I N S T I T U T D E S T A T I S T I Q U E
B I O S T A T I S T I Q U E E T
S C I E N C E S A C T U A R I E L L E S
(I S B A)

UNIVERSITÉ CATHOLIQUE DE LOUVAIN

D I S C U S S I O N
P A P E R

2014/03

Evaluation of the EU Proposed Farm Income Stabilisation
Tool by Skew Normal Linear Mixed Models

PIGEON, M., HENRY DE FRAHAN, B. AND M. DENUIT

Evaluation of the EU Proposed Farm Income Stabilisation Tool by Skew Normal Linear Mixed Models

Mathieu Pigeon · Bruno Henry de Frahan ·
Michel Denuit

the date of receipt and acceptance should be inserted later

Abstract The European Commission has introduced risk management tools in the rural development pillar 2 of the Common Agricultural Policy. One of them consists in providing co-financing support to mutual funds compensating farmers who experience a severe drop in their income. This paper analyses this income stabilisation tool for a region in Belgium by means of a Skew Normal linear mixed model. Relying on the Farm Accountancy Data Network (FADN), this analysis focuses on estimating the probability that such a fund would need to intervene and, in that case, the expected amount of each farm income compensation. The predictive distribution of future incomes given past revenues trajectory is derived and used for evaluation purposes. Particular attention is paid to additional requirements that could be imposed to the income stabilisation tool.

Keywords agricultural risk management · income stabilisation · Skew Normal distribution · mixed effects models · Common Agricultural Policy

Mathieu Pigeon
Institute of Statistics, Biostatistics and Actuarial Science (ISBA), Université catholique de Louvain, Louvain-la-Neuve, Belgium, E-mail: Mathieu.Pigeon@uclouvain.be

Bruno Henry de Frahan
Earth and Life Institute (ELI), Université catholique de Louvain, Louvain-la-Neuve, Belgium, E-mail: Bruno.Henrydefrahan@uclouvain.be

Michel Denuit
Institute of Statistics, Biostatistics and Actuarial Science (ISBA), Université catholique de Louvain, Louvain-la-Neuve, Belgium, E-mail: Michel.Denuit@uclouvain.be

1 Introduction and motivation

Among the European Union (EU)’s legislative decisions for the Common Agricultural Policy (CAP) after 2013, several new risk management tools are introduced in the rural development pillar 2 of the CAP as a response to greater price and yield variability [11]. One of them presented as an income stabilisation tool consists in providing co-financing support to mutual funds that would offer compensation to farmers who experience a severe drop in their farm income. If farm income drops by at least 30% below a 3-year average figure then the mutual fund would reimburse for not more than 70% of the farm losses. It is proposed that the rural development pillar 2 would contribute to these mutual funds at the rate of 0.65 euro for every 1 euro paid into the mutual fund by farmers and cover up to 65% of eligible costs. The current proposal leaves member states to define the rules for the constitution and management of such mutual funds. It is therefore too early to know whether participation to these mutual funds would be made on a voluntary or compulsory basis and delineated according to some criteria, such as farm type and geographic location.

This proposed whole-farm income stabilisation tool (IST) has several advantages. First, it is meant to provide an overall coverage to all farm risks. Farmers generally prefer to insure their whole-farm income than separate components of their income such as the yield or the revenue of a particular activity, as they see it closer to optimising their own economic well being [16]. Combining all farm’s insurable risks into a single contract provides a more efficient coverage than sorting them out into separate contracts when farm’s risks do not perfectly covariate [15, 8]. Indemnities from as well as financial contributions to the mutual fund are most likely to be lower compared to other risk management tools that would cover specific farm risks that are not perfectly correlated.

