Examples of statistical analyses with regard to policy assessment and farm performance

Werner Kleinhanss

Thünen Institute of Farm Economics

22th PACIOLI Workshop

Dublin (IRE) 29.09 – 01.10.2014

Structure

1. Introduction
2. Econometric estimation of standard income for taxation of small farms
3. Use of a Generalized Linear Model wrt analysis of income and farm performance
4. Conclusions

1 Introduction

- **FADN opens manifold options for classical statistical analysis, sophisticated modelling and simulations wrt to evaluation of policy measures, efficiency and performance of farms**
- **Two examples:**
  - Econometric estimation income
  - Use of a Generalized Linear Model wrt analysis of income and farm performance
  -

2 Estimates of standard income for taxation of small farms

According to § 13 Income Tax Act…

- **the profit of small farms can be determined based on standard rates (UAA <20 ha (excluding i.e. horticulture) , LU < 20, …)**
- **based on the 'economic farm value', where rates are differentiated by farm size (Wesche, 2009)**

Ministry of finances asks for a new methodology and calculation base for standard income

- **Agric. Ministry (BMEL) tested different measures based on TBN (proposal)**
- **Author's proposal to identify main influencing factors and to (econometric) estimate income based on TBN data**

→ Simulations of different schemes

Data and method

Farm sample selected form TBN data 2006/7 to 2012/13 according to the following criteria:

- **UAA < 20 ha, LU < 50, areas (ha) of hops <0.6, asparagus < 0.8, fruits < 1.85, strawberries < 0.6, horticulture open field 0.5, …, vines < 0.66**
- **Standard income (proposal) = UAA*300 + LU > 20*50/100 + UAA_hops*3815 + UAA_hortic_xx*yy (standard values)**

Econometric model

Pre-test of variables and data preparation

- Economic farm value used for § 13A (old) not significant
- Exogenous variables with high inter-correlation aggregated
- Outlier observations excluded

Stepwise (forward) regression model (no intercept):

- Profit = UAA_hops UAA_asparagus UAA_fruits UAA_strawberries UAA_nurseries UAA_flowers/ornamentalPlants UAA_vine LU_dairycow UAA_arable

Regression results – Model statistics

Summary of Forward Selection						
Step	Variables	n Var	Partial R-Square	Model R-Square	F-Value	Pr > F
1	LU_Dairy	1	0.4846	0.4846	3255.34	<.0001
2	UAA_Arable	2	0.0605	0.5451	460.50	<.0001
3	UAA_Flowers	3	0.0056	0.5507	42.79	<.0001
4	UAA_Vine	4	0.0008	0.5515	6.12	0.0134
5	UAA_Asparagus	5	0.0005	0.5520	3.77	0.0522
6	UAA_Strawberries	6	0.0003	0.5523	2.67	0.1024

