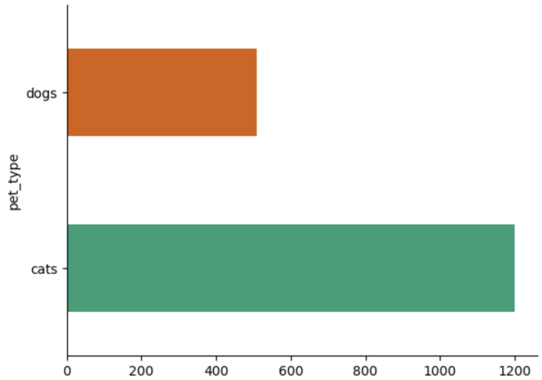
- *"Europe Pet Food Market Size and Share Analysis"*[3] — The report includes the pet market volume, the market size of pet food by products, distribution channels, etc.

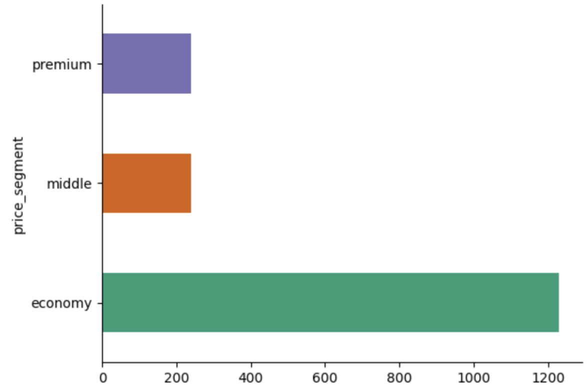
3 Data analysis

Suziria Distribution provided the datasets for this project. There are four large datasets that include various data on Ukrainian internal and export market for pets products, such as revenue and number of items/kilograms sold by a company during determined periods of time. The datasets one can find by the following ***link***. For answering the question about the impact of the war on export volume the data was cleaned and the ***new dataset*** was constructed. The data for this project contains the following columns:

- *export_volume_units*: the volume of pet food exports measured in units.
- *export_volume_tons*: the volume of pet food exports measured in tons.
- *price*: the price of pet food exports measured in millions of hryvnias.
- *price_per_unit*: the price of pet food exports per one unit.
- *war*: a binary variable indicating whether there was a war (1) or not (0) during the export period.
- *category*: the category of pet food product. The categories are *wet*, *dry_cats_and_dogs*, *treats*. Before building the model, the categories will be converted into the dummy variables.
- $manufacturer$: name of the company that is responsible for manufacturing the pet food products.
- *pet_type*: the type of pet for which the food is intended. The types are *cats*, *dogs*. Before building the model, the types will be converted into the dummy variables.
- *price_segment*: the price segment of the pet food product (*economy*, *middle*, *premium*). Before building the model, the segments will be converted into the dummy variables.
- *package_type*: the type of packaging used for the pet food products (*can*, *pack*, *pouch*, *sack*). Before building the model, the types will be converted into the dummy variables.
- *export*: A binary variable indicating whether the product was exported (1) or not (0).
- *observation*: a categorical variable indicating different observations.
- *time_period*: the time period during which the export occurred. There is available data for each month from January 2021 to June 2023.

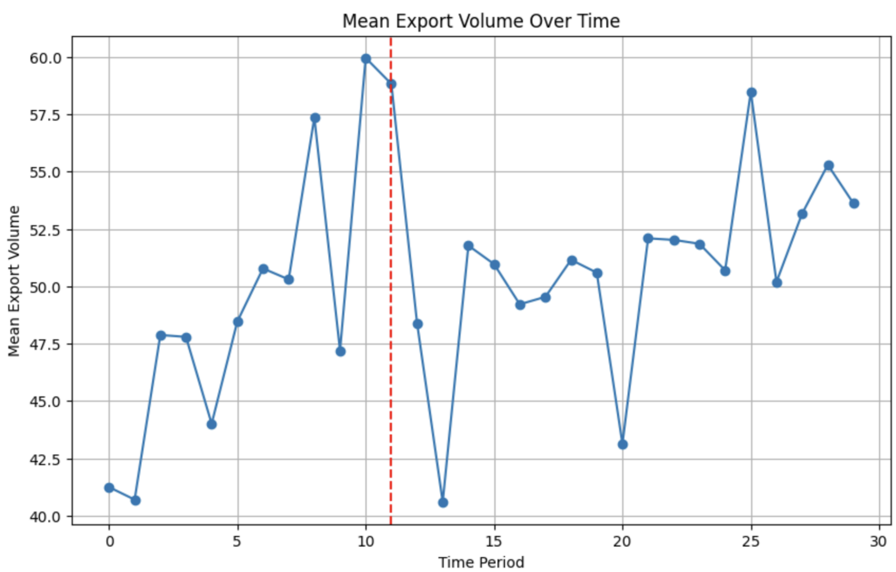
Let’s take a closer look at the data.



