DETERMINANTS OF WORK INCOME OF FAMILY WORK UNITS OF SWISS DAIRY FARMS

Sub theme: Entrepreneurship

Anke Schorr and Markus Lips

Agroscope, Tänikon, CH-8356 Ettenhausen, Switzerland

Abstract:

The paper analyses the determinants of annual income of a family work unit (FWU) for Swiss dairy farms based on data from the Farm Accountancy Data Network (FADN). Two regression types, namely a random-effects model and a quantile regression focusing on deciles, are applied. Organic and full-time farming, large farm size, high milk yield per cow and a diversification in either arable crops or agriculture-related activities are identified as attributes of a financially successful farm. In addition, the quantile regression reveals that the impacts of milk yield and concentrate inputs on annual income per FWU vary clearly among deciles, indicating that a high milk yield contributes most to the financial success of the best performing farms. The comparison between the two models shows that quantile regression is an appropriate analysis tool in the presence of substantial heterogeneity as is the case for Swiss dairy farms.

Key words: dairy, quantile regression, random-effects model, Switzerland

1 Introduction

Milk production in Switzerland is dominated by family farms. In addition, family members are carrying out the major part of the work. Accordingly, the remuneration of family work, i.e. of those mainly in charge of the work, as measured by annual income per family working unit (FWU), is a suitable indicator of the farm's financial performance. The annual report of the Swiss Farm Accountancy Data Network (FADN) regularly reveals a substantial heterogeneity. In 2015, the mean annual work income per FWU for dairy farms was CHF 37'600.-[3] (Dux *et*

[3] Average exchange rates 2016: CHF 1.00 = Euro 0.92 = USD 1.02 (https://data.snb.ch, accessed January 9, 2017)

al., 2016). While the mean income of the lowest quarter was CHF 14’200.-, that of the highest quarter was CHF 70’000.- or five times as much (Dux *et al.*, 2016).

Regression technique is a useful approach to identify determinants of the income per FWU. Applying a linear mixed regression model for Swiss dairy farms, Roesch (2015) shows that an increase in farm size, measured in utilized agricultural area and livestock units (LU), increases income, whereas the region determined by altitude above sea level, the number of family members and the area of wheat and maize cultivation have a negative influence. Hoop *et al.* (2015) examine the determinants of the production costs of one kilogram of milk for combined Swiss dairy and arable crop farms. They show that the farm size measured in LU and the milk yield per cow as explanatory variables have a negative influence on costs. In both of the mentioned studies, concentrate input is not considered.

To determine parameters which distinguish successful farms from their less successful counterparts, we carry out an analysis similar to those of Roesch (2015) and Hoop *et al.* (2015), addressing two additional issues. First, we introduce concentrate input as an explanatory variable, reflecting the increase of concentrate input in Swiss milk production in the last decade. Secondly, we apply a quantile regression to address the above-mentioned income heterogeneity. Furthermore, we apply a random-effects model to assess whether a quantile regression provides additional insights.

The remainder of this paper is organized as follows. Section 2 describes the data underlying our analysis, and section 3 outlines the methodology. Section 4 presents the results, section 5 provides a discussion and section 6 concludes.

2 Data

To perform the analysis, we use data from the Swiss FADN focusing on pure dairy farms only (farm type 21) between the years 2010 and 2014, a period in which no significant changes in

Swiss agricultural policy occurred for dairy farms. This data set yields 5’459 observations split between 1’832 farms, with an average of 3 observations per farm.

While the annual income per FWU serves as dependent variable, we use similar explanatory variables as Roesch (2015) and Hoop *et al*. (2015). In detail, the structural situation of the farm is represented by the farm size measured in LU, the stocking density, the location of the farm (plain, hill and mountain regions) and whether or not the farm is located in steep terrain (triggering direct payments). Production technique is addressed by milk yield per LU, organic production, usage of free-stall housing and input of concentrate feed per dairy cow. To assess diversification, the presence of arable crops and agriculture-related activities such as direct sales are introduced as dummy variables. A further dummy variable reports whether the farm is run full time or part time. With respect to education, a dummy variable is introduced which indicates whether the farm manager’s partner has any additional education in a non-agricultural subject[4]. Finally, the situation of the farm manager’s household is covered by the share of FWU in relation to the total work force on the farm and the household size measured in consumption units (adults).