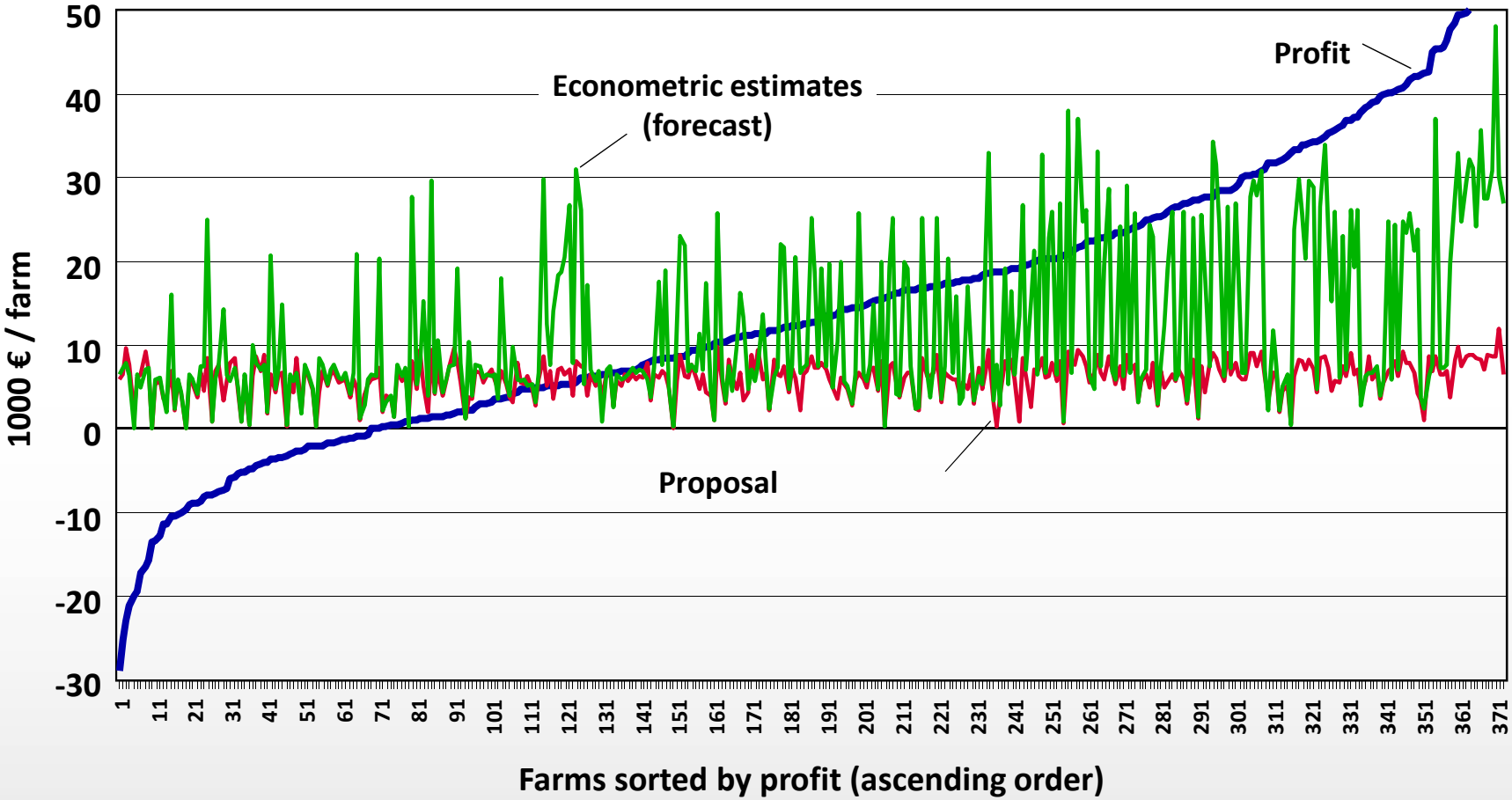
Parameter estimates

Parameter				
Variable	Parameter € / unit	STD	t-Value	Pr > |t|
UAA_Asparagus	19847	10216	1.94	0.052
UAA_Strawberries	38880	23795	1.63	0.102
UAA_Flowers	211573	35973	5.88	<.000
UAA_Vine	11803	4759	2.48	0.013
LU_Dairy	870	25	35.0	<.000
UAA_Arable	391	19	20.95	<.000

Parameters expressing income contribution per unit

- Dairy cows * 879 €, stat sign., greatest explanatory contribution
- Arable area * 391 €/ha, stat sign, 6 % contribution to R-Square
- Vines * 11,803 €/ha (seems to be realistic)
- Asparagus * 18,847 €/ha
- Strawberries * 38,880 €/ha (lower significance)
- Ornamental plants * 211,573 €/ha

Comparing profit, proposed standard and econometric forecast (2012/13)