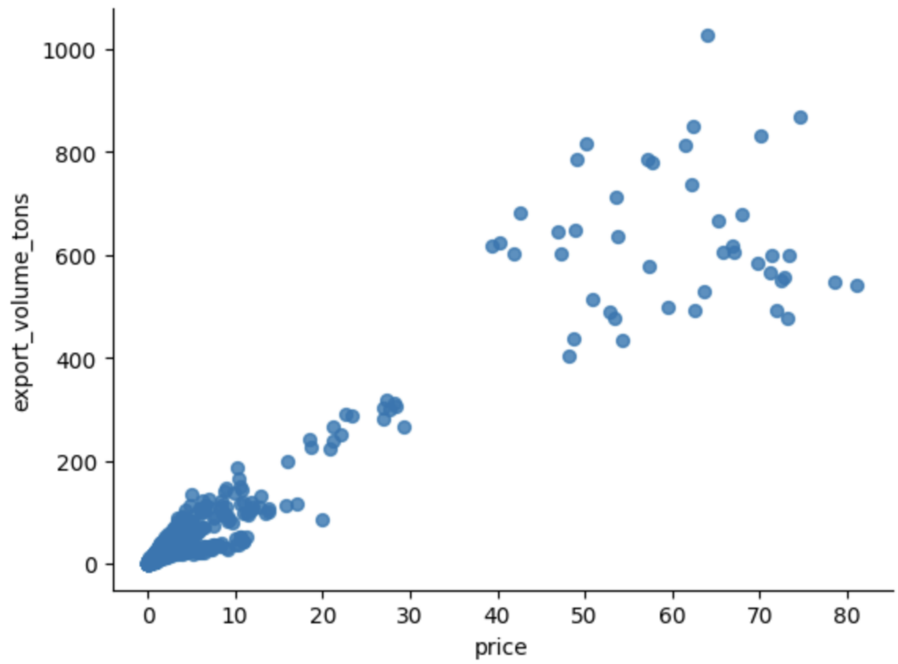


On the given plots, one can see that the cat food is the most popular for the export, considering price segment, economy one is the mostly exported.

The hypothesis that are going to be tested is whether the export volume of goods for pets decrease with the beginning of full-scale war.

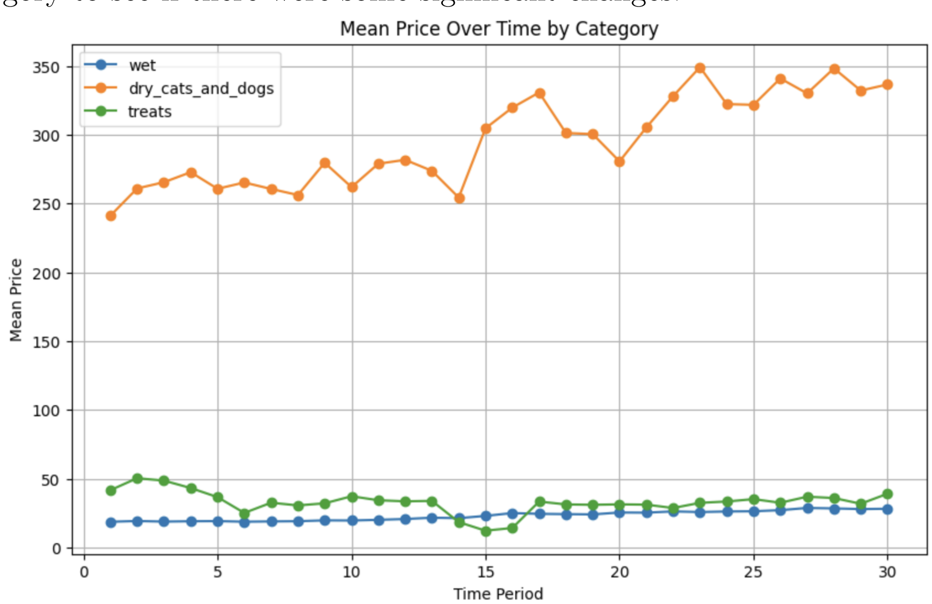


On this plot, the x-axis shows the months labeled from 1 to 30 where the 13 month is February 2022 and y-axis shows the mean value of export volume in tons within this time period. Looking at the plot, one can observe that there was a significant decrease in export volume in December 2022 (two months before the full-scale invasion) and then relatively quickly the price has stabilized.