Table 1 shows the distribution of the relevant variables in more detail: In addition to the mean values, values at the decile level are indicated. The values of the first decile are the ones of the farm at the first decile of the income distribution[5], whereas values of the ninth decile are the ones of the farm at the ninth decile. The value at the fifth decile represents the farm with the median income value.

Note that the values of the deciles for milk yield are steadily increasing, whereas those for concentrate inputs show a different tendency, almost steadily decreasing towards the ninth

[4] Note that any dummy variable related to different types of education of the farm manager and their partner is hard to validate for its significance.

[5] I.e. if we order our 5’459 observations by work income per FWU, from lowest to highest, the first decile would state the income (and values of all explanatory variables) of the 546th farm, whereas the ninth decile would reflect the values of the 4’913th farm.

decile. Although the first decile corresponds to the lowest milk yield, concentrate input for this decile is higher than for all other deciles.

Table 1: Deciles of the relevant explained and explanatory variables

Variable	Unit	Mean	Decile								
			1	2	3	4	5	6	7	8	9
Annual income per FWU	kCHF	**42.8**	−5.9	15.9	24.7	32.1	39.4	47.1	56.7	70.8	105.5
Farm size in livestock units (LU)	LU	**30.3**	24.8	22.9	25.6	26.4	29.6	31.1	32.8	35.9	44.0
Stocking density	LU/ha	**1.32**	1.34	1.26	1.31	1.28	1.28	1.31	1.34	1.37	1.38
Located in hill region	dummy	**0.41**	0.36	0.35	0.38	0.43	0.44	0.43	0.45	0.46	0.42
Located in mountain region	dummy	**0.40**	0.47	0.52	0.48	0.43	0.40	0.38	0.34	0.30	0.26
Located in steep terrain	dummy	**0.69**	0.70	0.73	0.73	0.72	0.70	0.68	0.69	0.64	0.59
Milk yield	kg/LU/ year	**6‘411**	6’155	6’047	6’188	6’291	6’421	6’455	6’609	6’665	6’874
Organic production	dummy	**0.16**	0.11	0.12	0.17	0.19	0.16	0.19	0.19	0.17	0.18
Free-stall housing	dummy	**0.29**	0.22	0.19	0.20	0.23	0.25	0.27	0.32	0.41	0.53
Input of concentrate feed per dairy cow	CHF/LU	**802**	961	807	796	783	797	782	793	740	757
Arable crops	dummy	**0.12**	0.09	0.10	0.12	0.13	0.14	0.11	0.13	0.14	0.16
Agriculture-related activities	dummy	**0.77**	0.71	0.70	0.76	0.76	0.79	0.80	0.83	0.78	0.79
Full-time farm	dummy	**0.31**	0.10	0.30	0.39	0.39	0.36	0.35	0.36	0.28	0.26
Farm manager’s partner with no additional education outside of agriculture	dummy	**0.29**	0.33	0.32	0.31	0.30	0.31	0.33	0.24	0.24	0.22
Share of non-family work units	%	**18.15**	20.77	12.21	12.62	12.83	14.39	16.40	18.65	22.76	32.72
Size of farm manager’s household	consump-tion units	**3.52**	3.48	3.31	3.54	3.56	3.71	3.57	3.63	3.49	3.41

3 Methodology

Two types of regression analysis are carried out, namely a random-effects model as a first step and a quantile regression as a second step. The latter serves as an indication of whether the explanatory variables contribute in the same manner to the financial success of the less and more successful farms. For both approaches, we use a panel data model, because our underlying data are based on two relevant dimensions, one spatial and one temporal.

