Every company should be directed to achieve income and reduce costs. Agricultural farms are no exception. In agriculture, achieved revenues are lower than in other sectors of the economy. Analysis of farm income should therefore be a key research issue (Zawadzka, Strzelecka 2012). Since the increase in the income of persons employed in agriculture allows one to provide an adequate standard of living (Babuchowska, Marks-Bielska 2011). Poland entered into the structure of EU agriculture in 2004, when Polish agriculture was incorporated in the common agricultural policy (CAP) and has become a recipient of Pillar I and II funds. An agricultural farm, however, is a different kind of a company. Innovations in farming are supported by new systems and management methods (for example: the introduction of new fertilisers and plant protection products, but also the purchase of new machinery). Farm investments should be conducted in the field of economics and organisation. This may involve increasing the area of the farm, using credits and introducing computer technology (van den Ban, Hawkins 1996). Making changes on farms is difficult because it can be associated with a change in the thinking of farm manager. The problem in action can also have no effect on the prices of agricultural products, which are determined by consumers and not producers. Changes in farm management should be an adjustment of character or planned. In contrast, the changes depend on the speed of response to emerging agricultural innovations and its relation to risk (Rizov et al. 2013). Figure 1 shows the percentages of groups of farmers according to response speed

* Warsaw University of Life Sciences – SGGW, Faculty of Applied Informatics and Mathematics.
** Warsaw University of Life Sciences – SGGW, Faculty of Economic Science.
to emerging innovations by Rogers (2003). Rogers divides farmers into several groups. The “Innovators” have large, specialised farms and are not afraid of risk. They are well educated and have a scientific approach to problem-solving. The second group of farmers is known as “Early Adopters”. This group includes well-educated farmers who are ready to implement new ideas proven by the innovators. The next two groups of farmers are described by Rogers as the “Early” and “Late Majority”. The first group introduces new solutions to the farm rather willingly. The second group makes changes only under economic or social pressure. The last group (“Laggards”– traditionalists) are not at all interested in making changes on their farms. The most common reason for this is that this group includes older farmers, producing only for their own needs.

In view of the above considerations, due to the CAP we can speak of the existence of new conditions for the development of agriculture in Poland. New financial instruments have permitted the development of farms. Due to the dynamic growth of subsidies, it should also be noted their role in increasing revenue in Poland’s agricultural sector. In the pre-accession period (1999–2003), subsidies accounted for an average of less than 9% of the income of farmers, while in the post-accession period this share exceeded 50%. The decisive role in this respect plays into how the subsidies are used at the farm level. These payments are not subject to clearing and verification due to their destination at the farm level. The freedom to dispose of these funds means that farmers only partially use them for the development of the farm or spend them on innovation. On small farms, subsidies are a component of income and are intended for living necessities.

Similar changes are related to province data. The largest share of subsidies during the period in question were received by the Mazowieckie and Małopolskie provinces (an average of 12–13% of all subsidies), while the lowest transfer of EU funds was received by the Śląskie, Podlaskie, Świętokrzyskie and Lubuskie provinces (average 2–3% of the total). In 2005–2008, changes occurred in the types of subsidies as dictated by the changing EU policy on agriculture, namely the increasing role of the environment and rural development. Farms increasingly depended on the receipt of payments from meeting environmental standards. Subsidies such as LFA, supporting agro-environmental and animal welfare, are to support the restructuring process of Polish agriculture and are not intended to stimulate the growth of agricultural production (Buks, Pietrzykowski 2014).
The primary aim of this study is to determine the changes in agricultural regions based on income in the period from 2004 to 2012. In the first part of the paper, we summarise the income dynamics in the different provinces in Poland. The dynamics are elaborated in terms of income level, pace and acceleration. In this part, we apply functional principal component analysis. As a result, we characterise the main sources of the income dynamics, which
allows us to find the provinces whose income dynamics can be regarded
as unusual. We support the results with a rate-depth analysis. In the second part,
we take into account some predictor variables. We investigate their time-varying
impacts on income using a functional concurrent regression model. Thus,
we reveal if the impact is decreasing or increasing.