Second, it is meant to cover severe if not catastrophic farm income risks that are most likely to be systematic within the same region where the mutual fund acts. For such risks there are reasons for a public support [16]. The systematic character of the risks covered justifies some form of public intervention, as insurers may be reluctant to enter such income insurance market. The severity of the risks also justifies some public intervention when the marginal utility of the income assistance is particularly high in case of low incomes and the political pressure on public authorities to intervene is soaring. Targeting public support on mutual funds that cover systematic and severe risks are therefore less likely to generate those crowding-out effects that are expected when subsidised income stabilisation programmes interfere with private risk management strategies or insurance markets. Third, the proposed IST is purposely designed to be classified as green box compatible with the Uruguay Round Agreement on Agriculture (URAA) of the World Trade Organization [23] since it supports mutual funds that release compensation payments to participants only in case of an income loss exceeding 30% of a 3-year average gross or net income in the preceding 3-year period. Those compensation payments are also not exceeding 70% of the income loss. As such, the proposed IST shares several features of the Canadian AgriStability income stabilisation programme except that this programme does not use mutual funds to channel its support. Instead, the Canadian federal government directly covers its share of the individual farm income losses. There is a participation fee to the

programme but at a low rate of CAD 3.80/CAD 1,000 of the individual reference income.

Finally, the proposed IST is well targeted to the individual income situations compared to other income related payments linked to aggregate indicator or index such as revenue, price or yield at the regional or national level, since its compensation payments are triggered when participants actually experience low income. This well targeting, however, brings several implementation problems that are more detailed in OECD [18].

The main problem of the whole-farm approach of the IST stems from the requirement to measure and collect accurately the farm income in a manner that avoids moral hazard and is acceptable for the management of the mutual fund. The Canadian AgriStability programme collects farm income information through the income tax system with supplementary information from farmers and the Farm Financial Survey [1]. Such reliable sources of information on the farm income situation are hardly available in a systematic manner at the individual farm level for most EU member states. Associated with the problem of measuring farm income is the delay between the incidence of low farm income and the actual compensation. Late compensation can be particularly detrimental to the survival of the farm for severe income losses that require a prompt response.

The proposed whole-farm IST has also the unavoidable problem of moral hazard due to changing incentives when individual outcomes are used to trigger compensation payments. This may eventually lead to crowding out risk management strategies at the farm level and give incentives to farmers to take on more risks that can yield higher returns [18]. Moral hazard is likely to be more prominent for this proposed IST where variability in farm incomes results from many various management decisions, not only from accidental and unintentional losses. One solution to lessen this typical moral hazard problem consists in making participation costs dependable on behaviour thanks to a bonus-malus system of counter-incentives. Other solutions consist in delineating mutual funds such that peer pressure can counter incentives towards more risky activities or adding an additional trigger for compensation based on regional references. Adverse selection can also be a problem when those participating to the mutual fund face a higher risk than those who do not, so that the expected compensation payments calculated from aggregate data underestimate their real cost.

It is, however, unclear to what extent these income stabilising mutual funds may be financially sustainable since, in contrast to the Canadian AgriStability programme, they solely rely on the financial contributions of their participants that are topped-up by financial contributions from the rural development pillar 2 at a fixed proportion. Following a preliminary evaluation of the European Commission [10], it is generally reported that such an income stabilisation scheme would be costly. For instance, in year 2006, 25% of EU-15 farms suffered an income loss greater than 30% and the total compensation for these farms would have amounted to approximately 9.9 billion euros if 70% of the loss would be compensated.

The main objective of this paper is to study the behaviour of the IST mechanism in various conditions. More precisely, we focus on estimating the probability of the activation of compensation payments for a typical regional mutual fund and, in that event, the expected amount of each farm's compensation and the total expected amount of compensation. To generate these three indicators and associated confidence intervals, we develop a Skew Normal linear mixed model that

decomposes farm incomes into a dynamic farm-specific effect that induces serial correlation and some explanatory control variables accounting for observed heterogeneity. It is a well-documented fact that the kind of evaluation conducted in the present paper requires farm-level observations since aggregated data at regional or national level may be misleading [6]. The model is fitted to the observations for Walloon farms included in the Farm Accountancy Data Network (FADN) from 1997 to 2007. The predictive distribution of future incomes given the observed trajectory of past revenues is derived and used for risk evaluation purposes. Simulations are also performed for future incomes and this stochastic information is used to generate the three indicators and evaluate the new whole-farm income stabilisation tool proposed by the European Commission.

