An application of item response theory for agricultural sustainability measurement

Supplementary materials

Brian Beadle[1,2*], Stephan Brosig[1] and Christoph Wunder[2]

[1*]Department of Structural Change, Leibniz Institute of Agricultural Development in Transition Economies (IAMO), Theodor-Lieser-Straße 2 Germany, Halle (Saale), 06120, Germany.
[2]Department of Economics, Martin Luther University of Halle-Wittenberg, Große Steinstraße 73, Halle (Saale), 06099, Germany.

*Corresponding author(s). E-mail(s): brian.beadle@abac.edu;
Contributing authors: brosig@iamo.de; christoph.wunder@wiwi.uni-halle.de;

Supplementary materials contents

1 Indicator literature review

To organize the review, we categorize indicators into broader themes within each of the sustainability dimensions (economic, environmental, and social). In the economic dimension (Table 1), the themes we use are productivity, profitability, resilience, operating costs, other income, and a final category for miscellaneous topics. The environmental dimension (Table 2) includes themes for water use, emissions, pesticide use, land use/biodiversity, intensity [of production], and organic factors. Finally, the social dimension (Table 3) is subdivided between internal indicators (i.e. indicators that measure impacts to the farm owners) and external indicators (that have an effect outside of the farm).

Theme	Indicator	Study
Productivity	Labor productivity	Van Passel and Meul (2012), van der Meulen, Dolman, Jager, and Venema (2014), Ryan et al. (2016), Vitunskiene and Dabkiene (2016), Dillon et al. (2015), Lynch et al. (2016)
	Land productivity	Van Passel and Meul (2012), Ryan et al. (2016), Vitunskiene and Dabkiene (2016), Dillon et al. (2015), Lynch et al. (2016)
	Capital productivity	Van Passel and Meul (2012), Vitunskiene and Dabkiene (2016)
	Return to labor	Povellato, Bodini, Longhitano, and Scardera (2012)
	Return to land	Povellato et al. (2012)
Profitability	Economic viability	Ryan et al. (2016), Dillon, Hennessy, and Hynes (2010), Dillon et al. (2015), Lynch et al. (2016)
	Net farm income	van der Meulen et al. (2014), Ryan et al. (2016), Ehrmann (2010)
	Profitability	Dillon et al. (2015), Lynch et al. (2016)
	Market return	Dillon et al. (2010)
Resilience	Solvency	van der Meulen et al. (2014), Vitunskiene and Dabkiene (2016)
	Farm diversification	Vitunskiene and Dabkiene (2016)
	Farm risk management	Vitunskiene and Dabkiene (2016)
	Liquidity	Ehrmann (2010)
Operating costs	Expenditure for energy	Povellato et al. (2012)
	Expenditure for services to third parties	Povellato et al. (2012)
	Total contracting costs to total variable costs	Barnes and Thomson (2014)
	Total output value to total fixed and variable costs	Barnes and Thomson (2014)
	Total paid labor to farm gross margin	Barnes and Thomson (2014)
	Total rent and interest paid to farm gross margin	Barnes and Thomson (2014)
Other income	Total subsidies to farm gross margin	Barnes and Thomson (2014), Dillon et al. (2010)
	Other gainful activities	Povellato et al. (2012)
Miscellaneous	Market orientation	Ryan et al. (2016), Dillon et al. (2015), Lynch et al. (2016)
	Fixed capital formation	Vitunskiene and Dabkiene (2016)

Table 1: Review of economic indicators using FADN data from selected studies.
Source: Author's compilation based on literature reviews by Dabkienė (2016) and Kelly et al. (2018)

Theme	Indicator	Study
Water use	Percent of UAA under irrigation	Povellato et al. (2012), Westbury, Park, Mauchline, Crane, and Mortimer (2011)
	Water units per UAA	Westbury et al. (2011)
Emissions	GHG emissions per farm	van der Meulen et al. (2014), Ryan et al. (2016), Vitunskiene and Dabkiene (2016), Dillon et al. (2015), Lynch et al. (2016)
	GHG emissions per kilogram of output	Ryan et al. (2016), Dillon et al. (2015), Lynch et al. (2016)
	Emissions from fuels and electricity	Dillon et al. (2015), Lynch et al. (2016)
	Energy use	van der Meulen et al. (2014), Westbury et al. (2011)
	Energy use efficiency	Van Passel and Meul (2012)
	Ammonia emissions	Ehrmann (2010)
	Methane emissions	Dillon et al. (2010)
Land use	Grassland	Povellato et al. (2012), Barnes and Thomson (2014), Westbury et al. (2011)
	Shannon index	Westbury et al. (2011), Gerrard, Padel, and Moakes (2012), Ehrmann (2010)
	Percentage of farm that is wooded	Barnes and Thomson (2014), Westbury et al. (2011)
	Percentage of farm that is rough grazing	Barnes and Thomson (2014), Westbury et al. (2011)
	Percentage of farm that is fallow (i.e. set aside)	Westbury et al. (2011)
	Percentage of spring crops	Westbury et al. (2011)
	Meadow and pastures	Vitunskiene and Dabkiene (2016)
	Biodiversity	Vitunskiene and Dabkiene (2016)
	Land use limitations	Povellato et al. (2012)
Chemical use	Nitrogen balance	Ryan et al. (2016), Ehrmann (2010), Dillon et al. (2015), Lynch et al. (2016)
	Nitrogen use efficiency	Van Passel and Meul (2012), Dillon et al. (2015), Lynch et al. (2016)
	Nitrogen content	Povellato et al. (2012)
	Nitrogen surplus	Van Passel and Meul (2012)
	Crop protection (i.e. pesticides) per UAA	Povellato et al. (2012), van der Meulen et al. (2014), Westbury et al. (2011)
Intensity	Livestock units per hectare	Povellato et al. (2012), Vitunskiene and Dabkiene (2016), Westbury et al. (2011), Gerrard et al. (2012)
	Total output per UAA	Barnes and Thomson (2014)
	Energy intensity	Vitunskiene and Dabkiene (2016)
	Intensification indicator	Gerrard et al. (2012)
Organic	Organic farming	Povellato et al. (2012)
	Humus balance	Ehrmann (2010)
	Organic phosphorous	Dillon et al. (2010)
	Organic nitrogen	Dillon et al. (2010)