On this plot, one can observe that the price is highly correlated with the export volume, which is actually pretty expected. The same observation are going to be seen while analyzing the results of fixed affects model.

For the internal market there is the same data. Let us construct the plot of price and category to see if there were some significant changes:



One can observe that the price of the dry pet food have started to increase from March 2022. It can be due to the fact that there was some companies, apparently on the frontline regions of Ukraine, that stopped producing pet food with the beginning of the war. So companies in other regions had the opportunity to raise prices because of increase in demand.

4 Methodology

The main methods that are going to be used in the research are the ones that are most commonly used to analyse penal data, such as Fixed Effects Models and Random Effects Models.

For testing the hypothesis about decrease in export volume due to the start of full invasion the **fixed effects model** and **random effects model** are used. A new dummy variable *war* are created to divide the data into two periods. Now the following model is created:

$$y_{it} = \beta_0 + \delta_0 war + \beta_1 x_{it} + a_i + u_{it}$$

For the following project, in the notation y_{it}, i denotes the particular brand of the pet goods and t denotes the time period. The variable a_i captures all time-constant factors that affect y_{it}. a_i is called an unobserved effect. u_{it} is often called idiosyncratic error or time-varying error, because it represents unobserved factors that change over time and affect y_{it}. Fixed effects removes the effect of time-invariant characteristics so one can assess the effect of the predictors that are dependent on time on the outcome variable.

In the random-effects model the assumption is that the effect varies from one observation to the next one, and that the observation are randomly sampled from the population. Also, it assumes that the true effect could vary from observation to observation because of heterogeneity. Thus, in such analysis time-invariant variables can be added to the model. For example, the effect of russian invasion can be different based on category of observation or manufacturer, as distinct manufacturers could handle the difficulties caused by the war differently.

In this project, two models are to be implemented. Then, the authors plan to determine which one is the better fit for the data. In the data there are 20 entities with complete data that correspond to Ukrainian export. To implement Fixed Effects Model and Random Effects Model plm model in R was used.

To determine which method among Fixed Effect Model and Random Effects model is better, the authors used Hausman test. The null hypothesis is that the individual effects are not correlated with the x_{it}. If the p-value is significant then fixed effects would be used, if not random effects would be used. Hausman test statistics is calculated the following way:

$$H = \frac{\beta_0 - \beta_1}{Var(\beta_0) - Var(\beta_1)}$$

The following hypotheses are constructed:

- H_0: The appropriate model is *Random effects.* There is no correlation between the error term and the independent variables in the panel data model.
- H_1: The appropriate model is *Fixed effects.* The correlation between the error term and the independent variables in the panel data model is statistically significant.

5 Results

5.1 Export

The Fixed Effect Model was constructed:

$$log(export\_volume\_units) = \beta_0 + \beta_1 war + \beta_2 price + a_i + u_{it}$$

The following results are obtained:

```
Coefficients:
       Estimate Std. Error t-value  Pr(>|t|)
war   -0.182369   0.063007 -2.8944  0.003942 **
price  0.255350   0.024162 10.5683 < 2.2e-16 ***
```

We get the p-value of 0.0039 while testing the hypotheses:

1. $H_0 : \beta_4 = 0$
2. $H_1 : \beta_4 \neq 0$

Under the null hypothesis we assume that, once all of the other independent variables have been accounted for, war has no impact on the export volume.

To test this hypotheses *t-statistics* is used:

$$t = \frac{\hat{\beta}_1}{SE(\hat{\beta}_1)}$$

- $\hat{\beta}_1$ is the estimated coefficient for the war variable.
- $SE(\hat{\beta}_1)$ is the standard error of the coefficient estimate.

t-statistic measures the number of standard deviations that the estimate $\hat{\beta}_1$ is away from zero. Intuitively, if the absolute value of the t-statistic is large, it indicates that the coefficient estimate $\hat{\beta}_1$ is far from zero relative to its standard error.

The $t-statistic$ is smaller than $t_{0.025}$ and p-value is small enough, so the null hypothesis would be rejected and one can say that the war has an impact on the export volume. The estimated coefficient near the war variable is -0.182 which means that the war has decreased the export volume by approximately 18% on average.

