ACTA SCIENTIARUM POLONORUM *Oeconomia 13 (2) 2014, 71–81*

DOES NATURAL HEDGE ACTUALLY WORK FOR FARMERS?

Paweł Kobus
Warsaw University of Life Sciences – SGGW

Abstract. The paper presents an analysis of the relation between yields and prices for major crop plants in Poland. The main objective of the paper was to examine the presence of natural hedging. For that purpose, yield value variances were calculated and compared with the theoretical variations of independent price and yield product. It was revealed that for sugar beet and rape a natural hedge could be observed, which leads to a 53% reduction in yield value variance. In case of wheat and barley, no natural hedge effect was observed. Practical implications of the conducted analysis are such that the tendency to consider production and price risk separately, could be very misleading in assessing the income risk of specific crop plants. It was also found that the negative correlation coefficient, commonly considered as being equivalent of a natural hedge effect, can be used only as a very rough measure of natural hedge strength.

Key words: natural hedge, income risk, agriculture, aggregation bias, price – yield correlation

INTRODUCTION

Price risk and yield risk are the most significant risks in the agricultural sector. Although it is generally believed that there exists a negative relation between the yield and price of crop plants. If this were true, it would work like a natural hedging mechanism. In years with lower yields, prices would increase and in years with higher yields, prices would decrease. Hence, the variability of their product, i.e., crop values would be lower than in case of unrelated, independent movements of yield and price[1]. Lower variance

[1] The probabilistic mechanism behind this so-called natural hedging [Finger 2012] is presented in the appendix.

Corresponding author – Adres do korespondencji: Paweł Kobus, Department of Agricultural Economics and International Economic Relations, Faculty of Economic Sciences, Warsaw University of Life Sciences – SGGW, Nowoursynowska 166, building 5, room 13B, 02-787 Warszawa, Poland
e-mail: pawel_kobus@sggw.pl

of crop value means lower income risk, and by the same token, lower demand for crop insurance. The proper assessment of crop plant specific income risk can help explain very low demand for crop insurance in Poland.

In many cases analysis of yield – price dependency is carried out with data aggregated on the national level. Using this kind of data may lead to serious overestimation of the correlation strength. In research by Coble et al. [2007] it was shown that the yield – price correlation for corn and soybeans in the USA are –0.381 and –0.386 when calculated on national level, and they drop to –0.064 and –0.096 when calculated on the farm level. According to Harwood et al. [1999], the strength of the negative correlation between corn yield and price estimated at farm level varies considerably across the counties in the United States. The strongest correlation is observed in counties that form part of the Corn Belt, while it is much weaker in the remaining counties. This could be explained by the fact that the Corn Belt supplies over half of total corn production in the USA and consequently, low yields in that part of the USA would seriously downsize the total supply of corn and cause an price increase. But even in counties from the Corn Belt, the average correlation at farm level is lower than the correlation at national the level.

The author's earlier work [Kobus 2011], which dealt with the influence of data aggregation levels on the variability of yields and prices of major crop plants also in Poland, showed that aggregation has a reducing effect on variability, although the degree of variability reduction was crop plant specific and in most cases bigger for yields than for prices. This suggests that the way in which aggregation affects the correlation should also be crop plant specific. Furthermore, the results from Harwood et al. [1999] imply that the strength of this correlation, even at the highest level of data aggregation, should vary in Poland. It is expected that the correlation should be low for those crop plants which for various reasons are easy to import to Poland and high in opposite cases.

The main objective of this paper is to examine the presence of the natural hedging mechanism for major crop plants in Poland. The secondary objective is to find whether it is justified to use the correlation between crop plant yields and their prices as an indicator of crop value variability reduction.

MATERIAL AND METHODS

Data

Two data sources were used in this analysis: the Central Statistical Office of Poland (CSO) and the Polish Farm Accountancy Data Network (FADN). The CSO data presented in Table 1 was used for assessing the strength of the natural hedging mechanism on the national level, however, the nominal prices shown in Table 1 were deflated using consumer goods and services price indices (CPI). The yields were not adjusted in any way, even though detrending had been recommended in the author's other paper [Kobus 2010]. The reason for this was the small length of time series, which involved a risk of serious overfitting.

