# 

KPI Analysis in one of the agro-industrial enterprises

# Introduction

The analysis was carried out in the R package. To analyse the impact of KPI indicators on profitability by units of one of the agro-industrial enterprises, data from budgets for the period from 2019 to 2022 were exported. The data were converted into panel data and mathematical calculations were performed (profitability, average subsidies per day per head were calculated). The panel is balanced and has a number of objects N=6 with the number of time cycles T=48. The panel is divided into 4 years of 12 months and into the following subdivisions:

1. Shchuchye
2. Kolybelka housing estate
3. Zaluzhnoye housing estate
4. Petrovskoye Residential Complex
5. Volchanskoye housing estate
6. MTF Shchuchye

# Results

The independent variables and their units of measurement are presented in Table 1.1. The dependent variable is profitability (Renrtabel)

|  |  |  |
| --- | --- | --- |
| **Name of variables in the original** | **Unit of measurement** | **Name of variables in R** |
| Gross production in the standard | т | GrossProduction\_t |
| marketability |  | Marketability |
| Farm loading |  | Farm\_capacity |
| Culling of cows |  | Culling |
| Milk yield in standard 1 FC | kg/goal | YealdPerCow |
| Feeds | rub/kg | FoodPerkg |
| Operating costs | rub/kg | OperationCostsPerkg |
| Cost of repairing the herd | rub/kg | Cow\_repear |
| Amortisation | rub/kg | Amortisation |
| Interest on loans | rub/kg | Credit\_procent |
| Sales price of milk in physical weight | rub/kg | Price\_inrealweight |
| Sales price of milk in hundredweight | rub/kg | Price\_instandartweight |
| Subsidies | rub./goal/day. | Subsidy |
| IOFC | rub./goal/day. | IOFC |

Table 1.1 - List of dependent variables

First, let's look at the trend in profitability by month from 2019 to 2022 in all divisions (Figure 1.1). In general, all divisions show a positive trend in all four years, although the month-to-month variability is quite high. Between 2019 and 2020, a strong dip is noticeable. This is due to a simultaneous increase in production costs and a fall in the realised price of milk in physical weight (Figures 1.2, 1.3). While in the first two years the main reason for the cost increase was depreciation and operating costs, in 2021 and 2022 feed prices held high, although offset by the high price of milk by weight (Figures 1.4-1.6). Figure 1.7 shows differences in profitability trends and trend slope angles between units, suggesting heterogeneity in the sample, albeit not as obvious. This heterogeneity should show up for some dependent variables when estimating the model using ordinary ANC.

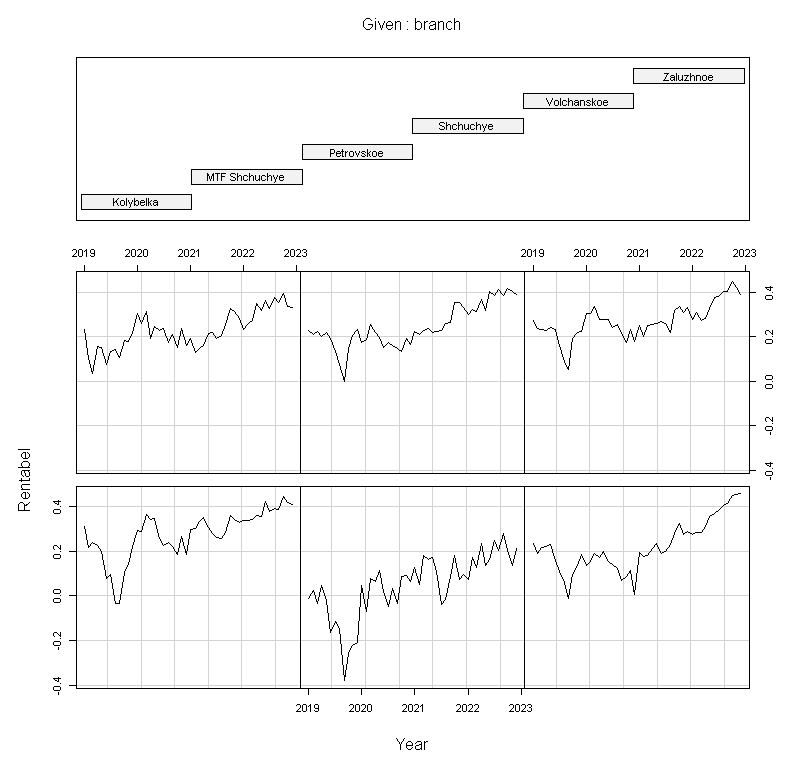


Figure 1.1 - Dynamics of profitability by month in different divisions (from left to right, starting from the bottom row)

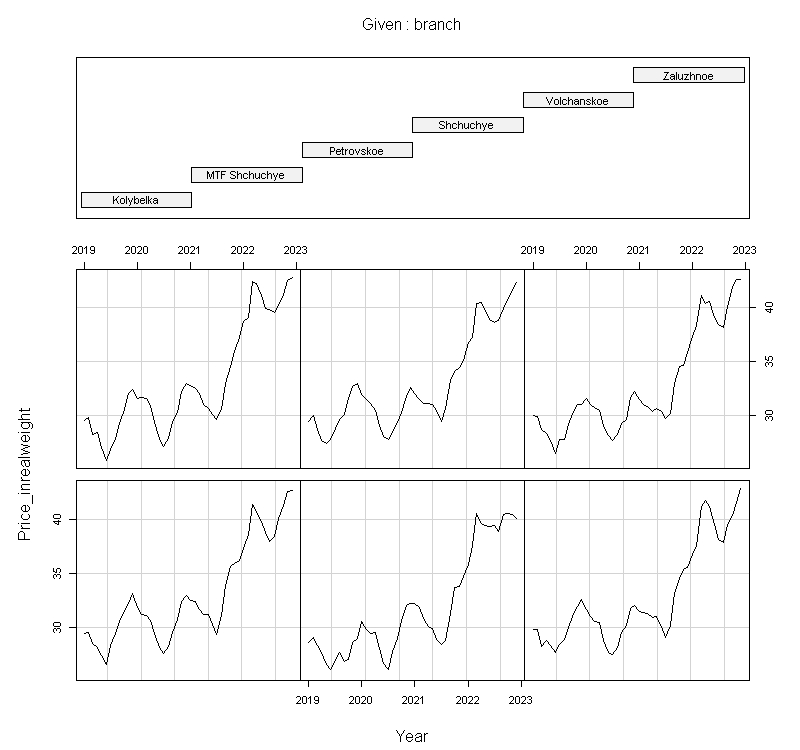


Figure 1.2 - Dynamics of milk selling price by month in different divisions (from left to right, starting from the bottom row)