Table 2: Environmental indicators using FADN data from selected studies. Source: Author's compilation based on literature reviews by Dabkienė (2016) and Kelly et al. (2018)

Level	Indicator	Study
Internal	Farmer age	Povellato et al. (2012), Vitunskiene and Dabkiene (2016), Dillon et al. (2015), Lynch et al. (2016)
	Isolation risk	Ryan et al. (2016), Dillon et al. (2010), Dillon et al. (2015), Lynch et al. (2016)
	Farmer education	Povellato et al. (2012), Ryan et al. (2016), Dillon et al. (2015), Lynch et al. (2016)
	Household vulnerability	Ryan et al. (2016), Dillon et al. (2015), Lynch et al. (2016)
	Work-life balance	Ryan et al. (2016), Dillon et al. (2015), Lynch et al. (2016)
	Demographic viability	Ryan et al. (2016), Dillon et al. (2010)
	Family labor	Povellato et al. (2012), Vitunskiene and Dabkiene (2016)
	Farmer gender	Povellato et al. (2012)
	Altitude	Povellato et al. (2012)
	Distance from inhabited center	Povellato et al. (2012)
	Networking	Povellato et al. (2012)
	Total farmer hours to total hours worked	Barnes and Thomson (2014)
	Pluriactivity	Vitunskiene and Dabkiene (2016)
	Workload exceeded	Vitunskiene and Dabkiene (2016)
	Continuity of the farm	Vitunskiene and Dabkiene (2016)
External	Labor supply	Povellato et al. (2012)
	Somatic cell count	van der Meulen et al. (2014)
	Cow lifetime	van der Meulen et al. (2014)
	Grazing hours	van der Meulen et al. (2014)
	Total hired labor to total hours worked	Barnes and Thomson (2014)
	Jobs on farm	Vitunskiene and Dabkiene (2016)
	Wage ratio on farm	Vitunskiene and Dabkiene (2016)

Table 3: Social indicators using FADN data from selected studies. Source: Author's compilation based on literature reviews by Dabkienė (2016) and Kelly et al. (2018)

2 Item selection considerations and definitions

In the literature, **profitability** is measured in a variety of ways such as income relative to operation income (Ehrmann, 2010) and gross margin per hectare (Ryan et al., 2016). Perhaps more appropriate in the context of sustainability, van der Meulen et al. (2014) calculates profitability as a ratio of farm net income to unpaid annual work units (AWU). The inclusion of unpaid (family) labor in the equation can (a) ensure adequate accounting for farmer(s) income sufficiency (Gómez-Limón, Arriaza, & Guerrero-Baena, 2020), (b) control for variations in farm size (van der Meulen et al., 2014), and (c) reflect the potential for contributing to social sustainability attributes such as inter-generational succession of the farm (see, e.g. Glauben, Tietje, & Weiss, 2005) and the ability for farmers to contribute to the local economy. This method, however, is unsuitable for corporate or cooperative farms that do not have unpaid labor. Since all labor costs are already compensated in these cases, the proposed solution to measure family and corporate farms equally is to subtract from family farms' income an allowance for unpaid (family) labor. We compute this allowance as family labor input quantity times the median wage $\tilde{w}$ for the federal state the farm is located in (denoted by subscript fs) FNI as:

$$i_{profit} = FNI - \tilde{w}_{fs}, \tag{1}$$

where i_{profit} is the farm's profit and FNI is calculated as total output plus the balance of current subsidies and taxes, less depreciation, intermediate costs (specific costs and farm overhead costs), and total external factors (wages, rent, and interest paid) (European Commission, 2000). In this calculation, the variable is a absolute sum for the whole farm, which is then converted to a relative calculation using AWU as a denominator when converted to an ordinal item.

Solvency is calculated using the same equation as Vitunskiene and Dabkiene (2016), the indicator uses a common formulation as the farms' total debts to total assets:

$$i_{solvency} = \frac{D_T}{A_T}, \tag{2}$$

where D_T is the combined total T of short-, medium-, and long-term loans, and A_T refers to the total of both long-term fixed assets (land, buildings, machinery, and breeding livestock) and short-term current assets (non-breeding livestock, the stock of agricultural products, and other circulating capital) (European Commission, 2000).

Wage ratio WR is measured as the ratio of the average hourly wage paid on the farm (total wages paid to total paid labor hours) to the median wage of the region. In contrast to Vitunskiene and Dabkiene (2016) who calculates this indicator as the ratio of average annual wages for farm workers to average wages in the whole country, the calculation for this index is the ratio of average wages on the farm to the median income in each NUTS3 region to capture regional differences in purchasing power and the cost of living:

$$WR = \frac{\overline{w}_h}{\overline{w}_h^{n3}}, \tag{3}$$

where $\overline{w}_h$ is the mean hourly h wage on the farm and $\overline{w}_h^{n3}$ is the mean hourly wage in the NUTS3 $n3$ region the farm is located in.

Whereas the item for **economic diversity** ED is calculated as a dummy variable used in SDG 2.4.1 to signal if a single agricultural product accounts for more than 66% of total output on the farm, we instead calculate the indicator as a continuous value:

$$i_{ED} = max\left(\frac{\phi_n}{O_T}\right), \tag{4}$$

where the value for economic diversity is the maximum max contribution of a single nth product ϕ to the farm's total output O_T, with n representing the 19 products specified in FADN such as grains, milk, wine, etc.

The **provision of employment** indicator PE is calculated as the ratio of total expenditure on wages w_T and contracted work c to total output O_T to control for farm size and reflect the intensity to which farms are providing income opportunities for agricultural workers and businesses in the region:

$$PE = \frac{w_T + c}{O_T}, \tag{5}$$

with contracted work including the expenditures for both contracted services and hired machinery (European Commission, 2000).