2. DATA

We used data from the Polish FADN for the period 2004 to 2012. We investigated farm net income as expressed per family labour unit. This variable is characterised by high variability across the different holdings, which is caused by some farms with extremely high income or debt (Goraj et al. 2010). Thus, we considered only 50% of the typical holdings in Poland selected in the range of the first and the second quartile of corresponding incomes (4897 observations in each investigated year). Taking into account the stratified sampling of the FADN project, we calculated the mean income level for each province. Thus, we dealt with trimmed means of the income curves. The explanatory variables taken into account are described as follows (compatible with FADN markings):

- **SE010**: Total labour input expressed in AWU (annual work unit = full-time person equivalent).
- **SE281**: Total specific costs. This includes crop-specific inputs (seeds and seedlings, fertilisers, crop protection products and other specific crop costs), livestock-specific inputs (feed for grazing stock, other specific livestock costs) and specific forestry costs.
- **SE025**: Total utilised agricultural area of the holding. Does not include areas used for mushrooms, land rented for less than one year on an occasional basis, woodland and other farm areas (roads, ponds, non-farmed areas, etc.). This consists of land occupied by the owner, rented land and land in sharecropping. This includes agricultural areas and areas temporarily not under cultivation for agricultural reasons or ones being withdrawn from production as part of agricultural policy measures. It is expressed in hectares.
3. STATISTICAL ANALYSIS AND RESULT

The first step of the analysis is to represent income as a smooth function of time. We want the function have a third derivative. The idea behind this is that income rate is driven by overall economic performance and should be filtered out of the noise coming from several shortcomings involved in data preparation. Moreover, we want to enquire about the dynamics of income rates in a unified way using the first and second derivatives. The first derivative corresponds to income pace and the second derivative to its acceleration. As a method of turning raw discrete data into smooth functions, we chose smoothing on a B-spline basis with a roughness penalty (the penalty is based on the integral of the square of the derivative of order 2).

Let \( y_{ij} \) denote income in time \( t_j \in T \) of the \( i \)-th district, where \( i = 1, \ldots, N \) and \( j = 1, \ldots, n \). Let \( c_i = [c_{i1}, \ldots, c_{ik}] \) be an unknown vector and \( \{\Phi_k\}_{k=1}^K \) be a set of known functions (B−spline basis of order 4). The smoothed income of the \( i \)-th district is expressed by \( x_i(t) = \sum_{k=1}^{K} c_i \Phi_k(t) \). Then, the estimate of the smoothed income is \( \hat{x}_i(t) = \sum_{k=1}^{K} \hat{c}_i \Phi_k(t) \), where \( \hat{c} = [\hat{c}_1, \ldots, \hat{c}_K] \) minimises:

\[
PENSSE(c) = \sum_i \sum_j \left( y_{ij} - x_i(t_j) \right)^2 + \lambda \sum_k \left[ \int D^2 x_i(s) \right]^2 ds .
\]  

The penalised parameter \( \lambda \) and the number \( K \) of basic functions are fixed. In fact, their values can be selected by generalized cross-validation criterion (see: Ramsay, Silverman 2005). The generalised cross-validation measure GCV for several \( (\lambda, K) \) combinations is depicted in Figure 3. The optimal \( (\lambda = \exp(-4), K = 8) \) point minimises the GCV measure, which is marked with a bullet in Figure 3. The original income series are exhibited in the left panel of Figure 4.

The smoothed income curves are in the right panel of Figure 4. We used them in the subsequent functional component analysis, which is designed to see what primary modes of variation are in the data and how many of them seem to be substantial.
3.1. Functional principal component analysis

The functional principal component analysis for functions $y_i(t), i=1, \ldots, N$ (see: Besse, Ramsay 1986) is designed to find weight functions $\xi_1, \ldots, \xi_M$, where each function $\xi_m, m \in \{1, \ldots, M\}$ is minimised:

$$\frac{1}{N} \sum_i \left( \int_{\tau} (\xi_m(s)x_i(s)ds)^2 \right)$$

and subject to:

$$\int_{\tau} \xi_m^2(s)ds = 1 \text{ and } \int_{\tau} \xi_m(s)\xi_k(s)ds = 0, \ (\forall \ k < m).$$
The vector \( f_m = (f_{lm}, \ldots, f_{Nm}) \) where \( f_{im} = \int_{T} \xi_m(s)x_i(s)\,ds \) is called the \( m \)-th principal component. The percentage of variability of the first \( m \) components is expressed as:

\[
V_m = \frac{\sum_{i=1}^{N} \sum_{l=1}^{m} f_{im}^2}{\sum_{i=1}^{N} \sum_{l=1}^{M} f_{il}^2}.
\]  

(4)

A method found to be helpful in interpreting the components is examining the plots of the overall mean function \( \hat{\mu}(t) = \frac{1}{N} \sum_{i} \hat{x}_i(t) \) and the functions obtained by adding and subtracting the suitable component functions \( \hat{\mu} \pm C_{k} \xi_{k}, k \leq m \), where \( C_{k} = \frac{1}{N} \sum_{i} f_{ik}^2 \). The plots of the components along with the principal components scores plots:

\[
\left\{ (f_{ik_1}, f_{ik_2}) : i = 1, \ldots, N \right\},
\]  

(5)

\( k_1 \leq k_2 \leq m \), giving good insight into the differences between objects \( y_i, i = 1, \ldots, N \).