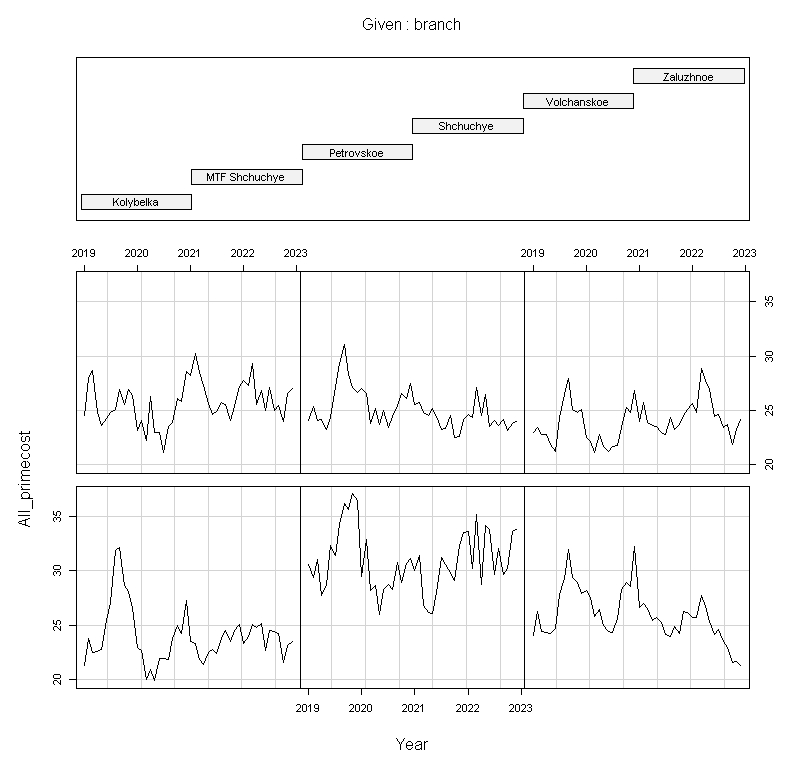


Figure 1.3 - Cost of sales by month in different divisions (from left to right, starting from the bottom line)

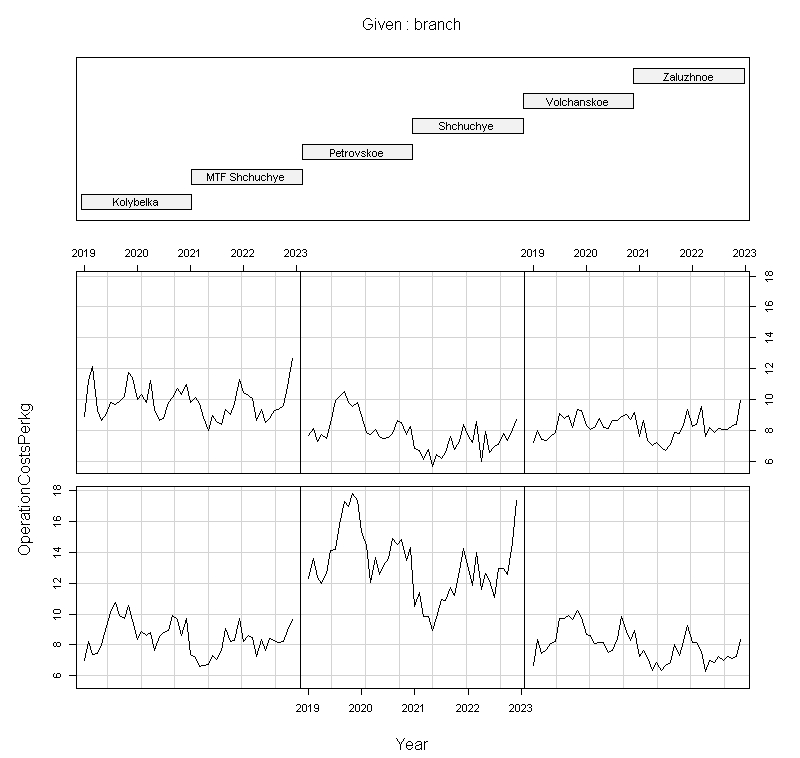


Figure 1.4 - Dynamics of operating costs by month in different divisions (from left to right, starting from the bottom row)

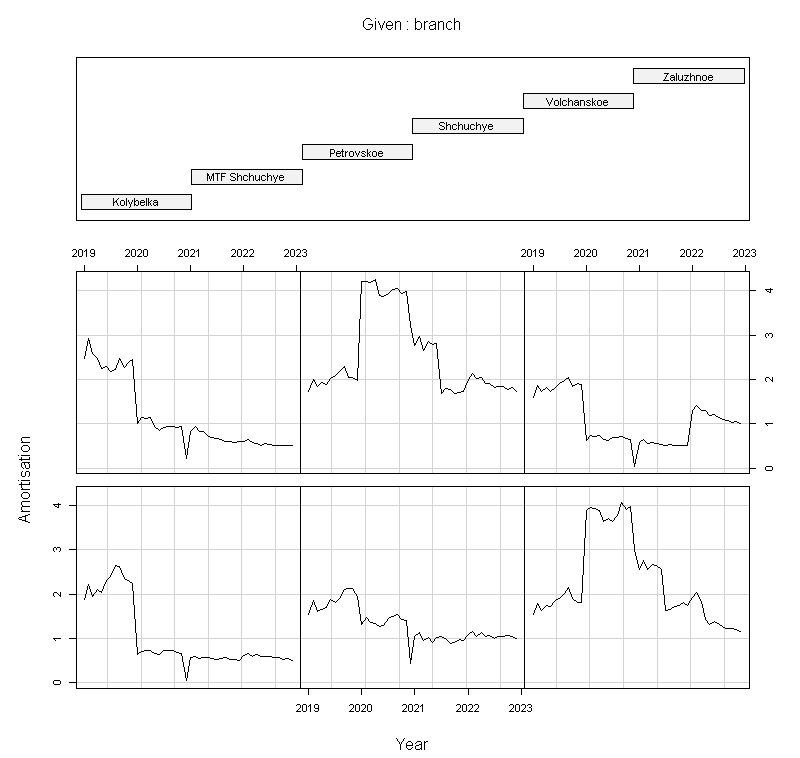


Figure 1.5 - Dynamics of depreciation by year in different divisions (from left to right, starting from the bottom row)