The *Random Effect Model* was constructed in the same way but it also takes into account the variables that are constant over time. The following results are obtained:

```
                               Estimate Std. Error z-value  Pr(>|z|)
(Intercept)                   7.0309115  0.7064366  9.9526 < 2.2e-16 ***
war                          -0.1763630  0.0672160 -2.6238  0.008695 **
price                         0.2382246  0.0602000  3.9572 7.583e-05 ***
export_volume_tons            0.0012536  0.0029561  0.4241  0.671502
category_dry_cats_and_dogsTrue 0.1593171  0.6199424  0.2570  0.797189
category_treatsTrue           0.1192782  0.8583505  0.1390  0.889480
pet_type_catsTrue             1.1383109  0.4562124  2.4951  0.012591 *
price_segment_economyTrue     0.3643420  0.5667240  0.6429  0.520294
price_segment_middleTrue      0.6306134  0.5864345  1.0753  0.282225
package_type_pouchTrue        2.9157245  0.6184863  4.7143 2.426e-06 ***
```

The p-value of the t-test corresponding to war variable is 0.008695 which is also small enough to reject the null hypothesis. The interpretation is pretty the same: the war has decreased the export volume by approximately 17.6% on average.

Let's conduct Hausman test to choose between two models. The p-value of 0.8104 was obtained, so the null hypothesis wouldn't be rejected. So the appropriate model for the project data is *Random Effects Model*.

The next step was to choose appropriate variables. As the model contains a lot of dummy variables, it is convenient to test joint significance of each group of these binary variables. For such task, in Random Effects Model, Mulivariate Wald test is used.

Let Let $\hat{\theta}_n$ be the estimator of P parameters. The distribution of this vector is supposed to converge Normal distribution (as parameters in Random Effects Model are estimated by MLE) with covariance matrix V. Then, by normalizing:

$\sqrt{n}(\hat{\theta}_n - \theta) \rightarrow N(0, V)$.

The test of Q hypotheses on the P parameters is expressed with a $Q \times P$ matrix R:

$H_0 : R\theta = r$

$H_1 : R\theta \neq r$.

The distribution of the test statistic under the null hypothesis is:

$(R\hat{\theta}_n - r)'[R(\hat{V}_n/n)R']^{-1}(R\hat{\theta}_n - r)/Q \quad \longrightarrow F(Q, n-P) \quad \xrightarrow[n\to\infty]{} \chi^2_Q/Q,$

where $\hat{V}_n$ is an estimator of the covariance matrix.

Wald test was conducted several times. The following results were obtained by the first time:

```
Wald test

Model 1: log(export_volume_units) ~ war + price + pet_type_cats + package_type_pouch +
    export_volume_tons + category_dry_cats_and_dogs + category_treats +
    price_segment_economy + price_segment_middle
Model 2: log(export_volume_units) ~ war + price + pet_type_cats + package_type_pouch +
    export_volume_tons + price_segment_economy + price_segment_middle
  Res.Df Df  Chisq Pr(>Chisq)
1    590
2    592 -2 0.0681     0.9665
```

```
Wald test

Model 1: log(export_volume_units) ~ war + price + pet_type_cats + package_type_pouch +
    export_volume_tons + category_dry_cats_and_dogs + category_treats +
    price_segment_economy + price_segment_middle
Model 2: log(export_volume_units) ~ war + price + pet_type_cats + package_type_pouch +
    export_volume_tons + category_dry_cats_and_dogs + category_treats
  Res.Df Df  Chisq Pr(>Chisq)
1    590
2    592 -2 1.1637     0.5589
```

Binary variables for category were dropped from the model, as the p-value is large, thus the null hypothesis cannot be rejected.

The following model was obtained:

```
Balanced Panel: n = 20, T = 30, N = 600

Effects:
                 var std.dev share
idiosyncratic 0.5599  0.7483 0.419
individual    0.7763  0.8811 0.581
theta: 0.8468

Residuals:
     Min.   1st Qu.    Median   3rd Qu.      Max.
-5.507851 -0.342806  0.090513  0.407251  2.300157

Coefficients:
                          Estimate Std. Error z-value  Pr(>|z|)
(Intercept)              7.1525965  0.4960203 14.4200 < 2.2e-16 ***
war                     -0.1762512  0.0671754 -2.6237  0.008697 **
price                    0.2373225  0.0599614  3.9579 7.560e-05 ***
pet_type_catsTrue        1.1091812  0.4165272  2.6629  0.007746 **
package_type_pouchTrue   2.8125010  0.4424363  6.3568 2.059e-10 ***
export_volume_tons       0.0013748  0.0029208  0.4707  0.637868
price_segment_economyTrue 0.3798398 0.5305040  0.7160  0.473993
price_segment_middleTrue 0.6147057  0.5159668  1.1914  0.233510
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares:    429.03
Residual Sum of Squares: 331.27
R-Squared:      0.22786
Adj. R-Squared: 0.21873
Chisq: 174.702 on 7 DF, p-value: < 2.22e-16
```

Variable *export_volume_tons* was dropped for the model. Then Wald test was performed for the second time.