We applied the analysis to \( y(t) = \dot{x}(t) \) (income), \( y(t) = \frac{d}{dt} \dot{x}(t) \) (pace of income) and \( y(t) = \frac{d^2}{dt^2} \dot{x}(t) \) (income acceleration).

### 3.1.1. Results of functional principal component analysis for income

The first two components explain 92\% of variability (\( V_2=92\% \)). The main variability is explained by the first component (\( V_1=79\% \)). Thus, the second component explains 13\% of total variability. Note that \( \hat{\mu}(t) - C_1 \xi_1(t) < \hat{\mu}(t) + C_1 \xi_1(t), \ (\forall t) \). Therefore, the first component shows that the main dissimilarity between the incomes of the provinces holds over time. Provinces with high/low income are placed to the right/left side of Figure 6. For example, the Kujawsko-Pomorskie, Pomorskie, Zachodniopomorskie and Opolskie provinces had, relatively, the highest level of income throughout the investigated period. In contrast, the Małopolskie province had the lowest income. Note also that the curve \( \hat{\mu}(t) - C_2 \xi_2(t) \) is crossing the curve
\( \hat{\mu}(t) + C_2 \xi_2(t) \) at the time \( t \), denoting the beginning of the year 2009. Thus, the second component shows that some provinces increased their incomes over time relative to other provinces. Examples of this are the Podkarpackie, Dolnośląskie and Opolskie provinces. The province with a relative decrease in income was Świętokrzyskie.

The diameter of each point in Figure 6 is proportional to the number of neighbours of the corresponding province. The provinces with relatively high numbers of neighbours are placed in the upper part of the Figure 6.

This suggests that the second source of variability relates to the number of neighbours. However, the points with large diameters (except the Świętokrzyskie province) are close to the horizontal axis, and we cannot claim that the relation is significant.

We agglomerated provinces into 6 groups: \{6\}, \{3, 9\}, \{5, 7, 10, 12\}, \{13\}, \{2, 4, 11, 15\}, \{1, 8, 14, 16\} (for details see: Figure 6). We then applied a complete agglomeration method with the PCA semimetric \( d_{PCA} \) to produce a dissimilarity structure where \( m = 2 \) and \( y_i = \hat{\lambda}_i \):
Spatial comparison of the level and rate of change…

\[ d_{PCA}(x_i, y_j) = \sqrt{\sum_{k=1}^{m} \left( \int (y_i(s) - y_j(s) \xi_k(s) \, ds \right)^2} = \sqrt{\sum_{k=1}^{m} (f_{ik} - f_{jk})^2}. \] (6)

The semimetric takes into account the variability as explained by the \( m \) principal components. The selected groups are noted in Figure 6 with different marks and can be easily interpreted in reference to Figure 6 and Figure 5. For example, the group \{2, 4, 11, 15\} and \{1, 8, 14, 16\} (for details see: Figure 6) correspond to provinces with high income levels.

The income curves for the Małopolskie, Opolskie and Świętokrzyskie provinces are suspected to be functional outliers. As a method of confirming or rejecting this hypothesis, we used the rate depth measures approach described in Febrero-Bande et al. (2008) and Cuevas et al. (2007). The curves with the lowest depth are likely to be functional outliers. For a depth, we used a random projection method where the random direction \( \alpha \) and projection of the data along the direction were taken. That is, we took the value:

\[ \int \alpha(s) y_i(s) \, ds. \]  

Then, the sample depth \( D(y_i) \) of data \( y_i \) is defined as the univariate depth of the corresponding one-dimensional projection expressed in terms of order statistics so that the median is the deepest point. A single representative value is obtained by averaging. The direction is chosen according to the Gaussian distribution. The bootstrap procedure is designed to select \( C \) so that, in the absence of outliers, the percentage of correct observations mislabelled as outliers is approximately equal to 1%:\[ P(D(y_i) > C) = 0.01. \] The value of \( C \) is found by estimating this percentile, using the observed sample curves. We used the smoothed bootstrap procedure based on the trimming described in Febrero-Bande et al. (2008). The results of the depth measure approach for income data are presented in Figure 7. The Małopolska province (label: 6–Małopolska) was chosen as an outlier. This is the province with the uniformly lowest income (note that it is placed to the left side of Figure 6).