Table 1. Yields and prices for major crop plants in Poland in period 2005–2011

Year	Winter wheat		Barley		Rape		Sugar beet	
	yield (dt)	price (zł)	yield (dt)	price (zł)	yield (dt)	price (zł)	yield (dt)	price (zł)
2005	39.5	36.69	32.2	37.34	26.3	77.33	416	17.57
2006	32.4	44.76	25.9	40.24	26.5	93.44	438	12.88
2007	39.4	70.68	32.5	64.11	26.7	95.66	513	10.83
2008	40.7	64.24	30.0	64.37	27.3	126.77	465	10.37
2009	41.7	48.26	34.4	40.80	30.8	108.24	543	11.57
2010	43.9	59.84	34.9	48.98	23.6	127.76	483	11.31
2011	41.3	81.99	32.7	75.38	22.4	183.91	574	14.40

Source: CSO, Local Data Bank.

As mentioned in the introduction, the author expects the level of data aggregation to have a significant influence on the strength of natural hedging. To examine this aspect of the problem, farm level data from the FADN sample were used. The process of data selection was as follows: samples from 2005–2011 were screened for farms which were present in the samples in all the years, and next, a separate selection was carried out for each crop from that pool. The selection criterion was availability of yield and transaction data for at least 6 years. Due to this, the sample sizes presented in Table 2 differ.

Table 2. The sizes of samples for each plant researched

Crop plant	Sample size
Winter wheat	1 647
Barley	512
Rape (with turnip rape)	858
Sugar beet	782

Source: Own calculations.

Obtaining crop plant yields is fairly straightforward, but prices need some clarification. Ideally, these should be transaction prices in-between harvests, but such data was not available. Hence, the average transaction prices for calendar years were used. In case of such crop plants like sugar beet or oil rape, when all production is sold directly after harvest, use of calendar years is totally justified. For cereals though, it could obscure the relation between yields and prices, in proportion to the amount of production sold next year following harvest.

The transaction prices on the farm level had been deflated with CPI, just like in the CSO national data case.

METHODS APPLIED

Let us denote crop yield by X and price by Y. The problem is to assess the relation of their product, i.e., the XY variance to the dependence between variables X and Y. An explanation of the random variables sum variance can be easily found in any introduction to probability textbook. However, the product variance is not only more complicated, but also less often found in books on probability. For this reason, the appendix contains the formal derivation for the variance of the random variables product. In this text, only the final formulas will be presented. The independence variance of product can be calculated using the formula (1):

$$D^2(XY) = D^2(X)D^2(Y) + D^2(X)E(Y)^2 + E(X)^2 D^2(Y) \tag{1}$$

where: $D^2(X)$, $D^2(Y)$, $E(X)$, $E(Y)$ – variances and expected values of variables X and Y, respectively.

When variables X and Y are dependent, the product variance should be calculated using the formula (2):

$$\begin{aligned} D^2(XY) &= D^2(X)D^2(Y) + D^2(X)E(Y)^2 + E(X)^2 D^2(Y) + \\ &+ C(X^2, Y^2) - C(X,Y)^2 - 2C(X,Y)E(X)E(Y) \end{aligned} \tag{2}$$

where: $C(X, Y)$ and $C(X^2, Y^2)$ are covariances of variables X and Y and their squares.

In case of a bivariate normal distribution, $C(X^2, Y^2) = 2C(X, Y)^2 + 4C(X, Y)\,E(X)\,E(Y)$, so formula (2) becomes reduced to:

$$\begin{aligned} D^2(XY) &= D^2(X)D^2(Y) + D^2(X)E(Y)^2 + E(X)^2 D^2(Y) + \\ &+ C(X,Y)^2 + 2C(X,Y)E(X)E(Y) \end{aligned} \tag{3}$$

This last formula could be treated as a quadratic function of the correlation – see Figure 1.

The position of the parabola vertex on Figure 1 depends on values of variation coefficients V_x and V_y, and if their product is outside of the <–1, 1> range, the parabola vertex is located within the range of values possible for the correlation coefficient (see solid line). Otherwise, the parabola vertex cannot be observed in that range (see dashed line). In fact, the product of variation coefficients is usually below 0.05 for yields and prices of crops, and consequently, the relation between the correlation coefficient and price-yield product is straightforward – the lower the correlation, the lower the variance.