Starting with all explanatory variables, the significance of each one is analysed by means of a t-test. The models in the results section (section 4) are presented with a minimum set of explanatory variables.

3.1 Random-effects model

We are interested in an analysis covering two types of differences: firstly, the differences which occur for single dairy farms over time, secondly, the differences which are present between different dairy farms within any given year. The latter type of difference warrants a cross-sectional analysis (Baltagi, 2013). Although for differences over time we could consider a fixed-effects panel model, incorporating both aspects requires us to use a random-effects model. A Hausman test can verify whether a random-effects model is appropriate to determine the contribution of the explanatory variables to the financial success of a dairy farm.

3.2 Quantile panel regression

Quantile regression allows us to estimate the significance and magnitude of the contribution of explanatory variables to the financial performance of a dairy farm in more detail than the random-effects model. With quantile regression, we can for example see if an explanatory variable contributes negatively for a lower quantile of the income distribution (i.e. the less successful dairy farms) and positively for a higher quantile of the income distribution (i.e. the successful farms).

As quantiles, we use decile intervals of the income distribution. This choice leads to a reasonable resolution (as opposed to e.g. quartiles with too little or centiles with too much detail) and a reasonably smooth path of resulting coefficients along the income distribution of successful dairy farms.

The literature presents several possible implementations or algorithms of a quantile panel regression. All algorithms have in common the minimization of a loss function $F(q; y(i,t), x_j$

(i,t)) which depends on the quantile q of the distribution, but there are several ways to perform the minimization. We choose the one proposed by Geraci and Bottai (2014), which is implemented in STATA and R. We choose this algorithm because – in line with the random-effects model – it does not rely on a fixed-effects transformation of the underlying data, as for example described by Powell (2016). Accordingly, the algorithm is based on a similar data treatment as the random-effects model, which also does not use a fixed-effects transformation. The way we choose whether to show one overall coefficient or a series of coefficients per quantile is determined as follows: Coefficients of explanatory variables which vary less than 0.5% are indicated with one overall coefficient – i.e. for any explanatory variable, the maximum absolute value of all coefficients in all quantiles v_{max} minus the minimum absolute value of all coefficients in all quantiles v_{min} divided by v_{min} should be less than 0.5%.

4 Results

Table 2 reports the results of the random-effects model. The statistical significance of the regression coefficients is indicated as follows: **Bold-faced coefficients** denote significance at a less than 1% probability level, i.e. a high statistical significance, whereas *coefficients in italics* denote a probability of greater than 10%, a low statistical significance. All other coefficients have probability levels between 1% and 10%. For the random-effects model, approximately a quarter of the variance can be explained (R^2 = 23.1%). The overall significance of the random-effects model is assessed by a Wald test and is very high (p-value < 0.001).

Table 2: Results of the random-effects model for the annual income per FWU (CHF/FWU)

Explanatory variable	**Unit**	**Coefficient**	**Standard error**	**p-value**
Farm size in livestock units (LU)	LU	**787**	45.3	<0.001
Stocking density	LU/ha	**−5’193**	1’160	<0.001
Located in hill region	dummy	*2’377*	1’800	0.19
Located in mountain region	dummy	*1’223*	2’031	0.55
Located in steep terrain	dummy	**−7’255**	750	<0.001
Milk yield	kg/LU/year	**3.4**	0.4	<0.001
Organic production	dummy	**7’731**	1’774	<0.001
Free-stall housing	dummy	3’236	1’406	0.02
Input of concentrate feed per dairy cow	CHF/LU	**−13.2**	1.1	<0.001
Arable crops	dummy	4’123	1’629	0.01
Agriculture-related activities	dummy	1’783	936	0.06
Full-time farm	dummy	**6’647**	883	<0.001
Farm manager’s partner with no additional education outside of agriculture	dummy	−3’761	1’339	0.01
Share of non-family work units	%	63.3	26.6	0.01
Size of farm manager’s household	consumption units	−761	367	0.04
Constant	-	**13’805**	3’834	<0.001

$R^2_{overall}$ = 23.1%; Hausman: p-value = 0.158

The farm size has a highly significant impact on the annual income per FWU. The contribution of an additional cow is CHF 787.- or approximately 2% of the annual income (CHF 37’600.-; Dux *et al.*, 2016). Both of the regional variables, i.e. hill and mountain region, show insignificant coefficients, whereas the location in steep terrain clearly reduces the annual income per FWU.