We emphasize that this analysis of a typical income stabilisation mutual fund does not aim at estimating the full operational cost of it, neither pricing the financial contributions to cover its cost. The full operational cost depends closely on numerous management specificities of the mutual fund that are not yet decided. Estimates that we generate are, however, the necessary preliminary step in that direction. Some suggestions on pricing are outlined at the end of the paper.

The remainder of this paper is organized as follows. After this introductory section, the next section of the paper specifies the Skew Normal linear mixed model used in the present paper to evaluate the EU proposed farm income stabilisation mechanism. The third section presents the farm dataset. The following sections use the model and this dataset to obtain estimations, validations and predictions. The last section discusses the results and ends with some policy and methodological recommendations. Proofs of the main results are gathered in an appendix at the end of this paper.

2 Skew Normal linear mixed model

2.1 Model specification

The model considered in this paper for annual farm incomes belongs to the class of general linear mixed models that are increasingly used in a wide variety of disciplines including econometrics, actuarial sciences and social sciences (e.g. [2] and [12]). Such models can be used in settings where repeated measurements are made on the same statistical units and when partial information is missing in the dataset resulting in unexplained heterogeneity.

Let Y_{it} be the annual income for farm i , $i = 1, \dots, n$ during calendar year t , $t = 1, \dots, T_i$. The trajectory $\mathbf{Y}_i = [Y_{i1} \dots Y_{iT_i}]'$ of annual revenues can be explained by a vector of p explanatory variables $\mathbf{x}_{it} = [x_{it1} \dots x_{itp}]$ with

$$\mathbf{X}_i = \begin{bmatrix} \mathbf{x}_{i1} \\ \vdots \\ \mathbf{x}_{iT_i} \end{bmatrix} = \begin{bmatrix} x_{i11} & \dots & x_{i1p} \\ \vdots & \ddots & \vdots \\ x_{iT_i1} & \dots & x_{iT_ip} \end{bmatrix}.$$

However, this information only partially explains the dynamics of the income amounts Y_{it} so that there remains heterogeneity in the data. The effect of the unknown characteristics relating to farm i during period t is represented by a random

variable Θ_{it} . Let $\Theta_i = [\Theta_{i1} \dots \Theta_{iT_i}]'$ be the vector of random effects for farm i . The sequences of income amounts $\mathbf{Y}_1, \dots, \mathbf{Y}_n$ are assumed to be independent.

The model used in the present paper is based on the following three assumptions:

- (H1) Given the random vector Θ_i , the random vector $\mathbf{Y}_i$ follows a Multivariate Skew Normal distribution with location parameter $(T_i \times 1) \mathbf{X}_i\beta + \mathbf{Z}\Theta_i$, where $\mathbf{Z}$ is the design matrix for Θ_i , scale parameter $(T_i \times T_i) \Sigma$ and shape parameter $(T_i \times T_i) \Delta$. Dropping the farm-specific subscript for simplicity, the conditional probability density function of the vector $\mathbf{Y}$ of annual farm incomes given individual risk factors Θ is

$$f_{\mathbf{Y}}(\mathbf{y}|\Theta = \theta) = 2^T \phi_T(\mathbf{y}; \mathbf{X}\beta + \mathbf{Z}\theta, \Omega) \times \Phi_T\left(\Delta' \Omega^{-1}(\mathbf{y} - \mathbf{X}\beta - \mathbf{Z}\theta); \mathbf{0}, \mathbf{I}_T - \Delta' \Omega^{-1} \Delta\right),$$

where $\Omega = \Sigma + \Delta\Delta'$ and $\phi_T(\cdot; \mu, \mathbf{S})$ and $\Phi_T(\cdot; \mu, \mathbf{S})$ respectively denote the probability density function and the distribution function of the T -dimensional Normal distribution with mean vector μ and variance-covariance matrix $\mathbf{S}$. Notice that the scale parameter Σ is not the usual variance-covariance matrix of $\mathbf{Y}$, contrarily to the Multivariate Normal distribution. A multivariate Skew Normal distribution is often specified by $\Sigma^{1/2}$ in place of Σ because of the plurality of the square roots of the matrix Σ . Henceforth, $\Sigma^{1/2}$ denotes the symmetric square root of Σ .