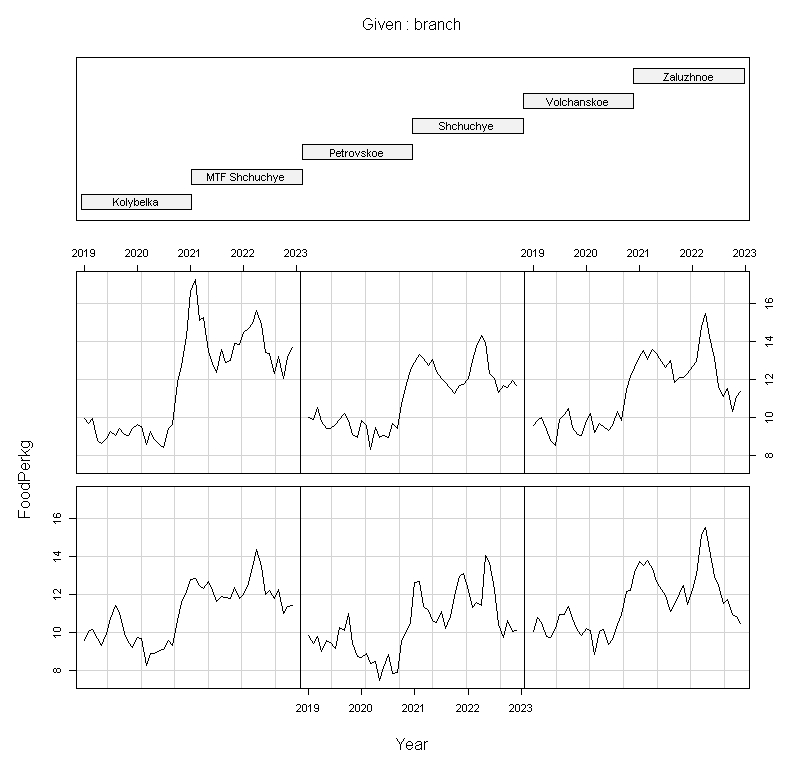


Figure 1.6 - Dynamics of feed costs by year in different divisions (from left to right, starting from the bottom row)

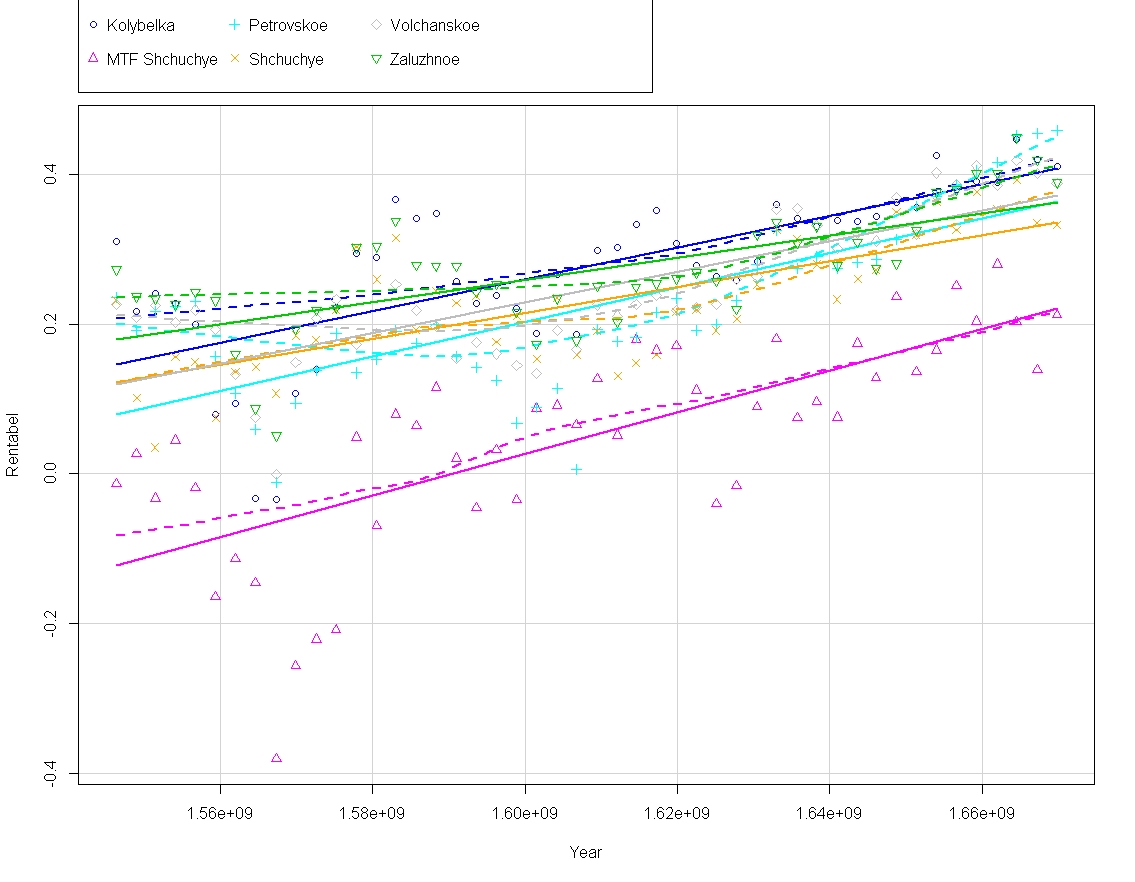


Figure 1.7 - Dynamics of profitability by month in different divisions (dotted lines indicate dynamics, straight lines - slope of the trend)