The following procedure has been adopted for evaluation of the natural hedging mechanism strength on the national level:

1. Calculation of the observed variance of yield – price product;
2. Calculation of the hypothetical variance of yield – price product assuming independency, according to formula (1);

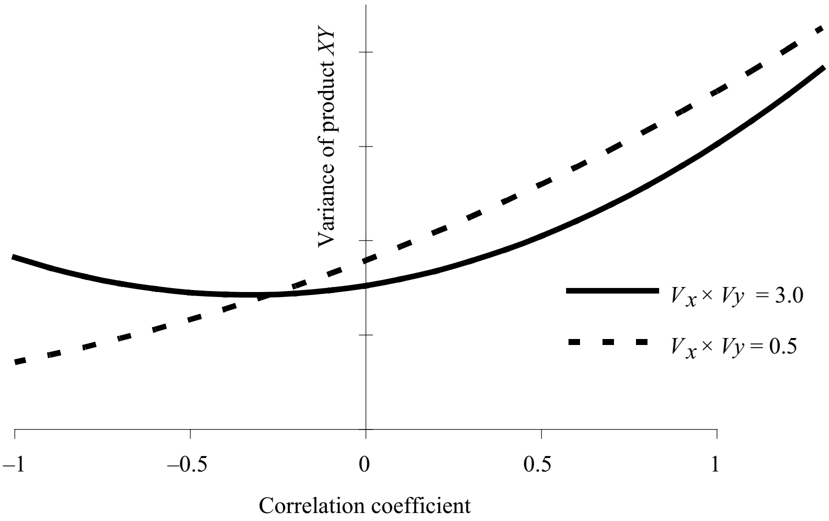

Fig. 1. Example random variables product variance as a function of correlation coefficient
Source: Own calculations

3. Calculation of the reduction of yield – price product variance (point 1) in comparison to the case of independency (point 2).

In order also to address the suitability of using the correlation coefficient as a proxy for assessing natural hedge strength, a similar procedure was used, whereby the observed variance of yield – price product was replaced by the hypothetical variance of yield – price product, assuming a bivariate normal distribution as in formula (3).

As mentioned in the introduction, it is self-evident that the yield – price dependency is stronger at higher aggregation levels. A single farmer is a price taker, only the aggregate supply of crop can affect price. Still, there are some open questions: how strong is the effect of this aggregation and which crops are affected more strongly. In this paper, six data aggregation levels are examined, i.e.: farm, gmina[2], powiat, voivodship, region and country.

The described above procedures for assessing strength of natural hedge were performed on each level of aggregation, the only difference being that instead of directly calculating the reduction of variability for each organization unit in point 3, the variances were first averaged across all units using the production area as weights, and then used for calculating the mean variability reduction.

[2] According to nomenclature of territorial units for statistics (http://epp.eurostat.ec.europa.eu/portal/page/portal/nuts_nomenclature/correspondence_tables/national_structures_eu) gmina is the lower level of Local Administrative Units (LAU2), while powiat is the upper level (LAU1) in Poland.

RESULTS

As mentioned earlier the characteristics of crop yields and prices probability distribution on farm level differ from those obtained at the country level of data aggregation. This is especially true for crop yields, due to the lower spatial autocorrelation in comparison to prices. The direction of divergence is quite clear, average variances are higher and correlations are lower on the farm level. Nevertheless, even the data aggregated at the country level could show which crop plant yield value variability is affected more strongly by the yield – price dependence.

Table 3. Strength of natural hedge effect for major crop plants in Poland on national level in period 2005–2011

Crop plant	Correlation of yield and price (–)	Observed variation of yield value (–)	Theoretical variation of yield value in case of independency	Observed reduction of variation (%)	Theoretical reduction of variation in case of bivariate normal distribution of yields and prices (%)
Winter wheat	0.26	282 118	250 453	–12.6	–17.1
Barley	0.00	144 770	148 996	2.8	0.0
Rape	–0.59	206 955	440 545	53.0	42.9
Sugar beet	–0.52	939 918	2 002 721	53.1	40.6

Source: Own calculations based on Central Statistical Office of Poland.