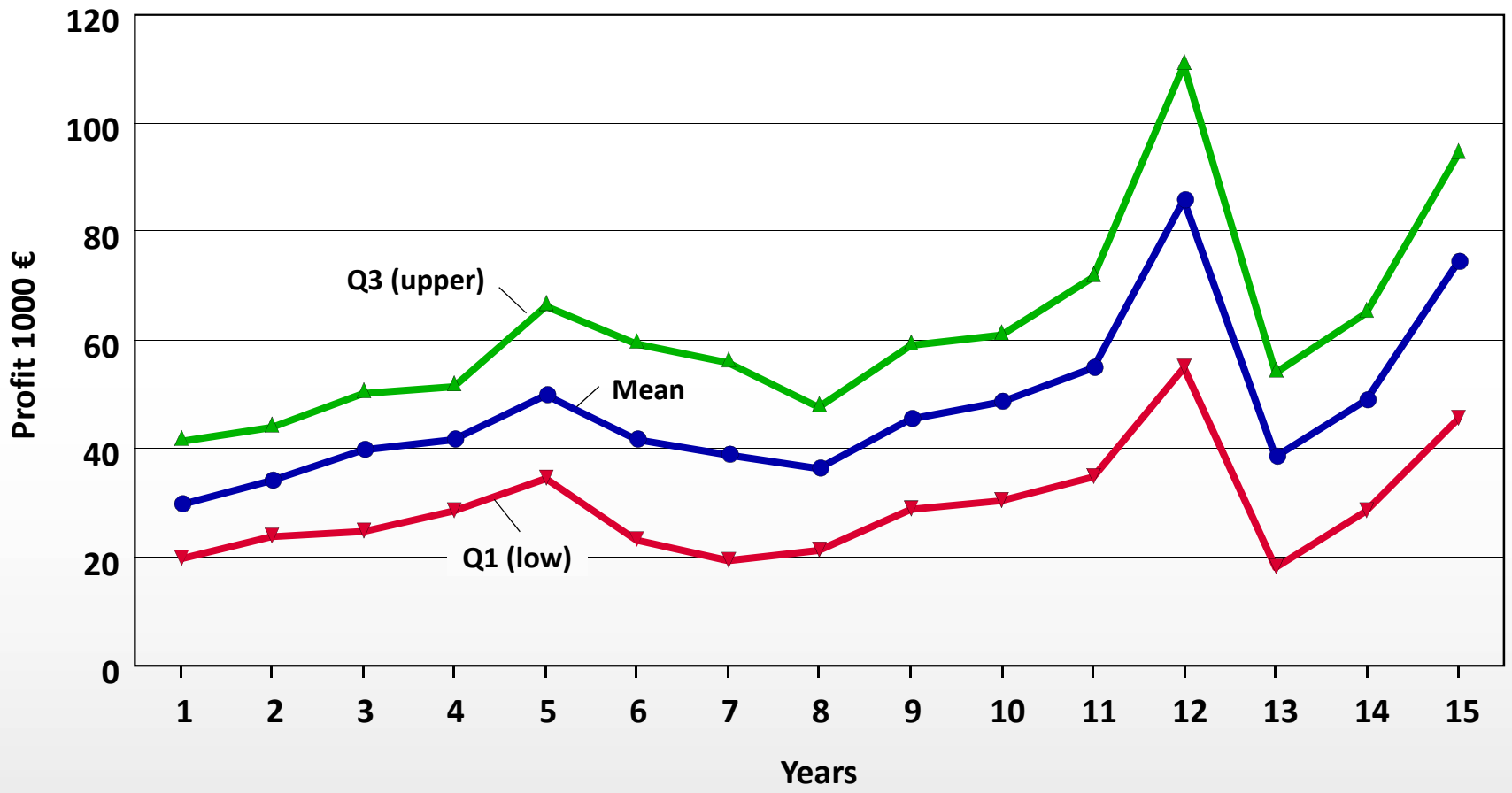
3 Use of a Generalized Linear Model (GLM) wrt analysis of income and farm performance

- GLM-Models, combining regression and variance analysis, allowing analysis of treatment effects based on the variance component
- Statistical procedures automatically defines dummy variables to measure fixed effects
- Different statistical tools (Univariate, Linear or Generalized Linear Models), the latter can be extended to random effects using OLS or ML estimators