The **expenditure on pesticides** EP is measured as the total pesticide expenditure EP_t (in €) over the farm's land area in UAA L_{UAA}:

$$EP_{UAA} = \frac{EP_T}{UAA_{ha,T}}. \tag{6}$$

The item for **GHG emissions** measures the total emissions produced by the farm and includes both indirect emissions (e.g. nitrogen from fertilizers) and direct emissions (e.g. CO_2 from fuel consumption). The framework of emission sources used for estimating total emissions developed by Coderoni and Esposti (2018) and Coderoni et al. (2013) is modified to suit German agriculture. For example, variables for CH_4 from rice cultivation and biological N fixation are not included in the German FADN data set and are thus omitted from the calculation. All calculations are performed in accordance to IPCC (2006) tier-1 and tier-2 estimates and summed to produce a total level of emissions for the farm:

$$GHG_{CO_2-eq} = \frac{\Sigma_{s=1}^{11} GHG_s \times GWP_v}{VA_G}, \tag{7}$$

where the level of CO_2 equivalent greenhouse gas emissions GHG_{CO_2-eq} on the farm is the sum over 11 sources s of greenhouse gases converted to CO_2 equivalents using their global warming potential values GWP_v (see Eurostat, 2017), as shown in Table 4. To reflect the farm's $CO_{2\text{-eq}}$ intensity, total emissions are divided by the farm's gross value added VA_G (see Umweltbundesamt, 2007), which is the total monetary value received by the producer including subsidies, less taxes and intermediate consumption (Eurostat, 2022).

We define **multi-factor productivity** MFP as the quotient of total value added VA over the total cost of using the three factors land, labor, and capital, denoted by X_A, X_L, and X_C:

$$MFP = \frac{VA}{X_A + X_L + X_C}, \tag{8}$$

where VA is computed as total output O plus the balance of subsidies and taxes S minus total expenditure on intermediate inputs X_{int} (i.e. specific costs and farming overheads):

$$VA = O + S - I_{int}. \tag{9}$$

Factor input costs are are estimated on an annual basis (i.e. one full agricultural season) (see FAO, 2018, p. 57) to reflect the opportunity costs of using the respective factor for production. The cost of land input X_A includes land rental expenditure X_Ar for rented land and the opportunity costs for owned land X_Ao. The latter are estimated as the quantity of owned land in hectares UAA_o (in ha) times the mean land rental price $\overline{r}$ (in Euro/ha) paid by all farms of the respective type (EU TF14 classification) in the respective region:

$$X_A = X_Ar + UAA_o \times \overline{r}. \tag{10}$$

Similarly, labor input X_L is an aggregation of the farm's total wage expenditure for paid labor X_Lp plus opportunity costs of unpaid labor estimated as the total quantity of unpaid labor in annual work units AWU_u times an expected return to labor based on the average wage per AWU of paid labor $\overline{w}_p$ on farms in the respective region:

$$X_L = X_{Lp} + AWU_u \times \overline{w}_p. \tag{11}$$

Finally, the value for capital input X_C is calculated as the total interest paid by the farm X_{Cb} plus the net worth of the farm capital C_o (without land values) multiplied by an assumed interest rate ι_b of 4%:

$$X_C = X_{Cb} + C_o \times \iota_b. \tag{12}$$

In contrast to other studies that use the Shannon Index (e.g. Gerrard et al., 2012; Westbury et al., 2011) measuring the number of and the evenness of distribution among different types of land use, **land ecosystem quality** LEQ assigns ecological quality scores to agricultural land based on the type and intensity of land use on the farm. The scale ranges from zero (for worst ecological quality) to one (best). The total ecosystem quality value LEQ_T is calculated as:

$$LEQ = \frac{\Sigma_{ip}(L_{ip} \times EQ_{lt})}{L_T}, \tag{13}$$

where L_{ip} is the total land in land use type ip (e.g. cereals, vineyards, etc. or woods, agricultural fallows, set aside), which is weighted by the ecological quality score EQ associated with that land use type ip (see Table 5). The sum of products is then divided by the total land area of the farm: L_T. All land values are reported in hectares and the EQ_{lt} values shown in Table 5 are derived from Reidsma, Tekelenburg, Van den Berg, and Alkemade (2006).

GHG type	GWP_v	GHG emission source (GHG_s)
CH_4	25	Manure management
		Enteric fermentation
N_2O	298	Manure management
		Synthetic fertilizers
		Crop residue
		Atmospheric deposition
		Leeching and run-off
CO_2	1	Energy
		Forest land
		Cropland
		Grassland

Table 4: GHG emission sources. The variable GWP_v is the multiplier to convert the GHG type to CO_2 equivalent units for equation 7. Source: Adapted from Coderoni et al. (2013)

EQ_{lt}	Description of land in production L_{ip}	O	Ir	G	LU	In
0.05	Irrigated		X			
	Highly intensive					>250
0.10	Intensive					80-250
0.15	Highly intensive organic	X	X			>250
	Intensive arable grazing livestock			<66%	>80	
0.20	Intensive organic	X				80-250
	Intensive pasture			>66%	<2	<250
	Highly intensive pasture			>66%	>2 OR	>250
0.25	Extensive					<80
0.325	Extensive arable grazing livestock			<66%	<1	<80
0.35	Extensive organic	X				<80
0.4	Extensive pasture			>66%	<1	<80
1	Natural grassland			>66%	<0.3	
	No production					

O = Denotes if the farm is organic, partially organic, or transitioning to organic

Ir = Denotes if the farm has installed irrigation

G = Percentage of land for forage crops (% of total used land

LU = Number of livestock units per hectare of used land

In = Value of direct inputs (fertilizers, pesticides, and feedstuffs for grazing) per hectare of used land

Table 5: Percentage values of land ecosystem quality. Source: Reidsma et al. (2006)