Milk yield, organic production, free-stall housing, diversification in arable crops, agriculture-related activities and full-time farming are determinants of successful farms. The size of the farm manager’s household and the input of concentrate feed per cow contribute negatively to the financial performance of a farm. The negative contribution of stocking density remains to be analysed further (e.g. by region). The share of non-family work units contributes positively to the annual income according to the random-effects model, but the quantile regression sheds a different light on this issue (Table 3).

Table 3: Results of the quantile regression for the annual income per FWU (CHF/FWU)

Variable	Unit	Coefficients for quantiles								
		10	20	30	40	50	60	70	80	90
Farm size in livestock units	LU	**836**	**821**	**839**	**840**	**841**	**844**	**841**	**857**	**851**
Stocking density	LU/ha	**−7’566**								
Located in hill region	dummy	3’861								
Located in mountain region	dummy	*−1’667*								
Located in steep terrain	dummy	**-8’588**								
Milk yield	kg/LU/ year	*−0.44*	*0.78*	1.63	**2.62**	**3.61**	**4.51**	**5.55**	**6.90**	**8.97**
Organic production	dummy	**6’957**								
Free-stall housing	dummy	4’364								
Input of concentrate feed per dairy cow	CHF/LU	**−15.8**	**−11.1**	**−10.9**	**−11.6**	**−12.5**	**−13.2**	**−13.1**	**−14.7**	**−14.0**
Arable crops	dummy	3’596								
Agriculture-related activities	dummy	*1’663*								
Full-time farm	dummy	**4’748**								
Farm manager’s partner with no additional education outside of agriculture	dummy	−3’491								
Share of non-family work units	%	*18.4*	*−4.3*	*23.5*	*24.6*	*25.3*	*28.3*	*25.5*	*51.0*	*43.8*
Size of farm manager’s household	consump-tion units	−1’017								
Constant	-	**15’518**								

Average pseudo-R^2 = 30.6%[6]

Four explanatory variables of the quantile regression model show differences above 0.5% – hence, a series of coefficients instead of a single one influence the annual income per FWU: farm size in LU, milk yield, concentrate input and share of non-family work units. Note that if we had used a criterion of non-overlapping 95% confidence intervals to decide whether to include a series of coefficients or a single coefficient, the series of coefficients would only have been applicable to the milk yield.

Whereas the estimated coefficients for farm size are in the range of a few percent, the other variables show highly differing results. Milk yield contributes clearly to the annual income per FWU for the higher deciles, whereas the impact is not significant for the lowest deciles. Concentrate inputs reduce income for all deciles, albeit to a varying degree, especially around the lower end of the income distribution.

[6] An average pseudo-R^2 computed for each decile according to the goodness-of-fit measure (Koenker and Machado, 1999) for quantile regressions (which corresponds to the same type of optimisation function as the quantile regression models) results in a value of 30.6%. It makes sense to consider the average because a different pseudo-R^2 value results for each of the deciles.

5 Discussion

The application of the two models can be seen as a sensitivity analysis. Whereas the random-effects model relates to the explanation of the mean value of the income distribution, the quantile regression at the fifth decile relates to the median.