Objectives and working procedure

- **Estimating profit functions over time, using farm dummies as indicators of long term farm (income) performance (use of SAS GLM)**
- **Data: Balanced sample of dairy farms over 15 years (1996/97 to 2010/11), Schleswig-Holstein, based on TBN data, removing outliers, 59 farms over 15 years**
- **Model: profit = O_crops O_milk O_otherLivstock O_subsidies&others I_Crops I_LivestSpec I_otherMaterials I_wages&deprec year farm_ID**
  - Covariates
  - fixed effects
  - +/- constant (different specifications)
- **Comparison with analyses based on income quartiles (number of farms being in income quartiles in each year**

Development and variation of income (profit €) 1996/97 (T=1) to 2010/11 (T=15)



Parameter estimates – covariates, time (1)

Covariates

	Coeff	Pr > ChiSq
Intercept	6896	0.1965
O_Crops	0.7966	<.0001
O-milk	0.7709	<.0001
O_oLiv	0.5683	<.0001
O_SubsiOth	0.6267	<.0001
I_Crops	-0.4654	0.0008
I_specLivest	-0.57	<.0001
I_othMateri	-0.8686	<.0001
I_WageDepr	-0.7214	<.0001

F-Statistic	50.04
R-Square	0.80

Fixed effect Time (T)

		Coeff	Pr > ChiSq
T	1	-1276	0.7058
T	2	-3691	0.2652
T	3	1922	0.5554
T	4	1746	0.5883
T	5	109	0.9717
T	6	-7252	0.0133
T	7	-3935	0.185
T	8	-8045	0.0059
T	9	-1336	0.6362
T	10	2518	0.3746
T	11	3408	0.2127
T	12	10307	<.0001
T	13	-3828	0.1517
T	14	-317	0.9029
T	15	0	.

Parameter estimates – farm_ID (2)

(best / worst performers wrt quartiles)

		Coeff	Pr > ChiSq
ID	10	-16833	0.0123
ID	71	-11877	0.017
ID	85	9468	0.0665
ID	87	-8742	0.1177
ID	93	-6991	0.2618
ID	121	7354	0.1392
ID	144	16057	0.0013
ID	*157*	*22290*	*0.0001*
ID	165	-11309	0.0194
ID	185	-5731	0.2471
ID	187	-1510	0.7854
ID	212	-676	0.8944
ID	213	994	0.8452
ID	233	5598	0.2676
ID	261	4964	0.3191
ID	263	10083	0.079
ID	265	-13646	0.0079
ID	266	-1378	0.7949
ID	273	9677	0.0563
ID	283	-9494	0.0827

		Coeff	Pr > ChiSq
ID	284	14217	0.0051
ID	305	-46296	<.0001
ID	306	4173	0.4092
ID	308	8415	0.0961
ID	321	-8842	0.0794
ID	324	7482	0.1409
ID	327	4670	0.3595
ID	331	-5283	0.317
ID	341	-122	0.9804
ID	342	-1623	0.7369
ID	*343*	*-12366*	*0.0139*
ID	350	4922	0.3277
ID	354	-7345	0.1664
ID	369	92	0.9857
ID	379	-4756	0.3427
ID	380	-11014	0.0323
ID	382	-15374	0.0016
ID	383	8028	0.1202
ID	403	-1412	0.8156
ID	405	-13723	0.0064