3 Item category distributions

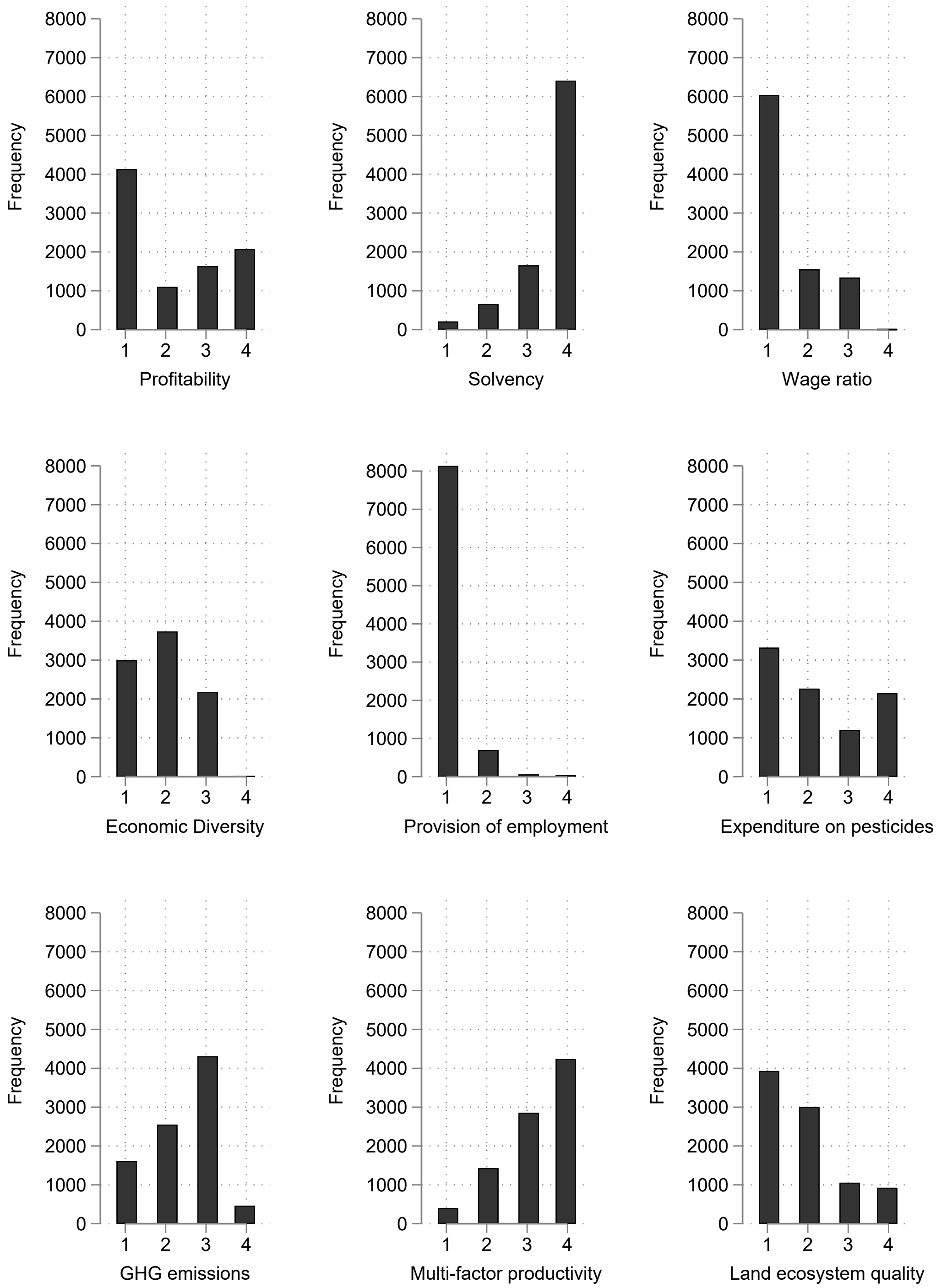

Fig. 1 Frequencies for each category of the nine sustainability items

4 Sensitivity analyses using 3- and 5-category items

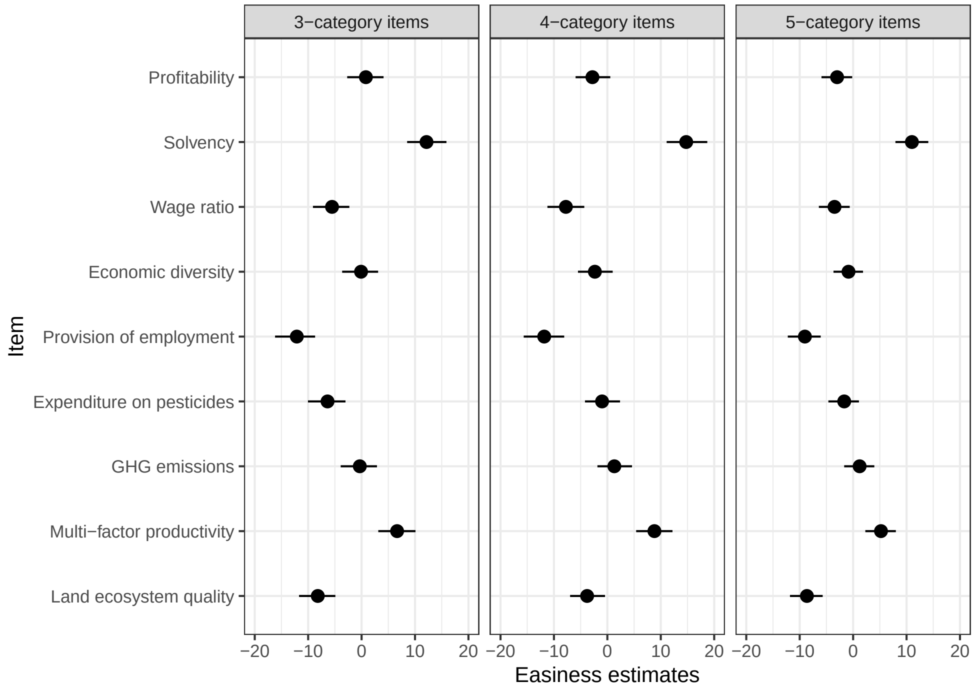

Fig. 2 Easiness parameters of the 3-, 4-, and 5-category models

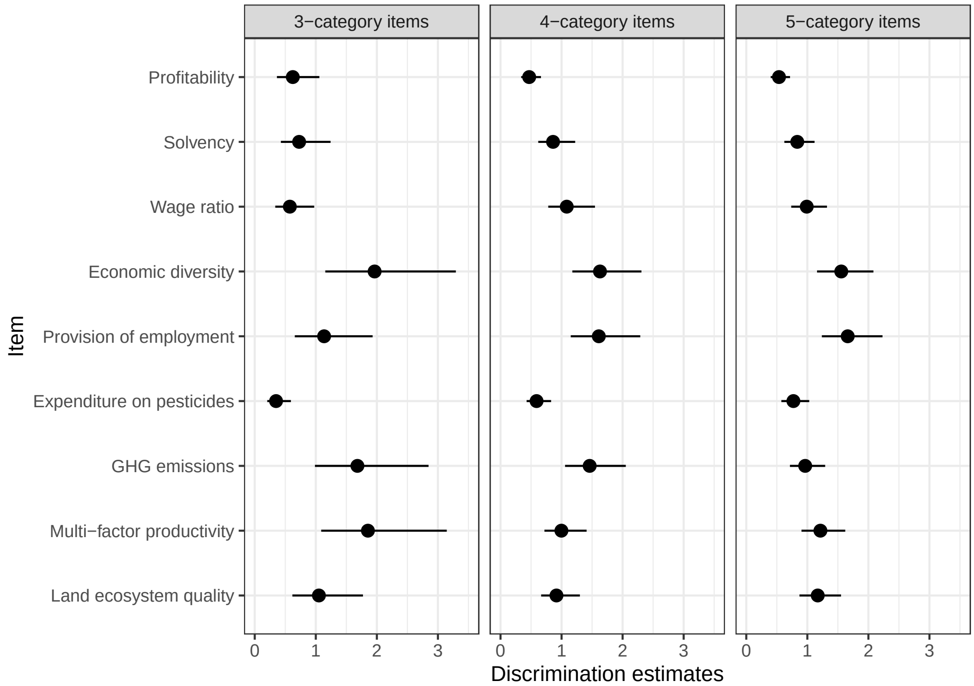

Fig. 3 Discrimination parameters of the 3-, 4-, and 5-category models

5 Results

5.1 Model comparison

In this section, we compare a restricted IRT model without a discrimination parameter and an unrestricted IRT model with a discrimination parameter to determine which model has a better fit. We use approximate leave-one-out cross-validation (LOO-CV) to measure the predictive accuracy for the purpose of model comparison Vehtari:etal:2017. LOO-CV works by leaving out one data point from the training set, fitting the model to the remaining data points, and then scoring the model on the left-out data point. This process is then repeated for each data point in the training set.

LOO-CV helps us identify potential issues with overfitting or underfitting our data while also providing an objective measure for choosing between models of different complexities. This measure is the expected log pointwise predictive density (ELPD), which is an estimate of the out-of-sample predictive accuracy for each model. The ELPD can be used to assess how well a probabilistic model can explain an observed set of data points. The model with the highest ELPD best explains our observations while minimizing overfitting or underfitting issues.

Table 6 presents the results of a comparison between the two models, one with and one without the discrimination parameter. The ELPD values for both models are shown along with their respective standard errors. Furthermore, the difference in ELPDs and its corresponding standard error is also reported. From these findings, we