The following results of Wald test were obtained by the second time:

```
Wald test

Model 1: log(export_volume_units) ~ war + price + pet_type_cats + package_type_pouch +
    price_segment_economy + price_segment_middle
Model 2: log(export_volume_units) ~ war + price + pet_type_cats + package_type_pouch
  Res.Df Df  Chisq Pr(>Chisq)
1    593
2    595 -2 1.6394     0.4406
```

Binary variables for price segment were dropped from the model, as the p-value is large, thus the null hypothesis cannot be rejected.

The following model was obtained:

```
Balanced Panel: n = 20, T = 30, N = 600

Effects:
                 var std.dev share
idiosyncratic 0.5591  0.7477 0.451
individual    0.6803  0.8248 0.549
theta: 0.8367

Residuals:
    Min.  1st Qu.   Median  3rd Qu.     Max.
-5.57827 -0.34133  0.11011  0.41207  2.22360

Coefficients:
                        Estimate Std. Error z-value  Pr(>|z|)
(Intercept)             7.590126   0.316144 24.0085 < 2.2e-16 ***
war                    -0.188781   0.062955 -2.9987  0.002712 **
price                   0.265652   0.023297 11.4029 < 2.2e-16 ***
pet_type_catsTrue       1.072831   0.384195  2.7924  0.005232 **
package_type_pouchTrue  2.731033   0.410950  6.6457 3.019e-11 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares:    434.91
Residual Sum of Squares: 333.48
R-Squared:      0.23324
Adj. R-Squared: 0.22808
Chisq: 180.988 on 4 DF, p-value: < 2.22e-16
```

The next step is to choose between this model and the Fixed Effects model. Hausman test was conducted one more time. The following results were obtained:

```
         Hausman Test

data:  log(export_volume_units) ~ war + price
chisq = 2.5853, df = 2, p-value = 0.2745
alternative hypothesis: one model is inconsistent
```

The p-value is large enough to conclude that after dropping insignificant variables, Random Effects Model is still more appropriate. From the model, one can draw such conclusion: the war has decreased the export volume by approximately 18.8% on average.

5.2 Internal Market

Similar approach was used to analyse internal market change. To find detailed description of the methods, look in the Export subsection. First of all, the Random Effects Model was constructed. The following result was obtained:

```
Balanced Panel: n = 37, T = 30, N = 1110

Effects:
                 var std.dev share
idiosyncratic 0.2612  0.5111 0.232
individual    0.8664  0.9308 0.768
theta: 0.9003

Residuals:
     Min.   1st Qu.    Median   3rd Qu.      Max.
-4.268846 -0.212960  0.065475  0.287367  1.660376

Coefficients: (2 dropped because of singularities)
                                        Estimate  Std. Error z-value  Pr(>|z|)
(Intercept)                          11.89257795  0.80006014 14.8646 < 2.2e-16 ***
volume_tons                           0.00307259  0.00040444  7.5972 3.026e-14 ***
war                                   0.08435035  0.03316940  2.5430 0.0109900 *
price                                 0.00425672  0.00346486  1.2285 0.2192447
category_dry_cats_and_dogsTrue       -2.75015047  0.42394234 -6.4871 8.751e-11 ***
category_treatsTrue                  -2.24439240  0.72895877 -3.0789 0.0020777 **
manufacturer_PP_Hal_Euro_ContractTrue  0.18756308  0.90481619  0.2073 0.8357802
manufacturer_TOV_KORMOTECHTrue       -0.56315244  0.59958448 -0.9392 0.3476086
manufacturer_TOV_PTF_TECHNOTrue      -1.90236774  0.94983814 -2.0028 0.0451952 *
pet_type_catsTrue                     1.19591837  0.32181619  3.7162 0.0002023 ***
price_segment_economyTrue            -0.54485175  0.48877266 -1.1147 0.2649642
price_segment_middleTrue              1.49922921  0.62439345  2.4011 0.0163460 *
package_type_canTrue                 -2.09831447  0.60053684 -3.4941 0.0004757 ***
package_type_glass_jarTrue           -2.83346322  0.93566294 -3.0283 0.0024594 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    374.36
Residual Sum of Squares: 286.36
R-Squared:      0.23506
Adj. R-Squared: 0.22599
Chisq: 336.801 on 13 DF, p-value: < 2.22e-16
```