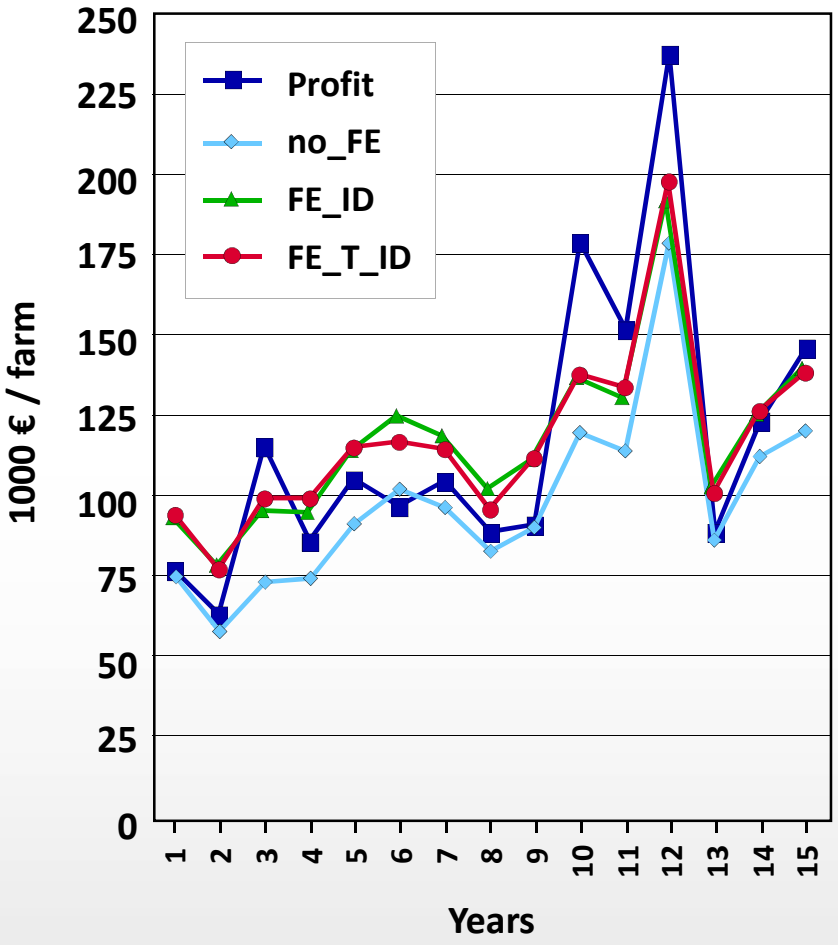
		Coeff	Pr > ChiSq
ID	409	-1239	0.8138
ID	417	-2884	0.5984
ID	418	-3359	0.5111
ID	430	-1469	0.7627
ID	436	7221	0.2586
ID	442	-1803	0.7166
ID	447	-3157	0.5569
ID	452	-3616	0.4686
ID	484	-1424	0.7751
ID	499	-501	0.924
ID	508	1838	0.7279
ID	524	-24133	<.0001
ID	541	-2989	0.5522
ID	542	-12254	0.0267
ID	564	-352	0.9457
ID	734	5817	0.2718
ID	735	6872	0.19
ID	736	-20291	0.0006
ID	741	0	.