- (H2) The random vector Θ follows the Multivariate Normal distribution with mean vector $(T \times 1) \mathbf{0}$ and variance-covariance matrix $(T \times T) \Psi$, i.e.,

$$f_{\Theta}(\theta) = \phi_T(\theta; \mathbf{0}, \Psi).$$

- (H3) Components of Ψ , Σ and Δ are selected in a way that the distribution of Θ depends only on T . Let $T_{\max} = \max_i(T_i)$ and the distribution of the random vector $(\Theta_1, \dots, \Theta_{T_{\max}})$ is stationary. We select a diagonal matrix with constant value for the shape parameter ($\Delta_1 = \Delta_2 = \dots$) and elements of matrices Ψ and Σ are given by (i^{th} row, j^{th} column)

$$\Psi_{ij} = \begin{cases} \Psi, & i = j \\ \Psi\rho^{|j-i|}, & i \neq j \end{cases}, \quad \Sigma_{ij} = \begin{cases} \sigma^2, & i = j \\ 0, & i \neq j. \end{cases}$$

The Skew Normal distribution is known to be particularly suitable in finance and insurance applications, see, e.g., Bolance et al [4], Eling [9] or Pigeon et al [19]. The probability density function of the univariate Skew Normal distribution is displayed in Figure 1 for different values of its shape parameter Δ . Contour plots of the bivariate Skew Normal probability density function are represented in Figure 2 for different values of its shape parameters Δ_1 and Δ_2 . It is clear from these two figures that the Skew Normal distributions are able to accomodate a variety of marginal behaviours and dependence structures. The Skew Normal linear mixed model has been proposed by Arellano-Valle et al [3] and appears to be relatively new compared to more conventional linear mixed models. For this class of models, maximum likelihood techniques are presented in Lin [14].

In this model, the correlation structure of $\mathbf{Y}_i$ is characterized by the latent unobservable process Θ_i . The components of the random vector Θ_i represent

Fig. 1 Univariate Skew Normal probability density functions $f(x) = 2\phi_1(x)\Phi_1(\Delta x)$ for different values of the shape parameter delta

effects that vary across individuals and the vector of fixed effects β represents effects that are the same across all individuals [12]. In our analysis, this class of models allow us to control for some characteristics of the farm and decompose the total risk into farm-specific risk (e.g., production risk) and random risk.

2.2 Properties and Inference

Now that we know the distribution of $\mathbf{Y}$ (dropping again the farm-specific subscript for readability), let us derive some elementary properties of it that are useful for risk evaluation. Precisely, the next result gives explicit expressions for the first moments of $\mathbf{Y}$ and establishes the closure of the Skew Normal family by marginalisation and by linear transformation. The proofs are given in appendix.

Proposition 1 *Let $\mathbf{Y}$ be a random vector with Multivariate Skew Normal distribution with location parameter $\mathbf{X}\beta$, scale parameter $\Sigma^* = \Omega + \mathbf{Z}\Psi\mathbf{Z}' - \Delta\Delta'$ and shape parameter Δ .*

(i) *The moment generating function is*

$$M_{\mathbf{Y}}(\mathbf{t}) = 2^T \exp\left(\mathbf{t}'\mathbf{X}\beta + 0.5\mathbf{t}'\Omega^*\mathbf{t}\right)\Phi_T(\Delta'\mathbf{t}),$$

Fig. 2 Contour plots of the bivariate Skew Normal probability density function for different values of the shape parameters δ_1 and δ_2

with $\Omega^* = \Sigma^* + \Delta\Delta'$. Straightforward calculations give

$$E[\mathbf{Y}] = \mathbf{X}\beta + \sqrt{\frac{2}{\pi}} \text{diag}(\Delta)$$

$$\text{Var}[\mathbf{Y}] = \Sigma^* + \left(1 - \frac{2}{\pi}\right) \Delta\Delta'.$$