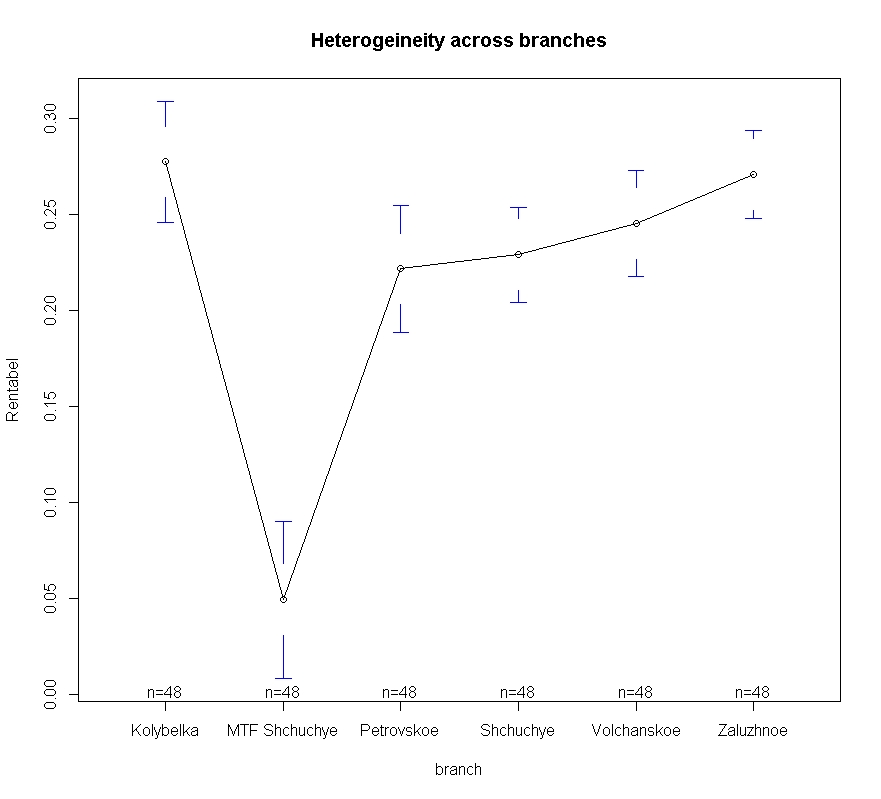


Figure 1.8 - Heterogeneity between different divisions

Figure 1.8 above shows the heterogeneity of the sample between the different divisions in terms of profitability, however, due to limited facilities, this visual heterogeneity may be spurious. This question will be more accurately answered by the Breusch-Pagan test, which will be conducted a little later. Figure 1.9 also shows the heterogeneity of profitability, but by month for the whole period. In this case, the heterogeneity is already more obvious, and there are signs of autocorrelation.

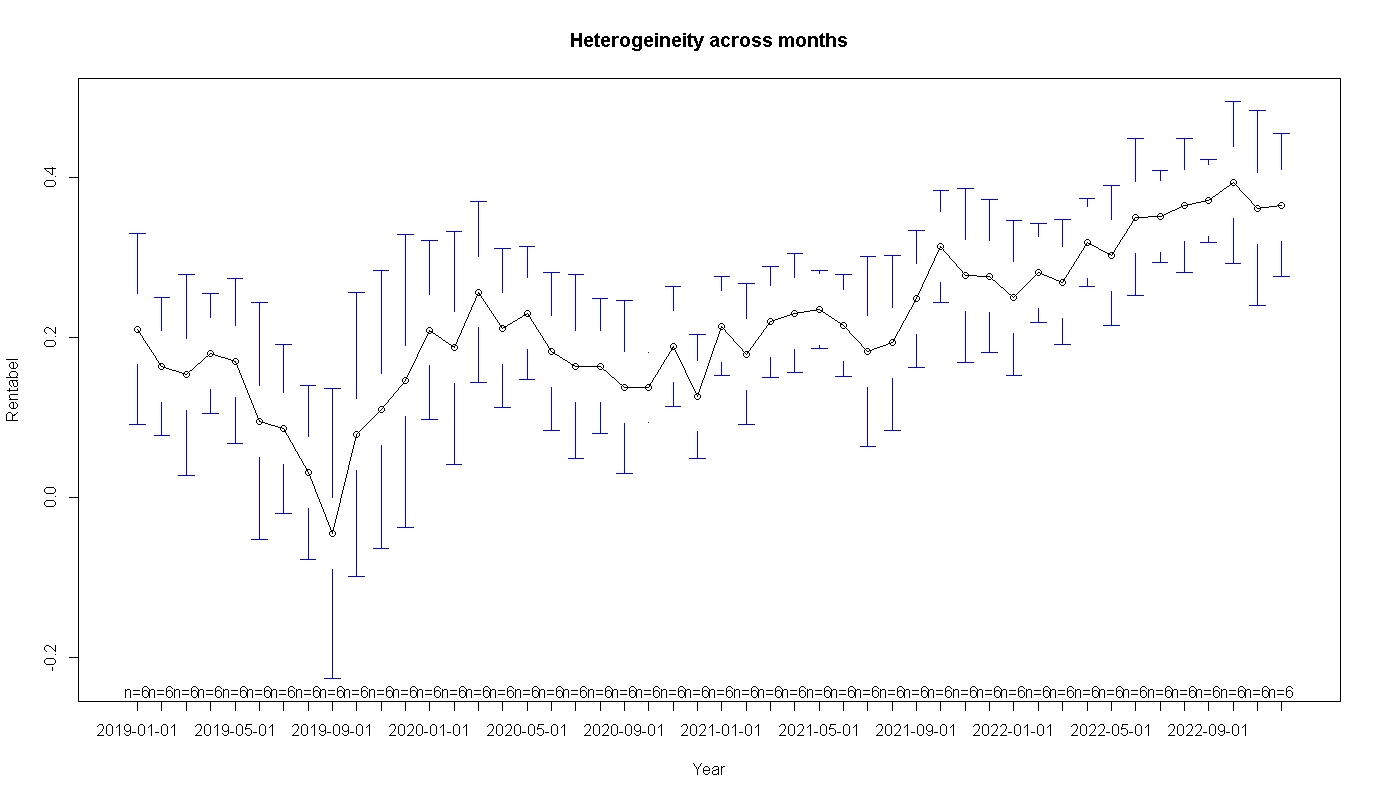
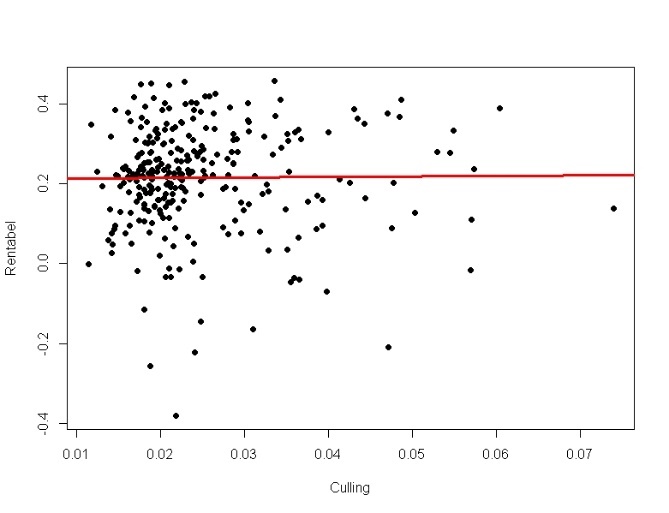
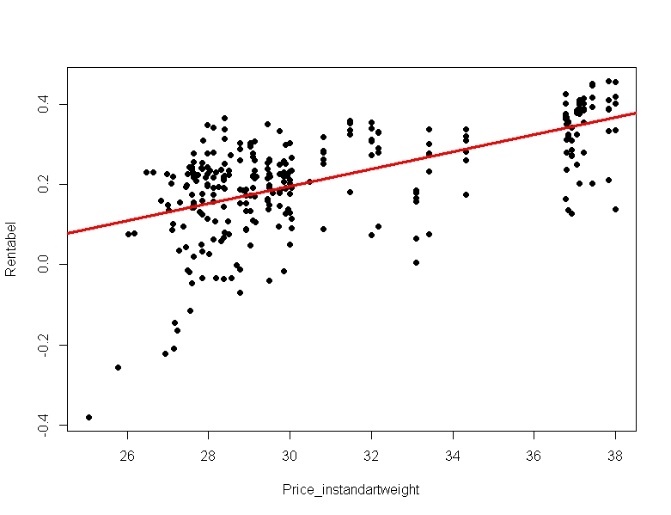