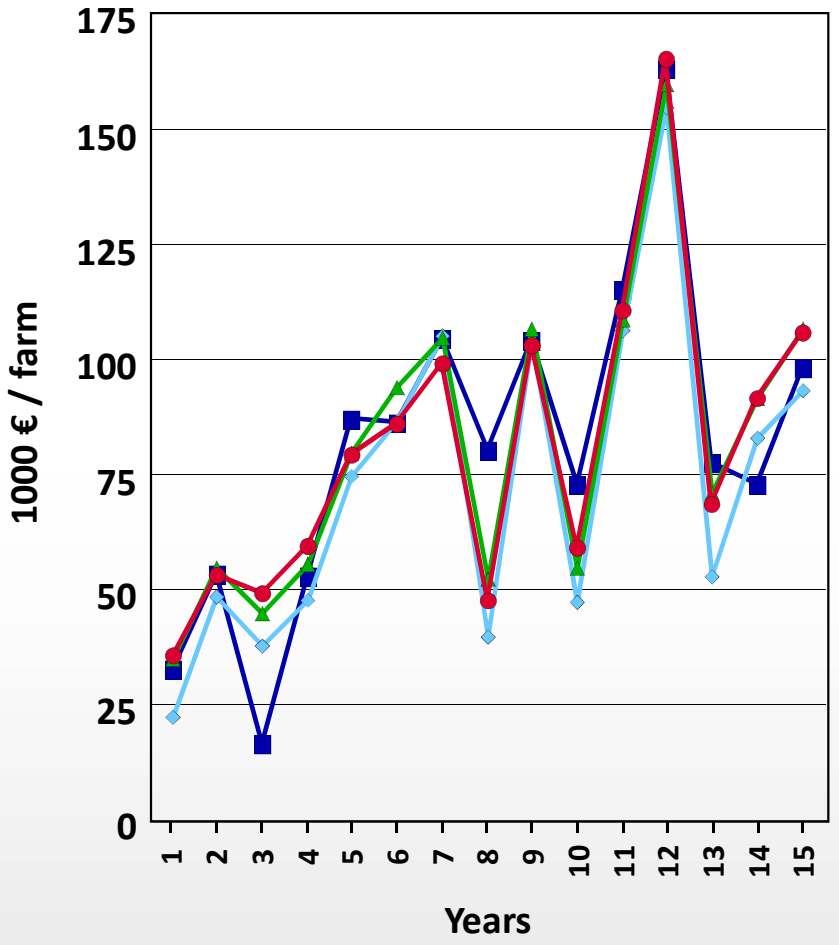
Comments

- **Statistically significant parameters and correct signs for covariates, partly for time, not at all for farm_ID's**
- **Coefficients for farm_ID not sufficient to identify / rank performance of farms**

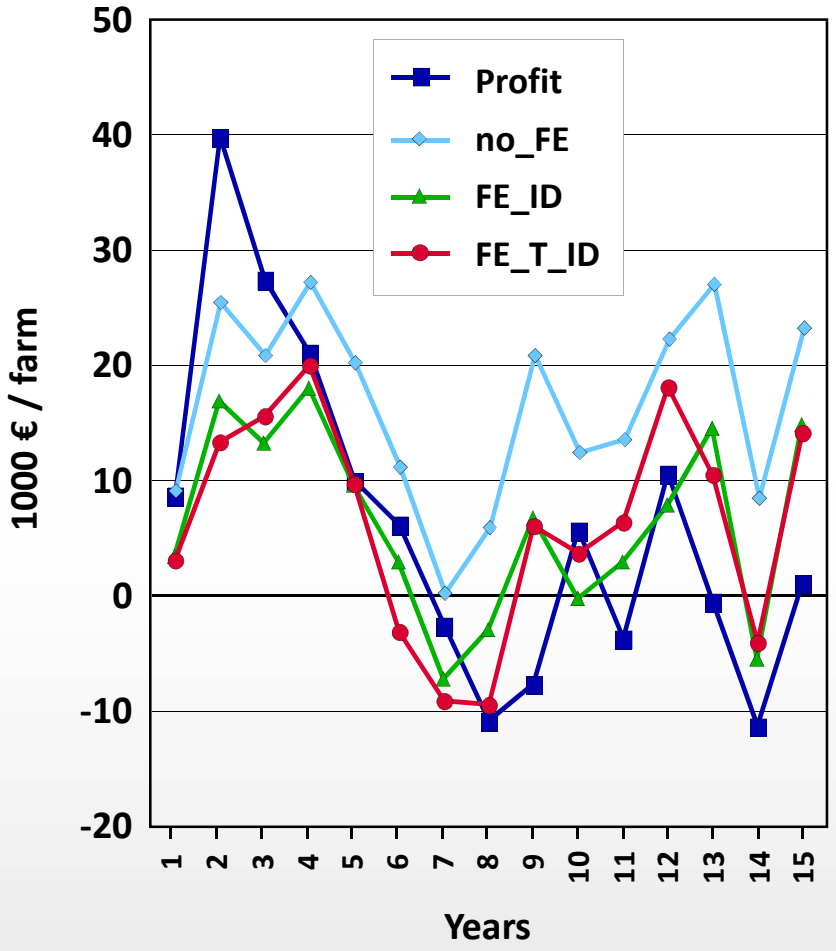
Appropriate ex-post income forecast for best performing farms, not at all for worst performers (see below)