Next, Wald test was used to choose the appropriate variables. First of all, the group of binary variables for *manufacturer* was tested. The following results were obtained:

```
Wald test

Model 1: log(volume_units) ~ volume_tons + war + price + category_dry_cats_and_dogs +
    category_treats + manufacturer_PP_Hal_Euro_Contract + manufacturer_TOV_KORMOTECH +
    manufacturer_TOV_PTF_TECHNO + pet_type_cats + price_segment_economy +
    price_segment_middle + package_type_can + package_type_glass_jar +
    package_type_pack + package_type_pouch
Model 2: log(volume_units) ~ volume_tons + war + price + category_dry_cats_and_dogs +
    category_treats + pet_type_cats + price_segment_economy +
    price_segment_middle + package_type_can + package_type_glass_jar +
    package_type_pack + package_type_pouch
  Res.Df Df Chisq Pr(>Chisq)
1   1096
2   1099 -3 4.699     0.1952
```

This group of binary variables was dropped from the model. After dropping, the model obtained is the following:

```
Balanced Panel: n = 37, T = 30, N = 1110

Effects:
                 var std.dev share
idiosyncratic 0.2612  0.5111 0.219
individual    0.9342  0.9665 0.781
theta: 0.9039

Residuals:
     Min.   1st Qu.    Median   3rd Qu.      Max.
-4.313378 -0.213261  0.066267  0.279950  1.660245

Coefficients: (2 dropped because of singularities)
                                 Estimate Std. Error z-value  Pr(>|z|)
(Intercept)                    11.0418629  0.4959219 22.2653 < 2.2e-16 ***
volume_tons                     0.0030428  0.0004053  7.5074 6.031e-14 ***
war                             0.0843575  0.0331638  2.5437 0.0109697 *
price                           0.0043366  0.0034654  1.2514 0.2107888
category_dry_cats_and_dogsTrue -3.0266201  0.3950753 -7.6609 1.847e-14 ***
category_treatsTrue            -2.7795779  0.7029056 -3.9544 7.672e-05 ***
pet_type_catsTrue               1.2241705  0.3305443  3.7035 0.0002126 ***
price_segment_economyTrue       0.0671679  0.4123240  0.1629 0.8705966
price_segment_middleTrue        2.0494289  0.5548524  3.6936 0.0002211 ***
package_type_canTrue           -2.2836255  0.5863102 -3.8949 9.824e-05 ***
package_type_glass_jarTrue     -2.4211575  0.7695999 -3.1460 0.0016552 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    370.32
Residual Sum of Squares: 287.01
R-Squared:      0.22497
Adj. R-Squared: 0.21792
Chisq: 319.017 on 10 DF, p-value: < 2.22e-16
```

Binary variable *price_segment_economy* was dropped from the model due to large p-value. Economy and premium price segments are now the base group. The following model was obtained:

```
Balanced Panel: n = 37, T = 30, N = 1110

Effects:
                 var std.dev share
idiosyncratic 0.2612  0.5111 0.224
individual    0.9047  0.9512 0.776
theta: 0.9024

Residuals:
     Min.   1st Qu.    Median   3rd Qu.      Max.
-4.319159 -0.212805  0.066857  0.279268  1.660043

Coefficients: (2 dropped because of singularities)
                                 Estimate  Std. Error z-value  Pr(>|z|)
(Intercept)                   11.09048925  0.38959713 28.4666 < 2.2e-16 ***
volume_tons                    0.00304938  0.00040413  7.5455 4.505e-14 ***
war                            0.08446375  0.03315929  2.5472  0.010859 *
price                          0.00429477  0.00346101  1.2409  0.214643
category_dry_cats_and_dogsTrue -3.02618875  0.38885731 -7.7823 7.124e-15 ***
category_treatsTrue           -2.79969438  0.68085759 -4.1120 3.922e-05 ***
pet_type_catsTrue              1.22386566  0.32538021  3.7613  0.000169 ***
price_segment_middleTrue       2.00684626  0.48278974  4.1568 3.228e-05 ***
package_type_canTrue          -2.28180923  0.57698230 -3.9547 7.662e-05 ***
package_type_glass_jarTrue    -2.40254883  0.74915888 -3.2070  0.001341 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    372
Residual Sum of Squares: 287.26
R-Squared:      0.22781
Adj. R-Squared: 0.22149
Chisq: 324.517 on 9 DF, p-value: < 2.22e-16
```