Figure 1.9 - Heterogeneity of profitability with respect to time.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | estimate | std.error | statistic | p.value |  |
| (Intercept) | 0,113291569 | 0,140717287 | 0,80510058 | 0,421462265 |  |
| Gross production in the standard | -5,28E-05 | 9,68E-06 | -5,455650011 | 1,09E-07 | \*\*\* |
| marketability | 0,398912079 | 0,12045921 | 3,311594692 | 0,001052698 | \*\* |
| Farm loading | 0,073438386 | 0,007944693 | 9,243703225 | 7,04E-18 | \*\*\* |
| Culling of cows | -0,139218029 | 0,156455329 | -0,889826064 | 0,374342736 |  |
| Milk yield in standard 1 FC | -0,00972936 | 0,001381898 | -7,040577185 | 1,55E-11 | \*\*\* |
| Feeds | -0,017049877 | 0,002442641 | -6,980100283 | 2,24E-11 | \*\*\* |
| Operating costs | -0,033805438 | 0,001222646 | -27,64940231 | 4,01E-81 | \*\*\* |
| Cost of repairing the herd | -0,024554195 | 0,00125178 | -19,61542393 | 4,66E-54 | \*\*\* |
| Amortisation | -0,033917414 | 0,001674747 | -20,25226567 | 2,62E-56 | \*\*\* |
| Interest on loans | -0,031292883 | 0,002268418 | -13,79502401 | 3,31E-33 | \*\*\* |
| Sales price of milk in physical weight | 0,008628915 | 0,001780584 | 4,846114624 | 2,11E-06 | \*\*\* |
| Sales price of milk in hundredweight | 0,002776396 | 0,001692183 | 1,64071791 | 0,102007579 |  |
| Subsidies | 0,000523076 | 0,000412212 | 1,268947718 | 0,205540546 |  |
| IOFC | 0,000434243 | 7,39E-05 | 5,872636418 | 1,24E-08 | \*\*\* |

Table 1.2 - Descriptive statistics of MNC regression (Pool)

Regression estimation showed the significance of most of the coefficients. The coefficients of determination are also close to unity (**Multiple R-squared: 0.978, Adjusted R-squared: 0.9767**). The p-value of the entire for the **F-statistic** is significant at any reasonable level of significance (**F-statistic: 752.5 on 16 and 271 DF, p-value: < 2.2e-16**), indicating that all regression coefficients are different from zero. To illustrate the insignificant coefficients, we can look at the plots of these variables (Figure 2.0). They show that for these coefficients, the problem of heterogeneity across individual effects and time is not solved by conventional ANC.

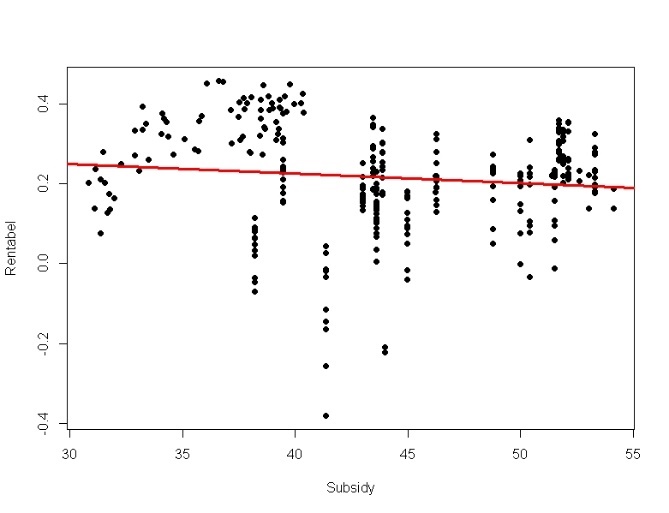


Figure 2.0 - Heterogeneity between profitability and attrition rate, standard weight milk price and subsidies