Besides the four variables with quantile-specific results, the remaining 11 coefficients are similar in terms of the level of significance in the random-effects model and the quantile regression. However, the values of several coefficients differ between the two models. In two cases, the coefficients of the quantile regression are outside the 95% confidence interval of the respective coefficients of the random-effects model: The coefficient of stocking density is smaller in the quantile regression (with a value of −7’566 as opposed to −5’193), as is the coefficient for full-time farming (with a value of 4’788 as opposed to 6’647). The remaining nine coefficients of the quantile regression are within the respective 95% confidence intervals of the random-effects model but still show differences: Organic production, free-stall housing and diversification in arable crops show smaller coefficients in the quantile regression, whereas location in steep terrain and size of the farm manager’s household show higher values.

We maintain the set of variables to a reasonable extent. For example, we keep all regions (plain region as a baseline, hill region and mountain region, although the mountain region never significantly contributes to a change in income) because the region is very important for the characterization of dairy farms. One reason for the insignificance of the mountain region might be the introduction of the variable ‘located in steep terrain’. Whereas the region is strongly connected to both altitude and the duration of the vegetation period, the presence of steep terrain occurs not only at higher but also at lower altitudes.

In terms of education of the farm manager or their partner, theoretically 36 different, simple dummy variables[7] can be constructed. Of these dummy variables, six yield significant coefficients of which three are of a higher explanatory power than in the case where no educational dummies are considered at all. The case of the highest significance of a single variable yielding a satisfactory result according to the Hausman test is the case where the partner of the farm manager has no additional education outside the area of agriculture. Thus, we include this educational dummy in our set of variables.

Further variables were tested for inclusion but proved insignificant (by means of a t-test). These included existence of different animal species on the farm (e.g. calves, sheep), existence of different crops (e.g. sugar beets, cereal), different levels of education of the farm manager and his or her partner and age of the farm manager. In addition, these variables did not improve the overall explanatory power of the regression model ($R^2 < 0.1\%$ per variable).

6 Conclusions

The paper analyses the determinants of annual income per FWU for Swiss dairy farms based on data from the FADN by means of two regression types: a random-effects model and a quantile regression focusing on deciles. Milk yield and concentrate input are shown to be significant explanatory variables in both the random-effects model and the quantile regression. Although the random-effects model shows a significant positive effect of milk yield on income, the quantile regression reveals a much more detailed picture. Whereas milk yield has no significant impact on low income deciles, its importance is increasing with income. With an increase of almost CHF 9.- per additional kilogram of milk yield, the best performing decile benefits greatly, suggesting a thorough understanding of production technology and economic performance on these farms.

[7] This number is obtained by considering types of education (agricultural, housekeeping, other), levels of education (0, i.e. lowest, through 5, i.e. highest), for the farm manager or their partner.

The quantile regression reveals differing effects (i.e. a shift in coefficients across the deciles) for four explanatory variables and yields a single coefficient for each of the remaining 11 variables of the model. Accordingly, the quantile regression provides valuable information about the work income heterogeneity in the sample used. In further studies, quantile regression could be applied to explain the performance of other farm types and enterprises.

References

Baltagi, B. H., 2013. Econometric Analysis of Panel Data. West Sussex, John Wiley & Sons.

Dux, D., Schmid, D., Jan, P., Hoop, D. & Renner, S., 2016. Die wirtschaftliche Entwicklung der schweizerischen Landwirtschaft 2015: Hauptbericht Nr. 39 der Zentralen Auswertung von Buchhaltungsdaten Stichprobe Einkommenssituation [The economic development of Swiss agriculture 2015: Main report no. 39 of the Swiss FADN]. Agroscope Transfer. 143, Ettenhausen.

Geraci, M. & Bottai, M. 2014. Statistics and Computing, 24:461-479.

Hoop, D., Zorn, A. Lips, M. & Gazzarin, C., 2015. Determinants of Full Costs of Dairy Production in Switzerland – a comparison of two disproportionate Joint Costs Allocations, 20th International Farm Management Congress, 12-17 July, Québec City.

Koenker, R. & Machado, J., 1999. Goodness of Fit and Related Inference Processes for Quantile Regression, *Journal of the American Statistical Association*, 94(448):1296-1310.