![Figure 7. Income outliers – depth measure approach](image)

Source: own calculations.
3.1.2. Functional principal component analysis for pace of income

We represented income as smooth differentiable curves and then estimated income pace as the first derivative of \( \hat{x}_i, \ i = 1, ..., N \). We utilised functional principal analysis for the estimates. The first three components explained 82% of variability (39%, 26%, 17% for the first, second and third component respectively). These are presented in Figure 8. The constants \( C_1, C_2 \) and \( C_3 \) of the component functions are slightly enlarged in Figure 8 because it visually improves the disparity between the presented curves. The pace of income is mainly positive; it corresponds to the increasing income trend.

The first functional component reflects the recent changes (from 2010 to 2012) in the provinces’ income paces. The second component indicates that we can separate two groups of provinces. The first group has a lower income dynamic in contrast to the second group, which has a higher dynamic. This is indicated by the fact that the curve \( \hat{\mu}(t) + C_2 \xi_2(t) \) is smoother than the curve \( \hat{\mu}(t) - C_2 \xi_2(t) \) (see: Figure 8). Provinces with smaller changes are placed at the bottom of Figure 9 (left).

The third component can help to distinguish two kinds of provinces. The first kind is characterised by an almost uniform greater pace of income vis-à-vis the second group of provinces.

![Figure 8. Component functions](source: own calculations.)

The provinces placed at the top of the functional principal components graph in Figure 9 (right) have a greater pace. The Lubuskie (4), Śląskie (12), Świętokrzyskie (13) and Zachodniopomorskie (16) provinces are suspected to be
outliers. The depth measure approach confirmed this with respect to the Lubuskie (4), Śląskie (12) and Świętokrzyskie (13) provinces, but not to Zachodniopomorskie (16).

Note that the highest variability of income pace across the provinces occurs when the pace is going to change the type of monotonicity of the overall income level. Moreover, the local maxima of the function $\hat{\mu}(t) - C_2 \xi_2(t)$ are shifted to the left in comparison to the local maxima of the overall level of income $\hat{\mu}$.

This means that the province labelled 16 (see: Figure 9 – left) preceded the monotonicity change.

![Figure 9](image)

**Figure 9.** Functional principal component plot for income pace: the second and the first component (left); and plot for income: the third and the first component (right)

Source: own calculations.

**3.1.3. Results of the functional principal component analysis for income acceleration**

Income acceleration is estimated as the second derivative of $\hat{x}_i$, $i = 1, ..., N$. The first three components explain 96% of variability (67%, 20%, 9% for the first, second and third component respectively). The first and second sources of variability describe the difference of points 4 and 13 (Lubuskie, Świętokrzyskie) against the others. This is seen from the allocation of points in the functional principal component plot (Figure 10). The source of variability between the other provinces is explained by the third principal component.

It was seen that the majority of variability was caused by two provinces. The provinces are outliers in the depth analysis. But it is difficult to characterise their income acceleration, because the first two sources of variability relate to the boundary points of the observational interval. It should be noted that the second derivative is hard to estimate at the points. However, this result can be caused by the non-standard income dynamics of the provinces.
4. FUNCTIONAL RESPONSES WITH FUNCTIONAL PREDICTORS: THE CONCURRENT MODEL

We put a model for the form:

$$ y_i(t) = \alpha(t) + \sum_{j=1}^{3} x_{ij}(t) \beta_j(t) + \varepsilon_i(t), \quad (7) $$

where: $j = 1, \ldots, N, \; N = 16$, $y_i$ – income of the $j$-th province, $x_{ij}$ – labour input of the $j$-th province averaged per holding [SE010], $x_{ij}$ – specific costs of the $j$-th province averaged per holding [SE281], $x_{ij}$ – utilised agricultural area of the $j$-th province averaged per holding [SE281], $\alpha$ – functional parameter representing overall income, $\beta_j$ – functional parameter of the $i$-th predictor, $i = 1, 2, 3$, $\varepsilon_i$ – functional error.