For the FE model, the values of some of the estimates, standard errors and t-statistics changed, although the significance of the coefficients remained essentially the same (Table 1.3). **R-Squared: 0.9686, Adj. R-Squared: 0.96612, F-statistic: 512.78 on 16 and 266 DF, p-value: < 2.22e-16**. The constants for each subdivision are presented in Table 1.4, from which individual differences can be seen.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| term | estimate | std.error | statistic | p.value |  |
| Gross production in the standard | -9,39E-05 | 1,88E-05 | -4,986942104 | 1,10E-06 | \*\*\* |
| marketability | 0,343741038 | 0,118326171 | 2,905029667 | 0,0039785 | \*\* |
| Farm loading | 0,079895978 | 0,008290278 | 9,637310409 | 4,59E-19 | \*\*\* |
| Culling of cows | -0,153500286 | 0,153394527 | -1,000689461 | 0,317879618 |  |
| Milk yield in standard 1 FC | -0,007846418 | 0,001480923 | -5,298330563 | 2,44E-07 | \*\*\* |
| Feeds | -0,01657613 | 0,002399608 | -6,907848726 | 3,57E-11 | \*\*\* |
| Operating costs | -0,033770877 | 0,001266724 | -26,66001108 | 2,38E-77 | \*\*\* |
| Cost of repairing the herd | -0,02533271 | 0,001468689 | -17,2485242 | 2,34E-45 | \*\*\* |
| Amortisation | -0,038315335 | 0,001939788 | -19,75233144 | 3,27E-54 | \*\*\* |
| Interest on loans | -0,032653579 | 0,002319969 | -14,07500608 | 4,62E-34 | \*\*\* |
| Sales price of milk in physical weight | 0,008507196 | 0,001754097 | 4,849900018 | 2,10E-06 | \*\*\* |
| Sales price of milk in hundredweight | 0,000587707 | 0,001728444 | 0,340020753 | 0,734107239 |  |
| Subsidies | 0,000359374 | 0,000429753 | 0,836233208 | 0,403768687 |  |
| IOFC | 0,000496222 | 7,33E-05 | 6,771233498 | 8,05E-11 | \*\*\* |

Table 1.3 - Descriptive statistics of FE (fixed effect) model

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Kolybelka | MTF Shchuchye | Petrovskoe | Shchuchye | Volchanskoe | Zaluzhnoe |
| 0.20101 | 0.19934 | 0.22970 | 0.20262 | 0.22914 | 0.20525 |

Table 1.4 - Display of fixed effects (constants for each unit)

The seemingly identical performance in the estimations of the conventional ISC and FE in the panel data analysis does not allow us to draw definitive conclusions in favour of one or the other model, an F-test is needed to compare both models. Results: **F = 4.3391, df1 = 5, df2 = 268, p-value = 0.0008179**. In this test, when the p-value is significant, we choose the FE model. This suggests that despite the good performance of the coefficients of determination and the significance of most of the coefficients, the estimates of the conventional ISC may not be valid due to the heterogeneity of the sample. On top of that, there is an obvious correlation between regressors in the ISC model, as well as time effects.

Time heterogeneity and possible autocorrelation as well as heteroskedasticity may give wrong results when using the conventional FE-model. However, since the number of independent variables is higher than the number of units, we cannot estimate the model with the usual RE-effect. We will use the RE- Wallace and Hussain model (Table 1.5). The estimates are similar to the previous ones. The same for the coefficients of determination (**R-Squared: 0.97631 Adj. R-Squared: 0.97491**) and the X2 statistic (**Chisq: 11168 on 16 DF, p-value: < 2.22e-16**)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| term | estimate | std.error | statistic | p.value |  |
| (Intercept) | 0,119790948 | 0,140435082 | 0,852998739 | 0,393660004 |  |
| Gross production in the standard | -5,54E-05 | 9,88E-06 | -5,603867492 | 2,10E-08 | \*\*\* |
| marketability | 0,394153107 | 0,12002888 | 3,28381892 | 0,001024107 | \*\* |
| Farm loading | 0,074509337 | 0,007963646 | 9,356184186 | 8,27E-21 | \*\*\* |
| Culling of cows | -0,14393159 | 0,156077819 | -0,922178378 | 0,356435539 |  |
| Milk yield in standard 1 FC | -0,009553098 | 0,001392556 | -6,860119621 | 6,88E-12 | \*\*\* |
| Feeds | -0,016977992 | 0,002436172 | -6,969126291 | 3,19E-12 | \*\*\* |
| Operating costs | -0,033807862 | 0,001228889 | -27,51090943 | 1,30E-166 | \*\*\* |
| Cost of repairing the herd | -0,024637617 | 0,001283565 | -19,19468009 | 4,10E-82 | \*\*\* |
| Amortisation | -0,034337063 | 0,001695621 | -20,25043108 | 3,52E-91 | \*\*\* |
| Interest on loans | -0,031367126 | 0,002264713 | -13,85037784 | 1,27E-43 | \*\*\* |
| Sales price of milk in physical weight | 0,008661659 | 0,001773752 | 4,883240148 | 1,04E-06 | \*\*\* |
| Sales price of milk in hundredweight | 0,002495277 | 0,001695604 | 1,471614846 | 0,141124914 |  |
| Subsidies | 0,000518709 | 0,000413187 | 1,255384051 | 0,209339376 |  |
| IOFC | 0,000440658 | 7,38E-05 | 5,973523897 | 2,32E-09 | \*\*\* |

Table 1.5 - Descriptive statistics of RE (random effect) model

The Hausman test was used to choose between RE and FE and the following results were obtained: **chisq = 19.671, df = 16, p-value = 0.2354**. The high p-value indicates that the RE model is preferred over FE, which is quite expected. There was no increase in the significance of any variables compared to the FE model. We need to conduct the Breusch-Pagan test to validate the choice of RE model. According to its results (**chisq = 0.35382, df = 1, p-value = 0.552**) we can conclude that there are no significant differences between the units, i.e. the objects are homogeneous. This is absolutely logical, as the subdivisions are located quite close to each other and the influence of external factors is almost the same. But we cannot make a choice in favour of the Pool-model, as the problem of heterogeneity is not the only one in this study. All those coefficients that were insignificant in ISC remained insignificant under FE and RE as well, and the difference in parameters is minimal. Figure 1.9 showed significant differences over time, so it is important to do tests for autocorrelation, heteroskedasticity and for detecting simultaneous correlation of residuals.