can conclude that the model including the discrimination parameter fits the data substantially better than that without the discrimination parameter; this is indicated by the substantially larger ELPD of the unrestricted model. Therefore, due to its superior performance on predictive accuracy, we proceed with the model that includes the discrimination parameter.

	ΔELPD	s.e.(ΔELPD)	ELPD	s.e.(ELPD)
Model with discrimination	0.0	0.0	-83991.7	192.8
Model without discrimination	-3148.7	77.1	-87140.4	188.8

ELPD is the expected log pointwise predictive density.

Δ ELPD is the difference in ELPDs.

s.e. denotes the standard error.

Table 6: Model comparison using LOO-CV

5.2 Item characteristic curves

Item characteristic curves (ICCs) are used to describe the relationship between the probability of responding in a *particular category* (or higher) relative to the latent sustainability of the farm, θ. We present ICC plots for the categories of “very sustainable”, “sustainable”, and “unsustainable”.

To interpret the ICC plots, we are interested in the slope of each curve as well as its location on the x-axis. The maximum of the curve's slope corresponds with the discrimination of the item, where a steeper slope corresponds to higher discrimination since a relatively small change in θ will cause a relatively large change in the predicted probability of a particular item category. The location of the curve corresponds to the item easiness for the category, where a curve shifted to the left side of the plot corresponds to an easier item category. The location of the ICCs on the x-axis not only tells us the easiness of each item, it also provides information on the variability of the easiness of the items. For example, if all ICCs were clustered to the left side of the x-axis, the model would be comprised solely of easy items and would not provide much information on farms with a high θ.