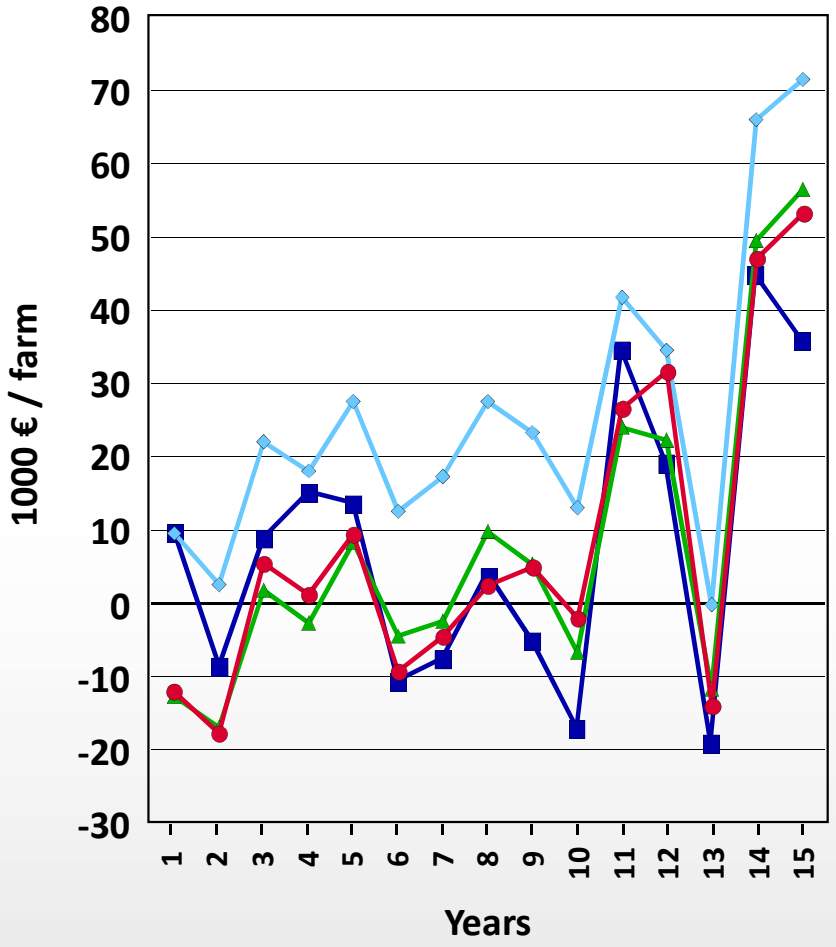
Profit (€) and forecast – example of best performers, different specifications wrt fixed effects