The functional parameters of the model are estimated by minimising the penalised sum of squares:

$$ \sum_{i=1}^{N} \int_{\tau} \left( y_i(t) - \alpha(t) - \sum_{j=1}^{3} x_{ij}(t) \beta_j(t) \right)^2 dt + \lambda_0 \int_{\tau} (\alpha(\tau))^2 d\tau + \sum_{i=1}^{3} \lambda_i \left( \beta_i(\tau) \right)^2 d\tau, \quad (8) $$
under the restriction that the observed functional variables and functional parameters are expanded by a B-spline basis of order 4. We fixed $\lambda_0 = 0.1$ and $\lambda_i = 5, \ i = 1, 2, 3, 4$. The estimated parameters (and point-wise confidence intervals under assumption of a Gaussian error) are depicted in Figure 12. The significance of the regression was checked by permutation tests. We used $F$-ratio statistics for each point $t$, denoted in the paper as $F(t)$. The statistic $F_{\text{max}} = \max_{t \in [2004, 2012]} F(t)$ was used to check the significance in the given range of time. The statistics and their permutation critical values are depicted in Figure 11 – left.

Observe that in recent years, income has become negatively associated with labour input and positively with the utilised agricultural area. At the same time, costs reduced their impact on income. The plot of residuals (see: Figure 11 – right.) shows how well the estimated model fits the data.

![Figure 11](image1.png)

Figure 11. $F$ statistics and their critical permutation values (left) and residuals plot (right)

Source: own calculations.

![Figure 12](image2.png)

Figure 12. Functional parameters of regression

Source: own calculations.
Figure 13. The curve \( \{(\hat{\mu}(t), \hat{\mu}(t)) : t \in T = [2004, 2012]\} \) with points (Year \( R^2 \)).

Source: own calculations.

The standard multiple correlation coefficients \( R^2(t) \) for \( t = 2004, 2005, ..., 2012 \) are equal respectively to 0.66, 0.56, 0.28, 0.50, 0.70, 0.78, 0.71, 0.68 and 0.62. The lowest coefficient is \( R^2(2006)=0.28 \) and the highest is \( R^2(2009)=0.78 \). We compared the coefficients with the pace and acceleration of income in a phase-plane plot (see: Figure 13). The lowest \( R^2(2006) \) preceded a period 2007–2008 of decreasing income. Note that 2006 shows (in the given model) no statistical significance between income and predictors.

5. CONCLUSIONS

This analysis relates to 50% of typical farms in Poland in terms of the time-varying mean income of the provinces. The results confirmed the growing impact of subsidies on farm income. The functional analysis allowed us to determine how varied farm income was in the period 2004–2012. Detailed conclusions can be specified for two on-going parts of the work as follows:

**Part one.** The main dissimilarity between provinces holds over time. We can distinguish provinces with relatively higher and lower income levels. The division has not changed since 2004. Some provinces improved their income relative to other provinces. At the beginning of 2007, their incomes converged and then crossed two years later at the beginning of 2009. The highest variability of pace of income across provinces occurred when the overall income was expected to change the type of its monotonicity. The change of the overall income pace could have been predicted by observing the behaviour of the pace of income of the 16\(^{th}\) province (Zachodniopomorskie). Two sinusoidal cycles of about 4 years in length in the income dynamics were observed. However,
the series of data in question seems to be too short to draw the conclusion that we could expect such cycles in next periods.

Part two. We managed to fit the functional regression model with time-varying parameters so we could analyse the impact of predictors on income in time. We found out that in recent years income has become negatively associated with labour input and positively with the utilised agricultural area. Specific costs became a non-significant factor. The parameter relating to utilised agricultural area has increased over time. The payment supports farmers’ incomes under the Common Agricultural Policy, which relies on granting financial support to a farmer proportionally to crop area, regardless of agricultural production volume. Because single area payment is increasing, this can be the reason that the parameter is also increasing. This is confirmed by research carried out by many authors showing that a significant part of income in Polish agriculture is payments. The study failed to confirm the use of payments for investment, but this was due to the specific nature of the data. This may also be a direction for further analysis.

REFERENCES

van den Ban A. W., Hawkins H. S. (1996), Agricultural Extension, Blackwell and Science, United Kingdom.


ABSTRACT

The primary aim of this paper is to determine the income changes in agricultural regions in the period from 2004 to 2012. In the first part of the paper, we have summarised the income dynamics in various provinces in Poland. The dynamics is elaborated in terms of income’s level, pace and acceleration. In the second part, we take into account some predictor variables. We have investigated the time-varying impact on income using a functional concurrent regression model. We used data from the Polish FADN. Our conclusions relate to 50% of typical farms in Poland in terms of the time-varying mean income of provinces. The analysis was performed in the R-CRAN.

PRZESTRZENNE PORÓWNANIE POZIOMU I TEMPA ZMIAN DOCHODÓW GOSPODARSTW ROLNYCH W LATACH 2004–2012

ABSTRAKT