The variable *price* was dropped from the model due to large p-value. Finally, after dropping several variables, such model was obtained:

```
Balanced Panel: n = 37, T = 30, N = 1110

Effects:
                 var std.dev share
idiosyncratic 0.2614  0.5113 0.224
individual    0.9075  0.9526 0.776
theta: 0.9025

Residuals:
     Min.   1st Qu.    Median   3rd Qu.      Max.
-4.324086 -0.215459  0.066394  0.281307  1.653202

Coefficients: (2 dropped because of singularities)
                                 Estimate  Std. Error z-value  Pr(>|z|)
(Intercept)                   11.07687663  0.38995338 28.4056 < 2.2e-16 ***
volume_tons                    0.00340845  0.00028214 12.0806 < 2.2e-16 ***
war                            0.09927392  0.03094401  3.2082 0.0013358 **
category_dry_cats_and_dogsTrue -3.02930926  0.38935840 -7.7803 7.238e-15 ***
category_treatsTrue           -2.81395252  0.68164922 -4.1282 3.657e-05 ***
pet_type_catsTrue              1.23261610  0.32572827  3.7842 0.0001542 ***
price_segment_middleTrue       2.04487586  0.48243745  4.2386 2.249e-05 ***
package_type_canTrue          -2.27944621  0.57773665 -3.9455 7.964e-05 ***
package_type_glass_jarTrue    -2.40187133  0.75014233 -3.2019 0.0013653 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    371.88
Residual Sum of Squares: 287.64
R-Squared:      0.22653
Adj. R-Squared: 0.22091
Chisq: 322.45 on 8 DF, p-value: < 2.22e-16
```

Also, Fixed Effects Model was constructed. The result:

```
Balanced Panel: n = 37, T = 30, N = 1110

Residuals:
     Min.   1st Qu.    Median   3rd Qu.      Max.
-4.136812 -0.202054  0.053407  0.273565  1.770359

Coefficients:
              Estimate Std. Error t-value  Pr(>|t|)
volume_tons 0.00336034 0.00030068 11.1757 < 2.2e-16 ***
war         0.09986127 0.03098092  3.2233  0.001305 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    318.14
Residual Sum of Squares: 280
R-Squared:      0.11987
Adj. R-Squared: 0.088646
F-statistic: 72.9356 on 2 and 1071 DF, p-value: < 2.22e-16
```

Hausman test was performed to decide between fixed effects and random effects model.

```
	Hausman Test

data:  log(volume_units) ~ volume_tons + war + category_dry_cats_and_dogs +  ...
chisq = 0.21416, df = 2, p-value = 0.8985
alternative hypothesis: one model is inconsistent
```

The p-value is large enough to conclude that Random Effects Model is still more appropriate in this case. From the model, one can draw such conclusion: the war has increased the sales by approximately 9.92% on average. This may happen due to the panic among population and efforts to buy as much goods as possible to simply survive.

To provide more detailed analysis, we also decided to look not on the sold units themselves, but on the increase of sold units during each month. So, we created new column *units_increase*. After constructing Random Effects Model, the obtained result turned out to be quite unusual:

```
Balanced Panel: n = 37, T = 30, N = 1110

Effects:
                  var  std.dev share
idiosyncratic 1.061911 1.030491 0.994
individual    0.006778 0.082330 0.006
theta: 0.08387

Residuals:
    Min.  1st Qu.   Median  3rd Qu.     Max.
-1.50720 -0.39580 -0.13223  0.10993 11.33844

Coefficients: (2 dropped because of singularities)
                                      Estimate  Std. Error z-value Pr(>|z|)
(Intercept)                         0.07550162  0.18004460  0.4193   0.6750
volume_tons                         0.00017516  0.00042340  0.4137   0.6791
war                                 0.28701134  0.06469894  4.4361 9.16e-06 ***
price                              -0.00421067  0.00502380 -0.8381   0.4019
category_dry_cats_and_dogsTrue     -0.03292790  0.09401385 -0.3502   0.7262
category_treatsTrue                 0.07928928  0.16735345  0.4738   0.6357
manufacturer_PP_Hal_Euro_ContractTrue -0.03632053  0.19952637 -0.1820   0.8556
manufacturer_TOV_KORMOTECHTrue      0.02820433  0.13238592  0.2130   0.8313
manufacturer_TOV_PTF_TECHNOTrue     0.05634037  0.20984807  0.2685   0.7883
pet_type_catsTrue                   0.00491344  0.07448325  0.0660   0.9474
price_segment_economyTrue           0.03514662  0.11157551  0.3150   0.7528
price_segment_middleTrue           -0.12952842  0.15186895 -0.8529   0.3937
package_type_canTrue               -0.02031021  0.13262961 -0.1531   0.8783
package_type_glass_jarTrue         -0.09074235  0.20610894 -0.4403   0.6597
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    1199.5
Residual Sum of Squares: 1171
R-Squared:      0.023748
Adj. R-Squared: 0.012169
Chisq: 26.6614 on 13 DF, p-value: 0.013838
```