The correlation coefficients in Table 3 are quite diverse. For sugar beet and rape, the correlations are negative and fairly strong, but for barley, it is almost exactly zero and for wheat, the correlation is even positive, although rather weak. These differences in correlation values can be attributed to several reasons, but three of them seem to be the most important ones: share of production in the global market, transport costs and possibility for safe and cheap storage. With respect to the production of the four crops being examined, in 2011 Poland had 16th place for wheat, 12th for barley, 8th for rape and 7th for sugar beet global production shares: 1.33, 2.5, 2.98 and 4.27% respectively. What is more, cereals can be easily stored even by farmers on their own, while rapeseeds and sugar beet roots must be sold by farmers immediately after harvesting. Consequently, we can observe quite a high negative correlation only in case of rape and sugar beet, whereas there is no correlation in case of cereals or it can in fact even be positive. It must also be noted that sugar beet is a special case because of the contracts signed by farmers with the sugar factories and any production above quota levels collects much lower prices.

Comparing the observed variation of yield value with the theoretical one, that has been calculated assuming independence, we are able to see that very roughly speaking, the reduction expressed in percentage value is equal to a correlation coefficient value multiplied by –100%. For example for rape, the correlation is –0.59 and for sugar beet it is –0.52, while at the same time, the variability reduction is nearly 53% for both of them. Given the positive correlation for wheat, it is not surprising that in case of this crop we do not observe a natural hedge effect and the variation of the observed yield value is higher than it would be in case of yield and price independence.

Although the correlation coefficient seems to approximate quite well the strength of the natural hedge effect, it by no means proves bivariate normality. In fact, the natural hedge effect is actually stronger than it should be if yield and price were to follow a bivariate normal distribution. As mentioned earlier in the paper, the variance of random variables product is just a function of correlation coefficient only in case of a bivariate normal distribution and even then this function's coefficients depend on variances and the expected values of price and yield. Therefore, one must remember that the correlation coefficient value is a very rough indicator of natural hedge strength.

So far, all considerations have been based on data aggregated on the national level by CSO. Since according to the author's previous work [Kobus 2011], the aggregation levels strongly influence the variances of yields and crop prices, therefore it should be expected that also the correlation and strength of a natural hedge effect shall be affected by the level of aggregation.

Table 4. Strength of natural hedge effect on various levels of data aggregation in period 2005–2011

Aggregation level	Mean unit area (ha)	Mean R (–)	Reduction of variability (%)	Mean unit area (ha)	Mean R (–)	Reduction of variability (%)
	Wheat			Barley		
Farm	15	0.048	–4.0	11	–0.102	13.1
District	31	0.056	–3.8	14	–0.118	14.3
Powiat	91	0.065	–2.3	25	–0.133	16.0
Voivodship	1 478	0.086	–3.9	309	–0.113	17.1
Region	5 912	0.079	–5.9	1 235	–0.140	15.7
Country	23 648	0.105	–5.0	4 938	–0.283	18.2
	Rape (with turnip rape)			Sugar beet		
Farm	15	–0.339	36.2	5	–0.374	47.8
District	28	–0.416	42.1	11	–0.463	51.0
Powiat	68	–0.492	51.3	31	–0.559	55.0
Voivodship	764	–0.654	65.0	263	–0.656	58.8
Region	3 055	–0.768	72.7	986	–0.700	60.4
Country	12 221	–0.823	74.2	3 943	–0.706	60.6

Source: Own calculations, based on FADN data.

There are notable differences on the national level of data aggregation between results from CSO and FADN, with generally the strength of the natural hedge effect being greater for the FADN data, in case of rape the increase is above 20%. The reason for those differences is such that while CSO data represents all farms in Poland, FADN data is only a sample, and what is more, it is representative only for farms of the economic size of at least 2 ESU.

Similarly to the calculation based on CSO data also the results from Table 4 show that the biggest reduction in variation is observed in case of rape and sugar beet. What could be surprising at first, is that the relation between the correlation coefficient values and reduction of variability is different for each crop plant, for example, on the level of voivodeship, the correlation coefficient takes the value of –0.65 for both rape and sugar beet, while the reduction of variability differs by 6 percent points. However, according to formula (2), the part of random variables product variance, which changes because of the dependence between variables, i.e.:

$$C\left(X^2, Y^2\right) - C\left(X, Y\right)^2 - 2C\left(X, Y\right)E\left(X\right)E\left(Y\right) \tag{4}$$

consists of three elements and can take on a totally different value for the same correlation coefficient.