We define

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{bmatrix} \quad \mathbf{X}\beta = \begin{bmatrix} \mathbf{X}_1\beta_1 \\ \mathbf{X}_2\beta_2 \end{bmatrix} \quad \Sigma^* = \begin{bmatrix} \Sigma_{11}^* & \Sigma_{12}^* \\ \Sigma_{21}^* & \Sigma_{22}^* \end{bmatrix} \quad \Delta = \begin{bmatrix} \Delta_1 & \mathbf{0} \\ \mathbf{0} & \Delta_2 \end{bmatrix}$$

where $\mathbf{Y}_1$ and $\mathbf{X}_1\beta_1$ are vectors with dimensions $(k_1 \times 1)$; and Σ_{11}^* and Δ_1 are matrices with dimensions $(k_1 \times k_1)$. By using

$$\mathbf{t}_{k_1} = \left[\underbrace{1 \quad \dots \quad 1}_{k_1} \quad 0 \quad \dots \quad 0 \right]'$$

in $M_{\mathbf{Y}}(\mathbf{t})$, we obtain that $\mathbf{Y}_1$ follows a Multivariate Skew Normal distribution with parameters $\mathbf{X}_1\beta_1$, Σ_{11}^* and Δ_1 .

- (ii) Define a vector $(T \times 1)$ $\mathbf{A}$ and a matrix $(T \times T)$ $\mathbf{B}$. Then, $\mathbf{A} + \mathbf{B}\mathbf{Y}$ follows the Multivariate Skew Normal distribution with location parameter $\mathbf{A} + \mathbf{B}\mathbf{X}\boldsymbol{\beta}$, scale parameter $\mathbf{B}\boldsymbol{\Sigma}^*\mathbf{B}'$ and shape parameter $\mathbf{B}\boldsymbol{\Delta}$.

The probability density function of the random vector $\mathbf{Y}$ of farm revenues is derived in Proposition 2. The proof is given in appendix.

Proposition 2 *In the model H1-H2-H3, the probability density function of $\mathbf{Y}$ is given by*

$$f_{\mathbf{Y}}(\mathbf{y}) = 2^T \phi_T(\mathbf{y}; \mathbf{X}\boldsymbol{\beta}, \boldsymbol{\Omega} + \mathbf{Z}\boldsymbol{\Psi}\mathbf{Z}') \\ \Phi_T\left(\boldsymbol{\Delta}'(\boldsymbol{\Omega} + \mathbf{Z}\boldsymbol{\Psi}\mathbf{Z}')^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}); \mathbf{0}, \mathbf{I}_T - \boldsymbol{\Delta}'(\boldsymbol{\Omega} - \mathbf{Z}\boldsymbol{\Psi}\mathbf{Z}')^{-1}\boldsymbol{\Delta}\right).$$

Hence, $\mathbf{Y}$ follows the Multivariate Skew Normal distribution with location parameter $\mathbf{X}\boldsymbol{\beta}$, scale parameter $\boldsymbol{\Sigma}^* = \boldsymbol{\Omega} + \mathbf{Z}\boldsymbol{\Psi}\mathbf{Z}' - \boldsymbol{\Delta}\boldsymbol{\Delta}'$ and shape parameter $\boldsymbol{\Delta}$.

The closed-form expression available for $f_{\mathbf{Y}}$ is particularly useful to estimate the model parameters by maximum likelihood.

3 Presentation of the data

Let us now describe the dataset at our disposal. Because the management of the rural development pillar 2 is often decentralized at the regional level, this paper analyses the new farm income stabilisation tool for one particular European region, the Region of Wallonia in Belgium. We use a sample of observations for Walloon farms included in the Farm Accountancy Data Network (FADN) from 1997 to 2007. The types of farms (TF) that are the most frequent in the Walloon region include specialist cattle-rearing and fattening farms (TF42) for 25%, field crop farms (TF1) with mainly cereals, sugar beets and potatoes for 22%, specialist dairying farms (TF41) for 15%, field crops-grazing livestock combined farms (TF81) for 13% and cattle-dairying, rearing and fattening combined farms (TF43) for 11% in number of farms [7]. Farms in this region as well as generally in Northwest Europe are more subject to price risks than production risks with the exception of the open-air horticultural farms (TF2).