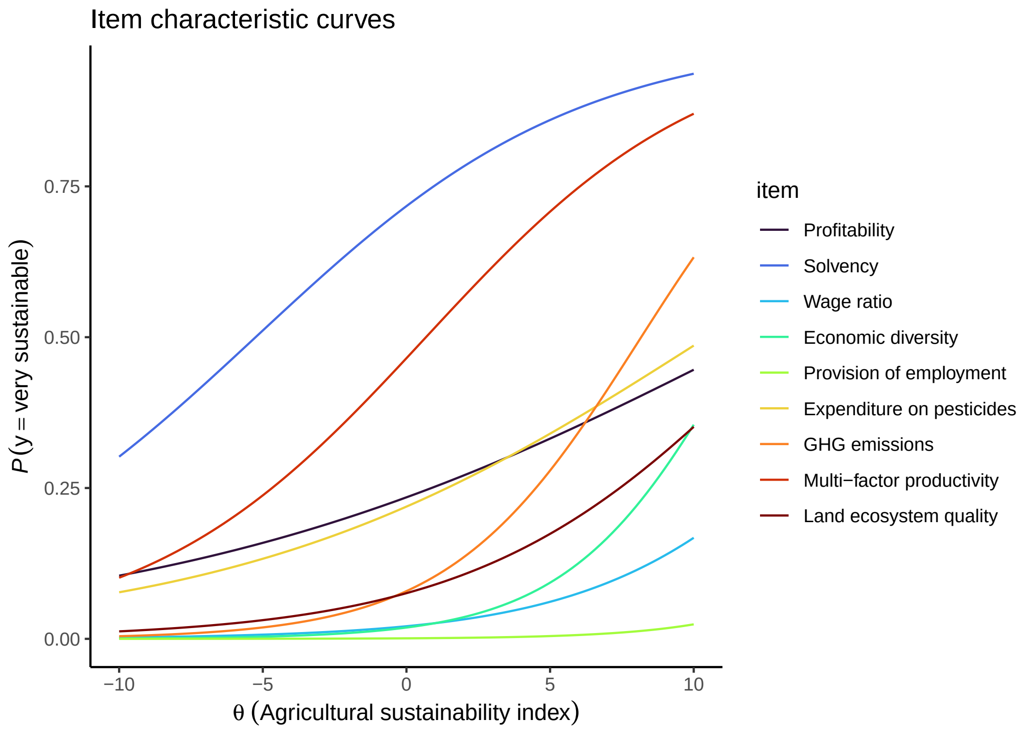

Fig. 4 Item characteristic curves for the "very sustainable" category in all nine sustainability items.

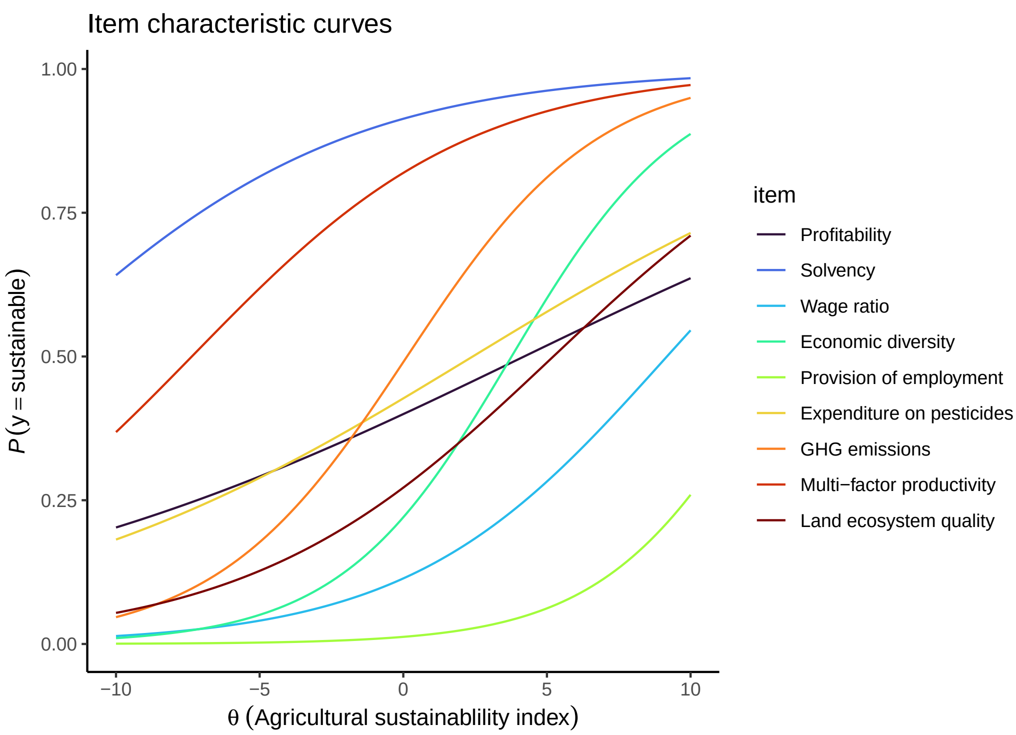

Fig. 5 Item characteristic curves for the "sustainable" category in all nine sustainability items.

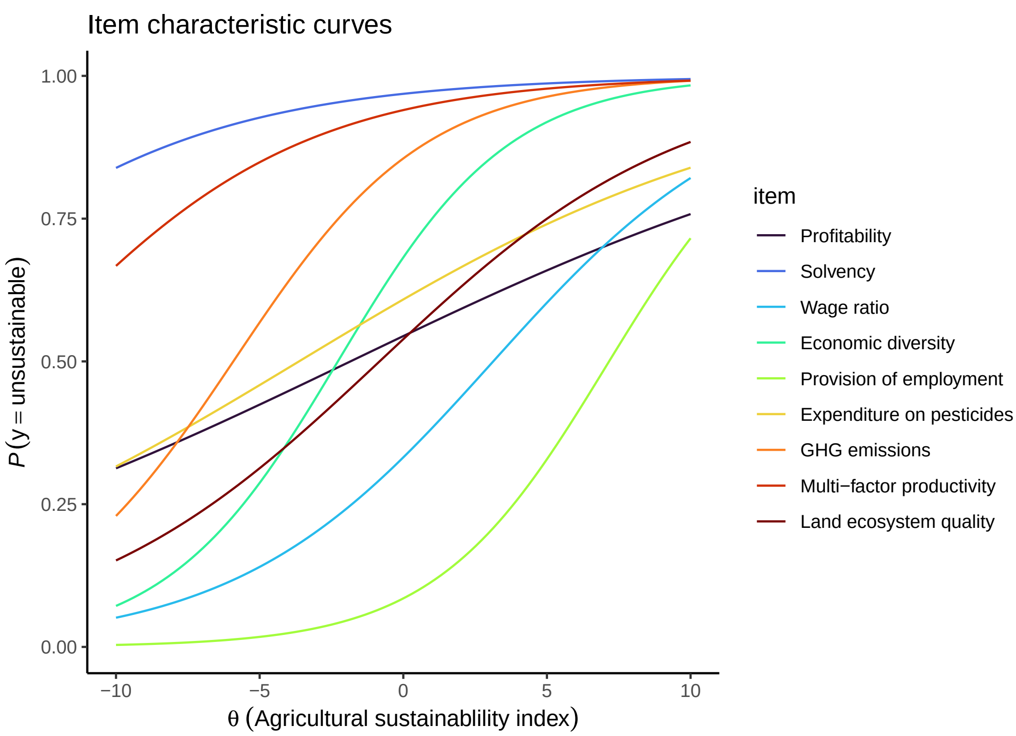

Fig. 6 Item characteristic curves for the "unsustainable" category in all nine sustainability items.