Powell, D., 2016. Quantile Regression with Non-additive fixed effects.

Roesch, A., 2015. Impact of the SO Threshold on the Statistics of Economic Variables for the Swiss Agricultural Sector, *German Journal of Agricultural Economics*, 64(1): 33-41.

Acknowledgements:

The authors are grateful to the FADN team of Agroscope for providing the data.

Dear reviewer we would like to thank you for your comments. Given the word limit of 3500 words, we were not able to expand our paper. (The paper has 3394 words.) However, below we answered your questions (**in bold letters**) carrying out additional analyses.

Despite only circa 30% of the variation in income per AWU being explained by the two regression methods the data set is so large and covers 5 years it should be possible to explore some common dairying questions as in the following suggestions:

1) The residuals could be used as a measure of variation of managerial competence. What do the residuals look like? Most simply, what are the co-variances with variables which might be linked to managerial capacity?

All of the managerial competence data that was available has been tested for inclusion into the model, based on overall improvement of the explanatory content (as measured by R^2) while not violating the model's underlying constraints (Hausman test for random-effects model). The following variables were excluded, based on these tests:

- Age of farm manager
- Age of farm (i.e. how long has the farm been operative)
- Training of the farm manager measured by 11 different variables: in agriculture (5 levels – low to high), in house economics (3 levels – low, middle, high) , outside the two before-mentioned sectors (3 levels – low, middle, high)
- Training of the farm manager's partner – measured by 11 variables, structured identically to the farm manager's training: here, one variable (low level training of farm manager's partner outside the agricultural or household sector) is kept in the model.

We currently do not have any other obvious variables at our disposal within the (however large) data set to measure managerial competence.

2) Can you differentiate between scale and size efficiencies? This is a common debating point in agricultural policy. I think your results suggest both are present as larger herds have higher incomes and are more efficient in having higher yields with no more concentrates.

We start with the following definition of the relevant terms: Economies of scale describe how much production increases when the firm increases its *scale* of production, i.e. increases all (both fixed and variable) inputs by a common proportionality factor. Economies *of* size describe what happens to *cost per unit of output* when production increases in a cost minimizing way.[8]

Indeed our results suggest the following:

There are *scale efficiencies* with respect to size of the farm in livestock units and milk yield. The analysis of the milk yield depending on concentrate and roughage input quantities would require a paper in itself which would be much more related to animal science.
This is suggested both by the random-effects model and the quantile regression.

There are *size efficiencies* with respect to concentrate use: Cost per kg of milk decreases (from lower performing to higher performing farms; see Table 1 below), if we look at it from a purely concentrate feed cost point of view.

Table 1: Output, concentrate costs and cost per kg milk for nine deciles

Decile	1	2	3	4	5	6	7	8	9
Output (kg milk)	152'644	138'476	158'413	166'082	190'062	200'751	216'775	239'274	302'456
Concentrate costs (CHF)	23'833	18'480	20'378	20'671	23'591	24'320	26'010	26'566	33'308

[8] *p. 111, Svend Rasmussen, Production Economics, Springer Verlag, Berlin Heidelberg 2013.*

Cost / kg milk	0.156	0.133	0.129	0.1245	0.1241	0.121	0.120	0.111	0.110

3) Are your lower income producers spending more on concentrates or using more - I think your units are monetary ones hence the question. Are they buying in food on their smaller farms in an attempt to gain the marginal gains of extra cows? If so, from the figures quoted it would appear to be likely to be profitable at the margin?

The question cannot be answered since the only figure, which is present in our data set is the monetary data. So it could either be that the lower performing farmers buy more concentrates or that they buy less (or an equal amount, say in terms of nutritional value, to their better performing counterparts), but for a worse price than their better performing peers.

4) A common long standing policy statement is that that the agricultural income problem can be overcome by shifting low performers into the higher performing classes! What do your results say about how his might be done? In that you have data over 5 years do individual farms stay in the same class each year? If not then performance between years may be random.