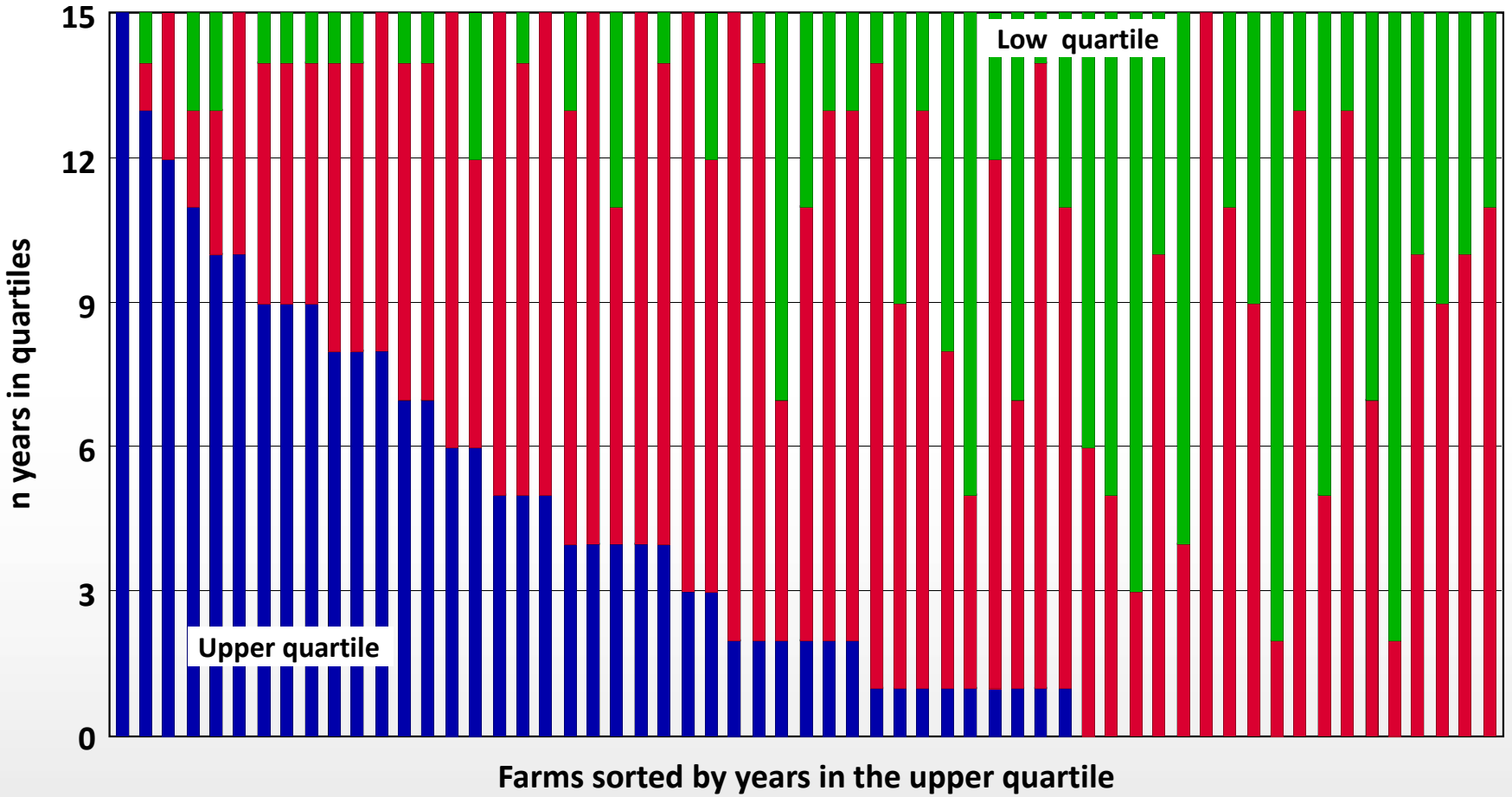


Profit (€) and forecast – example of worst performers





Share of farms being in quartiles



4 Conclusions

Estimation of standard income for taxation purposes

- **Existing standard method (economic farm value) not appropriate → statistically not significant**
- **Proposed standard income in most farms considered not higher than allowance → no income taxes**
- **Income based on regression results**
  - Positive, but higher variation, closer to ‚real' profit
  - If ‚real' income is larger than estimates farmers have to prove their ‚real' income based on bookkeeping data

Conclusions (2)

GLM Model application wrt farm performance

- **Statistically significant parameters and right signs for covariates, partially for time, not at all for farm dummies**
- **Coefficients for farm dummiy not sufficient to identify / rank performance of farms**

Delivers appropriate ex-post income forecasts for best performing farms, not at all for worst performers

- **Only 1 farm (of 59) being in the upper quartile in the whole period**
- **Staying in the same quartile is rare due to rather high variation of income in time**

Sources:

Wesche R (2009) Besteuerung der Land- und Forstwirtschaft. AID infodienst, Heft 1247 / 2009