Generally, the strength of the natural hedge effect diminishes at the farm level of data aggregation, but for obvious reasons, the degree of this decline is worth mentioning only when it has a fairly strong effect on the national level. In case of rape, the effect on the farm level is about 50% of that on the national level, while in case of sugar beet, it is about 75%.

It was shown earlier that natural hedge strength varies greatly between crop plants and that it is not a simple function of the correlation between yields and prices. But what are the practical implications of these findings?

The risk of any economic activity is usually measured by standard deviation or variation coefficient of income, and the lower variation coefficient, the less risky the activity. In case of crop plants, there is a tendency to consider production and price risk separately and if yields and prices were actually independent of one another, this would not be a problem. However, because of their dependency, this could be misleading.

Table 5. Variation coefficients for price, yield and yield value in period 2005–2011

Crop plant	Price	Yield	Yield value
		%	
Winter wheat	8.4	22.0	24.9
Barley	8.9	23.2	24.5
Rape (with turnip rape)	9.5	21.8	16.6
Sugar beet	10.7	21.8	16.9

Source: Own calculations based on Central Statistical Office of Poland.

Yield variation coefficients are almost the same for all crops (Table 5). However, in case of prices, variation coefficients for rape and sugar beet are considerably higher, so therefore it would be possible to conclude that rape and sugar beet are the most risky crop plants. But if we look at the variation coefficients for yield value (Table 5), we will see that the situation is in fact completely the opposite, with the risk for rape and sugar beet being two-thirds of that for cereals.

CONCLUSIONS

1. The existence and strength of the natural hedge effect depends on the crop plant species. Generally, the effect is stronger in case of those crop plants, the use or consumption of which is more limited by area and time.
2. The correlation coefficient can be used only as a very rough measure of natural hedge strength and cannot be treated as being equivalent of a natural hedge effect.
3. The lower the aggregation level, the weaker the natural hedge effect. On average, the strength on farm level is two-thirds of the strength on the national level of data aggregation.
4. Omission of the natural hedge effect leads to invalid estimations of income risk for specific crop plants.

REFERENCES

Bohrnstedt G.W., Goldberger A.S., 1969. On the Exact Covariance of Products of Random Variables. Journal of the American Statistical Association 64, 328, 1439–1442.

Coble K.H., Dismukes R., Thomas S., 2007. Policy Implications of Crop Yield and Revenue Variability at Differing Levels on Disaggregation. Selected Paper. American Agricultural Economics Annual Meeting, Portland, OR.

Finger R., 2012. How strong is the “natural hedge”? The effects of crop acreage and aggregation levels. 123rd EAAE Seminar, February 23-24, Dublin.

Frishman F., 1975. On the Arithmetic Means and Variances of Products and Ratios of Random Variables. A Modern Course on Statistical Distributions in Scientific Work NATO Advanced Study Institutes Series 17, 401–406.

Harwood J., Heifner V., Coble K., Perry J., Somwaru A., 1999. Managing risk in farming: concepts, research, and analysis. Economic Research Service, USDA [electronic resource] http://www.ers.usda.gov/publications/aer774/aer774.pdf.

Kobus P., 2010. Modelling Wheat Yields Variability in Polish Voivodeships. Problems of World Agriculture 10 (25), 33–40.

Kobus P., 2011. Influence of aggregation level on yield risk measures. EAAE 2011 Congress: Change and Uncertainty Challenges for Agriculture, Food and Natural Resources, August 30 to September 2, Zurich, 1–8.

CZY NATURALNE ZABEZPIECZENIE RZECZYWIŚCIE CHRONI ROLNIKÓW?