It is not easy to compare the overall set of data, since it is an unbalanced panel where farms enter and leave the panel. We can think of the following:

- We can divide the panel into slices where farms stay for n consecutive years (n=2, 3, 4, 5).
- We can further measure the mean and standard deviation of the maximum jump between deciles. The maximum jump is defined as follows: max_jump = max (decile,t) – min (decile, t).

E.g. if a farm was present in the data set for three years, and the deciles d(t) where it resides with respect to the income distribution of the year t were as follows: d(1) = 8, d(2) = 3, d(3) = 5. Then, the maximum jump would be: max_jump = 8 – 3 = 5.
We can then simulate a random distribution where each farm for a given number of farms (say, n=10’000) gets attached a decile (d =1, …, 9) over 5 years (t=1,…5). The assignment are independent. For the random sample, we can as well determine the average of maximum jumps and the respective standard deviation.
For the average values of the maximum jumps (for both our sample) and the random sample we can perform a t-test of testing whether the mean values are statistically different. The high values obtained by the test statistics (column to the far right of the table below in Table 2) show that the mean values can indeed be considered statistically different. Moreover, the mean values of the maximum jump obtained by our actual sample are lower than the ones of the random sample. This can be considered a confirmation that the farms are not showing a “random financial performance”, but the financial performance of the farms stays within a certain range of performance. (It is however possible that farms transition between deciles over time – due to economic development and due to the fact that the set of farms changes over time).

Table 2: Five consecutive years for a sample used and a random sample

Consecutive Years	Number of farms (n)	Average maximum jump	*Std. deviation (max jump)*	Average max. jump – random sample	*Std. deviation (max jump) – random sample*	t-value
2	249	1.2	*1.3*	3.0	*2.1*	20.5
3	193	2.2	*1.7*	4.4	*1.7*	18.6
4	199	2.1	*1.6*	5.3	*1.5*	28.2
5	461	2.5	*1.7*	6.9	*1.3*	55.9

The analysis of the change of the relative performance would be another paper. As far as we know there is no consistent approach in the literature to this issue.

5) ***Can you use your equations to examine the resilience of low and high performance when price and other shocks occur? Low input farmers using a high proportion of non-tradable inputs such as family labor are possibly more secure. Please explore.***

Doing this kind of analysis is not easily based on our data without a major effort (warranting a new paper), as price shocks would have to be translated into shocks of the variables used within our model and, in addition, a realistic model of price shocks would have to be developed.

6) Finally if your FADN results have balance sheet data then are your better performing income farmers becoming relatively richer? (cf. Piketty on Capital)

It is not easy to measure what "relatively richer" could mean. There are several ideas how to measure it, but theoretically these measures could fail. The main reason is that money gained as income could be taken out of the farming business and thus not be reflected in the balance sheet of the farm. If we assume that the amount of money not covering personal expenses of the family of the farmer (or manager of the farm) is not taken out of the business, we could assume (and measure whether our assumption is correct) that the farms with relatively more income grow relatively more in terms of total farm assets or total farm equity.

We restrict ourselves again, as in question 4, to the part of the data which is consecutively present for a certain number of years (n=2,3,4,5).

For these, we analyze total equity and total assets of the farm over the years for which the farm is present within the data set. For different time frames, the results of this analysis are pictured below in Figure 1.
We measure the difference of the average equity between the last year in which each farm is present in the data set and the first year for which each farm is present in the data set, per average decile to which the farm belongs during its presence in the data set.
If the line displays a positive slope, it means that the equity for the higher decile (i.e. the more successful farms) grows more than for the lower decile (the relatively less performing farms). In that sense, we could say for a positively sloping line that the more successful farmers get relatively richer than their counterparts who do perform less well.
The same kind of argument could be used for Figure 2, where we analyze total assets. The analysis total assets could of course be obfuscated by a large inflow of external capital (debt).