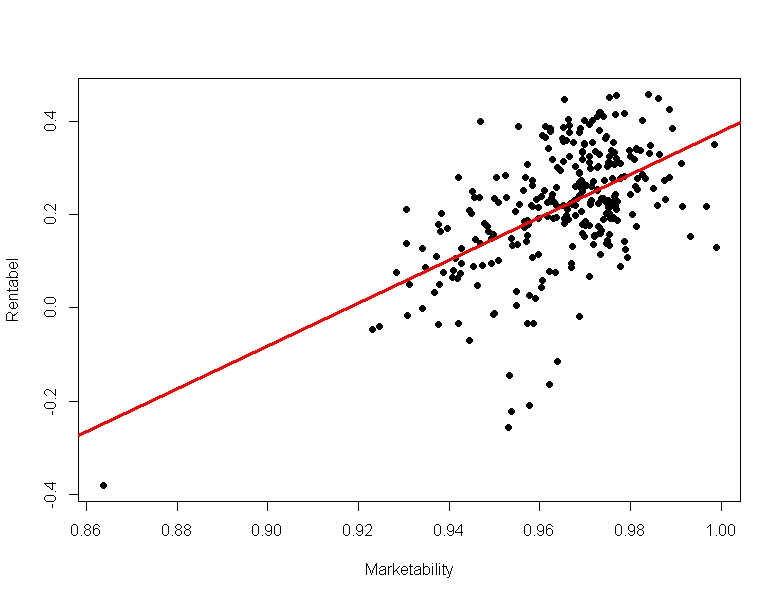
The use of a model with long time series is associated with the problem of simultaneous correlation of residuals between objects, so we may need to replace with robust errors. To diagnose such a problem, we will perform the Breusch-Pagan-Lagrange and Pasaran tests. Both tests signal that there is a correlation between the errors of the objects (**chisq = 124.84, df = 15, p-value < 2.2e-16, z = 8.2445, p-value < 2.2e-16**). We also run the Breusch-Godfrey test to diagnose temporal correlation (**chisq = 133.11, df = 48, p-value = 6.246e-10**). In this case, both tests show identical results for the FE model. The presence of temporal correlation is evident from the low p-value. The **Dickey-Fuller** test shows, albeit weakly, a lack of stationarity (**Dickey-Fuller = -4.4804, Lag order = 2, p-value = 0.01**), which is not surprising given the variability of the variables over time. The Breusch-Pagan heteroscedasticity test showed (**BP = 338, df = 19, p-value < 2.2e-16**) a difference among the variances in the matrix. That is, all tests indicate that conventional MNC regression or FE model cannot be used.

To address these defects we need to estimate the regression with Driscoll-Kraay standard errors , as they are robust to common forms of cross-sectional and temporal dependence, and robust to heteroskedasticity (Table 1.7).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Estimate | Std.Error | t-value | Pr(>|t|) |  |
| Gross production in the standard | -0,00009 | 0,00003 | -3,17910 | 1,651E-03 | \*\* |
| marketability | 0,34374 | 0,31155 | 1,10330 | 2,709E-01 |  |
| Farm loading | 0,07990 | 0,01482 | 5.3275 | 1,546E-07 | \*\*\* |
| Culling of cows | -0,15350 | 0,16140 | -0,95110 | 3,424E-01 |  |
| Standard milk yield per 1 FC | -0,00785 | 0,00202 | -3,89380 | 1,247E-04 | \*\*\* |
| Feeds | -0,01658 | 0,00496 | -3,34460 | 9,417E-04 | \*\*\* |
| Operating costs | -0,03377 | 0,00221 | -15,28200 | 2,200E-16 | \*\*\* |
| Cost of repairing the herd | -0,02533 | 0,00214 | -11,81810 | 2,200E-16 | \*\*\* |
| Amortisation | -0,03832 | 0,00214 | -17,90690 | 2,200E-16 | \*\*\* |
| Interest on loans | -0,03265 | 0,00444 | -7,35790 | 2,287E-12 | \*\*\* |
| Sales price of milk in physical weight | 0,00851 | 0,00309 | 3.3178 | 6,295E-03 | \*\* |
| Sales price of milk in hundredweight | 0,00059 | 0,00294 | 0,20010 | 8,416E-01 |  |
| Subsidies | 0,00036 | 0,00079 | 0,45640 | 6,485E-01 |  |
| IOFC | 0,00050 | 0,00012 | 4,07260 | 6,126E-05 | \*\*\* |

Table 1.7 - Descriptive statistics of FE (fixed effect) model with standard errors Driscoll-Kraay

The table shows that marketability has become insignificant. The standard error of the indicator increased and t-statistics decreased. To understand why this happened, let us look at Figure 2.1. We can see the presence of heteroscedasticity, i.e. the points have different distance from the regression (red line) when moving along the horizontal line. Consequently, when we introduce robust errors, this distance (distance of the points) from the red line becomes the same, i.e. larger. Those standard errors that were closer to the red line due to the large weights become further away. This increases the error, decreases the t-statistic and, consequently, decreases the significance.



Our model is almost ready, we just need to check whether there are variables in our data that change over time but are constant for different units. To do this, we need to use the time fixed effect model.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| term | estimate | std.error | statistic | p.value |  |
| Gross production in the standard | -6,61E-05 | 2,60E-05 | -2,538834056 | 0,011810172 | \* |
| marketability | 0,27887975 | 0,120278879 | 2,318609494 | 0,021330012 | \* |
| Farm loading | 0,062248186 | 0,00827605 | 7,521485233 | 1,35E-12 | \*\*\* |
| Culling of cows | -0,131687297 | 0,182432082 | -0,721842865 | 0,471154133 |  |
| Milk yield in standard 1 FC | -0,006398783 | 0,001883386 | -3,397488453 | 0,000806478 | \*\*\* |
| Feeds | -0,025213574 | 0,003158739 | -7,982165024 | 7,78E-14 | \*\*\* |
| Operating costs | -0,031719949 | 0,001865319 | -17,00510791 | 5,05E-42 | \*\*\* |
| Cost of repairing the herd | -0,027449525 | 0,001539702 | -17,82781576 | 1,18E-44 | \*\*\* |
| Amortisation | -0,03640799 | 0,002086407 | -17,45009234 | 1,90E-43 | \*\*\* |
| Interest on loans | -0,022167963 | 0,005943765 | -3,729616262 | 0,00024377 | \*\*\* |
| Sales price of milk in physical weight | 0,006406148 | 0,002634459 | 2,431674599 | 0,015825478 | \* |
| Sales price of milk in hundredweight | 0,021269243 | 0,003631208 | 5,857346411 | 1,69E-08 | \*\*\* |
| Subsidies | -0,001547696 | 0,001308987 | -1,182362073 | 0,238332603 |  |
| IOFC | 0,00024065 | 9,40E-05 | 2,560310164 | 0,011124703 | \* |

Table 1.6 - Descriptive statistics of time FE (fixed effect) model

As can be seen from Table 1.6, the values of some variables have changed, but we are interested in those values that were insignificant in other models without the use of robust errors, but have become significant. In this case, the realised price of milk in standard weight became a significant coefficient, and more significant than the realised price of milk in physical weight. Looking at the data on this variable, it is clear that indeed, the realised milk price at certain time intervals was the same among the different units. Simply put, the previous models could not take this into account in any way, so the price of milk in standard weight remained an insignificant coefficient in all previous models, although in fact it has a direct impact on profitability. Also in this model, the coefficients of determination increased slightly (**R-Squared: 0.98238 Adj. R-Squared: 0.97712**) and the **F-statistic** increased (**F-statistic: 201.995 on 61 and 221 DF, p-value: < 2.22e-16**). Although the F-test and the Lagrange test showed a clear priority in favour of the time FE model (**F = 3.7182, df1 = 47, df2 = 221, p-value = 2.499e-11**, **chisq = 26.5, df = 1, p-value = 2.635e-07**), we cannot give it priority, as this way we cannot correct the problem with heteroscedasticity and autocorrelation. The final model should be the FE model with robust Driscoll-Kraay errors.