5.3 Predicted probabilities by farm type, size, and region

Farm type	Frequency	P(Y = very unsustainable)	P(Y = unsustainable)	P(Y = sustainable)	P(Y = very sustainable)
Mixed	1621	0.2373	0.2725	0.3335	0.1568
		(0.0661)	(0.0545)	(0.0421)	(0.0430)
Granivores	884	0.3523	0.2937	0.2554	0.0986
		(0.0654)	(0.0550)	(0.0425)	(0.0444)
Other grazing livestock	810	0.3180	0.2922	0.2776	0.1122
		(0.0640)	(0.0551)	(0.0439)	(0.0476)
Milk	2341	0.3396	0.2935	0.2635	0.1034
		(0.0649)	(0.0552)	(0.0429)	(0.0456)
Other permanent crops	190	0.2437	0.2748	0.3290	0.1525
		(0.0603)	(0.0507)	(0.0507)	(0.0555)
Wine	498	0.3682	0.2932	0.2455	0.0931
		(0.0661)	(0.0545)	(0.0421)	(0.0430)
Horticulture	450	0.3014	0.2900	0.2888	0.1198
		(0.0634)	(0.0548)	(0.0452)	(0.0493)
Fieldcrops	2134	0.2231	0.2661	0.3434	0.1673
		(0.0579)	(0.0483)	(0.0530)	(0.0576)

Table 7: Posterior means of the predicted probabilities for each sustainability category with standard errors in parentheses, by farm type.

Farm economic size class	Frequency	P(Y = very unsustainable)	P(Y = unsustainable)	P(Y = sustainable)	P(Y = very sustainable)
25,000-50,000	820	0.2979	0.3003	0.2879	0.1140
		(0.0711)	(0.0536)	(0.0498)	(0.0510)
50,000-100,000	1631	0.3073	0.3015	0.2815	0.1097
		(0.0718)	(0.0536)	(0.0495)	(0.0498)
100,000-250,000	3133	0.2884	0.2987	0.2945	0.1183
		(0.0700)	(0.0534)	(0.0501)	(0.0520)
250,000-500,000	2040	0.2732	0.2955	0.3053	0.1261
		(0.0685)	(0.0531)	(0.0509)	(0.0538)
500,000-750,000	579	0.2491	0.2885	0.3226	0.1399
		(0.0661)	(0.0523)	(0.0524)	(0.0570)
750,000-1,000,000	197	0.2105	0.2717	0.3504	0.1674
		(0.0616)	(0.0503)	(0.0555)	(0.0627)
1,000,000-1,500,000	149	0.1875	0.2581	0.3664	0.1880
		(0.0583)	(0.0487)	(0.0579)	(0.0661)
1,500,000-3,000,000	230	0.1434	0.2230	0.3922	0.2415
		(0.0510)	(0.0443)	(0.0628)	(0.0733)
>3,000,000	149	0.1272	0.2067	0.3981	0.2680
		(0.0477)	(0.0427)	(0.0644)	(0.0760)

Table 8: Posterior means of the predicted probabilities for each sustainability category with standard errors in parentheses, by farm size.

Region	Frequency	Est	Error	Q10	Q90
Stuttgart	550	0.1207	0.0508	0.0582	0.1879
Karlsruhe	174	0.1278	0.0521	0.0625	0.1976
Freiburg	97	0.1263	0.0523	0.0613	0.1973
Tübingen	379	0.1246	0.0515	0.0615	0.1934
Oberbayern	285	0.1224	0.0512	0.0589	0.1909
Niederbayern	290	0.1139	0.0492	0.0536	0.1798
Oberpfalz	192	0.1274	0.0521	0.0618	0.1967
Oberfranken	181	0.1167	0.0498	0.0555	0.1842
Mittelfranken	252	0.1268	0.052	0.0616	0.1958
Unterfranken	231	0.134	0.0535	0.0668	0.206
Schwaben	278	0.1085	0.0478	0.0496	0.1718
Brandenburg	286	0.2043	0.0642	0.1198	0.2874
Hamburg	99	0.0883	0.0431	0.0379	0.1465
Darmstadt	181	0.1273	0.0522	0.0626	0.1965
Gießen	113	0.1177	0.0502	0.0558	0.1852
Kassel	274	0.1095	0.0482	0.0507	0.1748
Mecklenburg-Vorpommern	272	0.1862	0.0621	0.1062	0.2662
Braunschweig	126	0.1521	0.0569	0.0799	0.2277
Hannover	214	0.1313	0.053	0.0653	0.2025
Lüneburg	474	0.1307	0.0528	0.0646	0.2013
Weser-Ems	519	0.1017	0.0463	0.0458	0.163
Düsseldorf	204	0.1183	0.0503	0.0563	0.1859
Köln	131	0.1223	0.0513	0.0596	0.1904
Münster	344	0.0955	0.0447	0.0421	0.1554
Detmold	225	0.1199	0.0506	0.0574	0.1878
Arnsberg	123	0.1163	0.0497	0.0543	0.1826
Koblenz	212	0.1158	0.0497	0.0538	0.1825
Trier	227	0.1004	0.046	0.0447	0.1624
Rheinhessen-Pfalz	463	0.1151	0.0494	0.0548	0.1813
Saarland	90	0.118	0.0506	0.0564	0.1849
Dresden	128	0.1696	0.0597	0.0913	0.2474
Chemnitz	123	0.1639	0.0588	0.0882	0.2431
Leipzig	67	0.197	0.0638	0.1125	0.2804
Sachsen-Anhalt	270	0.1942	0.0631	0.1114	0.2757
Schleswig-Holstein	565	0.1236	0.0513	0.0599	0.1923
Thüringen	279	0.1984	0.0635	0.1155	0.2792

Table 9: Regional averages of the predicted probability for a random farm achieving the "very sustainable" category, with standard errors and scores for top and bottom 10% of the sample.