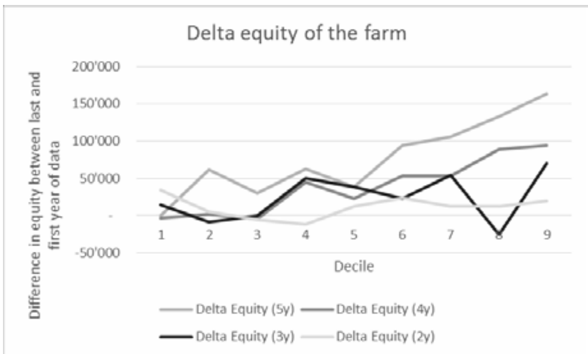

Figure 1: Delta in equity of the farm

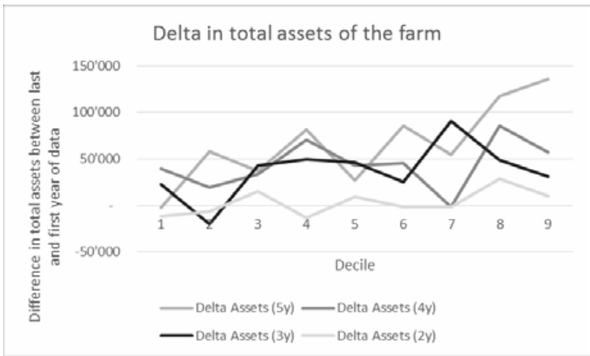

Figure 1: Delta in total assets of the farm

7) *How relatively and financially stable are your high and low performers as measured by their debt to asset ratios and their ability to withstand adverse external financial events while holding so called "growth stocks" such a land.*

Again, introducing shocks is not easily possible, but assessing the indicated ratios per decile and year is feasible and indicated below.

The debt-to-equity ratio and the amount of land held by the farmers can be analyzed in analogy to question 6.
The farms which are part of the data set for n (n=2,3,4,5) consecutive year are considered separately. For each separate set of farms, the average debt-to-equity ratio (Table 3) as well as the amount of land (Table 4) is computed for each "average decile" where each farm belongs to for each year. Then, the stability of these values – as well as the level of these value among the deciles of the distribution – is analyzed.
Results for the amount of land (Table 4) tend to be very stable over the years, and the land owned by the (on average) better performing farms is consistently more than what the lower performers own.
Results for the debt-over-equity ratio are less stable (Table 3).

Table 3: debt-to-equity ratio for all deciles

D/E				
Decile	Mean (5y)	Mean (4y)	Mean (3y)	Mean (2y)
1	1.51	0.82	0.91	0.15
2	1.63	0.84	-0.08	3.92
3	1.42	1.45	2.35	3.16
4	1.94	1.44	1.40	1.84
5	-0.87	0.58	0.88	1.99
6	1.37	0.91	1.56	2.12
7	1.23	-4.16	1.38	1.56
8	**-13.81**	2.13	0.07	5.55
9	0.85	-2.04	1.50	4.31
Average coefficient of variation	***-0.77***	***0.19***	***1.40***	***0.74***

The low value of the debt-to-equity ratio (highlighted in bold) is due to one outlier which could arguably be excluded, e.g. by restricting our analysis to farms which display a positive equity figure in the balance sheet data.

Table 4: Total owned land / ha for all deciles

Total owned land / ha				
Decile	Mean (5y)	Mean (4y)	Mean (3y)	Mean (2y)
1	8.9	10.1	10.7	13.7
2	11.7	15.2	14.5	14.4
3	15.1	10.7	17.7	17.0
4	15.8	14.3	12.7	16.6
5	15.6	15.8	14.6	15.9
6	18.8	15.3	15.8	15.9
7	18.9	20.2	21.9	16.7
8	22.5	21.9	16.0	17.0
9	20.6	24.7	32.6	19.1
Average Coefficient of Variation	***0.01***	***0.02***	***0.01***	***0.01***