In the final fixed-effects model with the application of robust Driscoll-Kraay errors, the marketability indicator ceased to be significant in terms of its impact on profitability in the units of an agro-industrial organization on a par with subsidies, cow culling and the selling price of milk in standard weight. But as the time FE model shows the realized price of milk in standard weight is actually a significant coefficient and its impact on profitability is high. Non-significant indicators have weak variability over time and across facilities, which is due to good standardization of production (exception: subsidies). The effect of these indicators on profitability would be better tested among units with wider coverage, i.e. the sample should contain more facilities. Although such a narrow analysis with respect to the number of units allows to identify the effects of KPIs on profitability more precisely in a certain area of the territory.

# Discussion

The study revealed that while analysing the factors affecting profitability, the main focus should be on:

* gross production in standard weight has a negative impact on profitability
* the level of farm utilisation (the closer the utilisation is to the maximum, the lower the economic (opportunity) costs)
* the amount of milk yield in standard weight per one forage cow (there is a negative dependence of profitability on milk yield)
* feed, operating costs, herd repair costs, depreciation, amortisation, interest on loans also have a negative correlation, but it is quite expected (depreciation has the greatest impact, followed by operating costs and interest on loans).
* sales price of milk in physical weight and IOFC positively affect the profitability of an agro-industrial organisation
* the realized price of milk in standard weight also has a positive effect on profitability

It is worth discussing separately the dependence of profitability on milk yield in standard weight and gross production in standard weight. The negative dependence seems strange if we take into account the individual dependence (Figures 2.2 and 2.3).

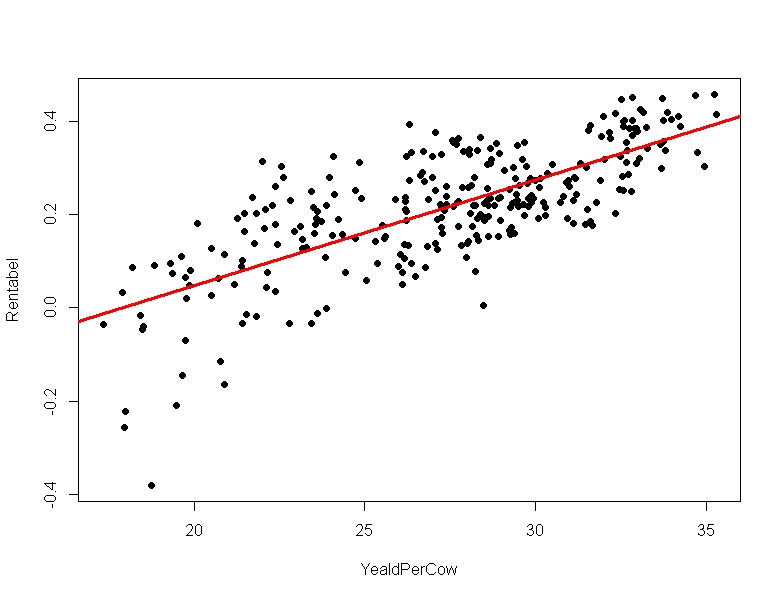


Figure 2.2 - dependence of profitability on milk yield in standard weight per 1 fodder cow

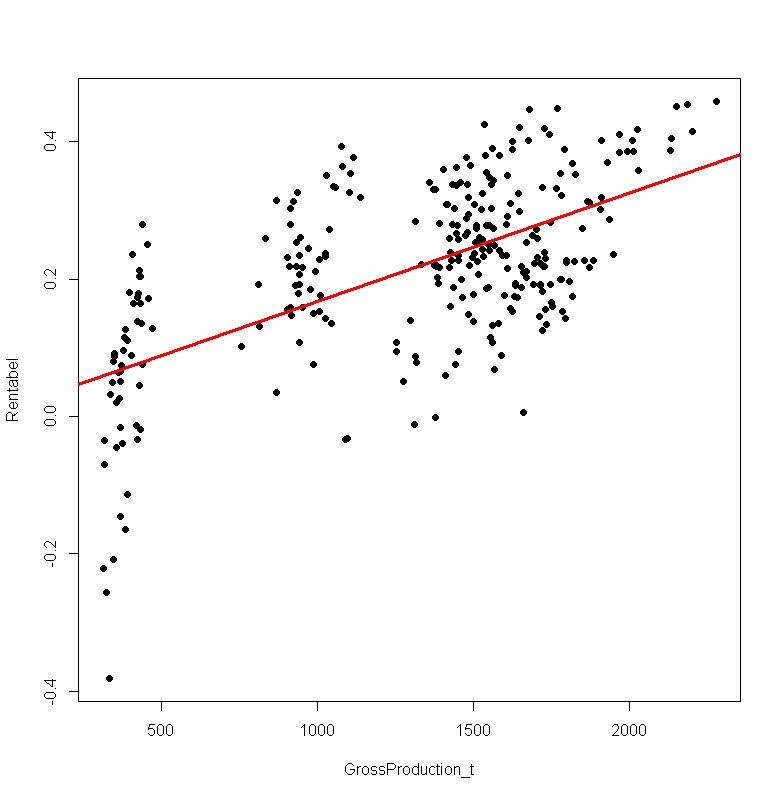


Figure 2.3 - dependence of profitability on gross production in the standard

Nevertheless, each of the models shows that when gross production in standard weight and milk yield in standard weight per cow increases, profitability decreases, albeit slightly. It can be assumed that this is due to increased load on farms. At increased load, the growth rate of costs is higher than the growth rate of profit. In fact, there is a shift of the optimum. This phenomenon is peculiar to firms with ever increasing production. In order to meet demand, the firm has to overload its production capacity, which ultimately increases costs.