The only significant variable in this model is *war*, which is time-dependent. From such result one might conclude that Fixed Effects Model would be a better fit. The result for Fixed Effects Model (all time-dependent variables included):

```
Balanced Panel: n = 37, T = 30, N = 1110

Residuals:
     Min.   1st Qu.    Median   3rd Qu.      Max.
-1.772141 -0.403682 -0.098432  0.159640 10.874375

Coefficients:
               Estimate  Std. Error t-value  Pr(>|t|)
volume_tons  0.00235620  0.00084972  2.7729  0.005652 **
war          0.29859948  0.06694248  4.4605 9.039e-06 ***
price       -0.01276198  0.00702877 -1.8157  0.069700 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    1164.6
Residual Sum of Squares: 1136.2
R-Squared:      0.024376
Adj. R-Squared: -0.011184
F-statistic: 8.91123 on 3 and 1070 DF, p-value: 7.7982e-06
```

From this result, we can see that the war has increased the increase of sales by approximately 0.299 on average. As before, such situation may happen due to the widespread panic all over Ukraine.

Also, the result of Hausman test confirms the hypothesis of Fixed Effects Model being better in case of sales increase:

```
        Hausman Test

data:  units_increase ~ volume_tons + war + price + category_dry_cats_and_dogs +  ...
chisq = 8.7783, df = 3, p-value = 0.03239
alternative hypothesis: one model is inconsistent
```

6 Conclusions

On this stage of the project, the hypothesis about the dependence between the start of full-invasion and export volume was tested. And indeed it was shown that the war has dependence on the export volume of Ukraine. It has decreased the export volume by approximately 18.8%. Also the package type of pouch has positive impact on the units exported comparing to other package types. It can be because pouches are more easily to deliver than, for example, cans.

As for the internal market, a lot of characteristics, of product influenced the sales:

1. Holding all factors zero, the sales are $e^{11.08}$ on average;
2. On average, the increase of sales by one tonne increases the sales in units by approximately 3.4%;
3. On average, the war increased the sales by approximately 9.92%;
4. On average, sales of dry food are less than sales of wet food by approximately 302.9%;
5. On average, sales of treats are less than sales of wet food by approximately 281.3%;
6. On average, sales of food for cats are greater than sales of food for dogs by approximately 123.3%;
7. On average, sales of middle price segment are greater than sales of premium and economy segments by approximately 204.5%;
8. On average, sales of cans are less than sales of pouches, sacks and packs by approximately 227.9%;
9. On average, sales of glass jars are less than sales of pouches, sacks and packs by approximately 240.2%. This may be due to the inconvenience of glass transportation;

One can observe that certain products had much greater sales than others.

As for increase of sales, one may conclude that price was not the most significant factor at all. But war increased the increase of sales by 0.299 on average.

Overall, it can be stated that the russian invasion did have impact on Ukrainian market for pets food. As for export, the war had negative impact on sales, but the sales increased on Ukrainian internal market. It could be because of the panic around and the fact that people were buying everything.

References

[1] A full course in econometrics-undergraduate level. Retrieved from: https://www.youtube.com/playlist?list=PLwJRxp3blEvb7P-7po9AxuBwquPv75LjU.

[2] Analysis of the Future Development Trend of the Pet Industry. Retrived from: https://drive.google.com/drive/folders/1lfJLwo6ml1bSzpN6nwQGrId2EK1QCzQL?usp=share$_link$.

[3] Europe Pet Food Market Size and Share Analysis. Retrived from: https://www.mordorintelligence.com/industry-reports/pet-food-market-in-europe-industry.