References

Barnes, A., & Thomson, S. (2014). Measuring progress towards sustainable intensification: How far can secondary data go? *Ecological indicators*, *36*, 213–220, https://doi.org/10.1016/j.ecolind.2013.07.001

Coderoni, S., Bonati, G., D'Angelo, L., Longhitano, D., Mambella, M., Papaleo, A., Vanino, S. (2013). Using FADN data to estimate agricultural greenhouse gases emissions at farm level. *Pacioli*, *20*, 13–054, Retrieved from https://hdl.handle.net/11566/250330

Coderoni, S., & Esposti, R. (2018). CAP payments and agricultural GHG emissions in Italy. A farm-level assessment. *Science of the Total Environment*, *627*, 427–437, https://doi.org/10.1016/j.scitotenv.2018.01.197

Dabkienė, V. (2016). The scope of farms sustainability tools based on FADN Data. *Scientific Papers Series Management, Economic Engineering in Agriculture and Rural Development*, *16*(1), 121–128, (ISSN 2284-7995)

Dillon, E.J., Hennessy, T., Buckley, C., Donnellan, T., Hanrahan, K., Moran, B., Ryan, M. (2015). Measuring progress in agricultural sustainability to support policy-making. *International Journal of Agricultural Sustainability*, *14*(1), 31–44, https://doi.org/10.1080/14735903.2015.1012413

Dillon, E.J., Hennessy, T., Hynes, S. (2010). Assessing the sustainability of Irish agriculture. *International Journal of Agricultural Sustainability*, *8*(3), 131–147, https://doi.org/10.3763/ijas.2009.0044

Ehrmann, M. (2010). Assessing ecological and economic impacts of policy scenarios on farm level. *50th Annual Conference, Braunschweig, Germany, September 29-October 1, 2010.*

European Commission (2000). *Community committee for the Farm Accountancy Data Network (FADN): Definitions and variables used in FADN.* European Commission Directorate-General Agriculture, Brussels.

Eurostat (2017). *Glossary: Carbon dioxide equivalent.* https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Carbon_dioxide_equivalent. (Accessed 2023-10-13)

Eurostat (2022). *Gross value added of the agricultural industry - basic and producer prices.* https://data.europa.eu/data/datasets/bwzxcrbwhsymehubjmz6mw?locale=en. (Accessed 2023-10-13)

FAO (2018). *Guidelines for the measurement of productivity and efficiency in agriculture.* Food and Agriculture Organization of the United Nations, Rome.

Gerrard, C.L., Padel, S., Moakes, S. (2012). The use of farm business survey data to compare the environmental performance of organic and conventional farms. *International Journal of Agricultural Management*, *2*(1), 5–16, https://doi.org/10.22004/ag.econ.159244

Glauben, T., Tietje, H., Weiss, C.R. (2005). *Analysing family farm succession: a probit and a competing risk approach* (Tech. Rep.). Department of Food Economics and Consumption Studies, University of Kiel. (No. 724-2016-49080)

Gómez-Limón, J.A., Arriaza, M., Guerrero-Baena, M.D. (2020). Building a composite indicator to measure environmental sustainability using alternative weighting methods. *Sustainability*, *12*(11), 4398, https://doi.org/10.3390/su12114398

IPCC (2006). *IPCC guidelines for national greenhouse gas inventories.* Prepared by the National Greenhouse Gas Inventories Programme. Retrieved from http://www.ipcc-nggip.iges.or.jp/public/2006gl/index.html

Kelly, E., Latruffe, L., Desjeux, Y., Ryan, M., Uthes, S., Diazabakana, A., … Finn, J. (2018). Sustainability indicators for improved assessment of the effects of agricultural policy across the EU: Is FADN the answer? *Ecological Indicators*, *89*, 903–911, https://doi.org/10.1016/j.ecolind.2017.12.053

Lynch, J., Hennessy, T., Buckley, C., Dillon, E., Donnellan, T., Hanrahan, K., … Ryan, M. (2016). Teagasc National Farm Survey 2015 sustainability report. *Athenry, Co. Galway: Teagasc*, , (ISBN 978-1-84170-631-3)

Povellato, A., Bodini, A., Longhitano, D., Scardera, A. (2012). Assessing Farm Sustainability. An Application with the Italian FADN Sample. *2012 First Congress, June 4-5, 2012, Trento, Italy.* Retrieved from https://ageconsearch.umn.edu/record/124381

Reidsma, P., Tekelenburg, T., Van den Berg, M., Alkemade, R. (2006). Impacts of land-use change on biodiversity: An assessment of agricultural biodiversity in the European Union. *Agriculture, Ecosystems & Environment*, *114*(1), 86–102, https://doi.org/10.1016/j.agee.2005.11.026

Ryan, M., Hennessy, T., Buckley, C., Dillon, E.J., Donnellan, T., Hanrahan, K., Moran, B. (2016). Developing farm-level sustainability indicators for Ireland using the Teagasc National Farm Survey. *Irish Journal of Agricultural and Food Research*, *55*(2), 112–125, https://doi.org/10.1515/ijafr-2016-0011

Umweltbundesamt, S.B., Bundesanstalt für Geowissenschaften und Rohstoffe (2007). Umweltbundesamt.

van der Meulen, H., Dolman, M., Jager, J., Venema, G. (2014). The impact of farm size on sustainability of Dutch dairy farms. *International Journal of Agricultural Management*, *3*(2), 119–123, https://doi.org/10.5836/ijam/2014-02-07

Van Passel, S., & Meul, M. (2012). Multilevel and multi-user sustainability assessment of farming systems. *Environmental Impact Assessment Review*, *32*(1), 170–180, https://doi.org/10.1016/j.eiar.2011.08.005

Vitunskiene, V., & Dabkiene, V. (2016). Framework for assessing the farm relative sustainability: a Lithuanian case study. *Agricultural Economics*, *62*(3), 134–148, https://doi.org/10.17221/125/2015-AGRICECON

Westbury, D., Park, J., Mauchline, A., Crane, R., Mortimer, S. (2011). Assessing the environmental performance of English arable and livestock holdings using data from the Farm Accountancy Data Network (FADN). *Journal of Environmental Management*, *92*(3), 902–909, https://doi.org/10.1016/j.jenvman.2010.10.051