Streszczenie. W pracy przedstawiono analizę relacji między plonami i cenami podstawowych roślin uprawnych w Polsce. Głównym celem pracy było sprawdzanie występowania efektu naturalnego zabezpieczenia. Dla wybranych roślin uprawnych obliczono wariancje wartości plonów jednostkowych i porównano z teoretyczną wariancją iloczynu ceny i plonu, przy założeniu niezależności. Wykazano, że wśród rozpatrywanych roślin tylko w przypadku rzepaku i buraków cukrowych można mówić o występowaniu efektu naturalnego zabezpieczenia. W wyniku negatywnej relacji między cenami i plonami dla obydwu roślin zaobserwowano redukcję zmienności wartości plonu o 53%. Praktyczną konsekwencją przeprowadzonej analizy jest stwierdzenie, że tendencja rozłącznej analizy cenowego i produkcyjnego ryzyka może prowadzić do całkowicie błędnej oceny ryzyka dochodowego, które z punktu widzenia rolnika jest najważniejsze. Wykazano również, że ujemna korelacja, popularnie utożsamiana z efektem naturalnego zabezpieczenia, może być tylko nieprecyzyjnym przybliżeniem siły efektu naturalnego zabezpieczenia.

Słowa kluczowe: naturalne zabezpieczenie, ryzyko dochodowe, rolnictwo, błąd agregacji, korelacja plonu i ceny

Appendix: Derivation of the expected value and variance of the random variables product

Let *X* and *Y* be a pair of jointly distributed random variables, with expected values *EX* and *EY*, variances D^2X and D^2Y, and covariance *C*(*X*, *Y*). The expected value of the *XY* product can be calculated relatively easily using the formula:

$$E(XY) = E(X)E(Y) + C(X, Y) \tag{A.1}$$

where covariance *C*(*X*, *Y*) is given by formula:

$$C(X, Y) = E(XY) - E(X)E(Y) \tag{A.2}$$

The formula for variance is more difficult to derive. When applying directly the definition variance, we get:

$$D^2(X, Y) = E\left[XY - E(XY)\right]^2 \tag{A.3}$$

Bohrnstedt and Goldberger [1969] proposed first to decompose variables into centred variables and expected values, such as ($\Delta X + EX$) instead of *X*, where $\Delta X = X - EX$. Frishman [1975] used the following property of variance as the starting point:

$$D^2(X) = E(X^2) - [EX]^2 \tag{A.4}$$

but applied it to product *XY*:

$$D^2(XY) = E(X^2Y^2) - \left[E(XY)\right]^2 \tag{A.5}$$

Using formulas (A.1), (A.2) and (A.5), Frishman showed that:

$$D^2(X, Y) = \left[E(X^2)E(Y^2) + C(X^2Y^2)\right] - \left[E(X)E(Y) + C(X, Y)\right]^2 \tag{A.6}$$

By expanding formula (A.6), we arrive at:

$$\begin{aligned}D^2(XY)=&\left\{\left[D^2(X)+E(X)^2\right]\left[D^2(Y)+E(Y)^2\right]+C\left(X^2,Y^2\right)\right\}-\\&-\left[E(X)^2E(Y)^2+2C(X,Y)E(X)E(Y)+C(X,Y)^2\right]=\\&=\left[D^2(X)D^2(Y)+D^2(X)E(Y)^2+E(X)^2D^2(Y)+E(X)^2E(Y)^2+C\left(X^2,Y^2\right)\right]-\\&-\left[E(X)^2E(Y)^2+2C(X,Y)E(X)E(Y)+C(X,Y)^2\right]\end{aligned} \tag{A.7}$$

Simplifying formula (A.7) results in:

$$\begin{aligned}D^2(XY)=&D^2(X)D^2(Y)+D^2(X)E(Y)^2+E(X)^2D^2(Y)+\\&+C\left(X^2,Y^2\right)-C(X,Y)^2-2C(X,Y)E(X)E(Y)\end{aligned} \tag{A.8}$$

In case of the bivariate normal distribution $C\left(X^2,Y^2\right)=2C(X,Y)^2+4C(X,Y)E(X)E(Y)$ and formula (A.8) becomes reduced to:

$$\begin{aligned}D^2(XY)=&D^2(X)D^2(Y)+D^2(X)E(Y)^2+E(X)^2D^2(Y)+\\&+C(X,Y)^2+2C(X,Y)E(X)E(Y)\end{aligned} \tag{A.9}$$

.

Both formulas for the expected value (A.1) and for variance (A.8) of the random variables product become considerably simplified in the case of independence:

$$E(XY)=E(X)E(Y) \tag{A.10}$$

$$D^2(XY)=D^2(X)D^2(Y)+D^2(X)E(Y)^2+E(X)^2D^2(Y) \tag{A.11}$$

Accepted for print – Zaakceptowano do druku: 24.04.2014