The class of farming is known to have an influence on income variability [13]. This is the reason why we split the population of farms in three categories: crop, dairy and cattle farms, as if we establish three regional mutual funds, one for each farm category in the Walloon region. The crop farm category includes TF1310, 1410, 1420 and 1443; the dairy farm category TF4110, 4120 and 4310; and the cattle farm category TF4210, 4220 and 4320.

The farm's net value added measures the amount available for remuneration of the fixed production factors (labour, land and capital). The European Commission, in the context of the income stabilisation tool proposal, has adopted it because it is the most comparable indicator of farm incomes between member states. Whether production factors are external or family factors does not modify the net value added. Note that net value added is here measured on a per farm basis and that all incomes are expressed in current euros, without any adjustment for inflation.

Figure 3 displays box-plots for each calendar year and each type of farm summarizing the corresponding income distributions. Median farm incomes are displayed for each calendar year, surrounded by the first and third quartiles defining the central box. This indicates income range excluding the 25% of farms with the lowest income and 25% with the highest income. Figure 3 reveals an increasing trend as well as a marked cyclic behaviour that may be due to natural events affecting production or movements in prices. These cycles will be accounted for by means of sine and cosine functions included in the covariates.

Fig. 3 Box plots for farm's net value added for crop farms (left), dairy farms (middle) and cattle farms (right)
Source: FADN

The observable characteristics of the farm are also important and are reflected in the information incorporated in the model. Specifically, explanatory variables include:

1. Type of farm (TF): which is a categorical variable representing the economic-technological orientation of the farm.
2. Total utilized agricultural area (UAA) of holding (TotUAA) and its square: it consists of land in owner occupation, rented land and land in sharecropping. It includes agricultural land temporarily not under cultivation for agricultural reasons or being withdrawn from production as part of agricultural policy measures. It is expressed in hectares (10,000 square meters).
3. Livestock units (LU) and its square: annual average number of equines, cattle, sheep, goats, pigs and poultry present on holding, converted into livestock units (LU). Animals that do not belong to the holder but are held under a production contract are taken into account according to their annual presence.
4. Number of dairy cows (Cows): the annual average number of female bovine animals, including female buffaloes, which have calved and are held principally for milk production for human consumption, is also included for dairy farms only.

Figure 4 shows descriptive statistics for explanatory variables and Figure 5 illustrates relations between farm incomes and explanatory variables for crop, dairy and cattle farms respectively.

4 Estimation results

All parameters of the model are estimated in a one-step procedure using the maximum likelihood method thanks to the explicit expression for the probability density function of the response derived in Proposition 2. Calendar time as well as sine and cosine functions are included in the covariates to account for the increasing trend and the cyclic behaviour which is five years in the final model as it appeared in Figure 3. Validation issues are discussed in subsection 5.3. Estimation results are presented in Table 1 for the three farm categories where non-significant effects have been removed from the model. The length of cycle for crop, dairy and cattle farms and the selection of the explanatory variables are performed by means of a stepwise procedure implemented with the help of statistical tests and the Bayesian information criterion (BIC). All results are consistent with the descriptive analysis presented in Figures 3, 4 and 5.

Fig. 4 Descriptive statistics for crop farms (top), dairy farms (middle) and cattle farms (bottom) for explanatory variables LU (left), TotUAA (middle) and Cows (right)
Source: FADN

Fig. 5 Relations between crop farm (top), dairy farm (middle) and cattle farm (bottom) net value added and explanatory variables

Source: FADN

Variables	Class		
	Crop (s.e.)	Dairy (s.e.)	Cattle (s.e.)
Intercept	−23,679 (3,036)	−8,776 (825)	4,308 (227)
$t = \text{Year} - 1997$	1,921 (211)	738 (73)	845 (138)
$\sin(2\pi t/5)$	8,641 (875)	1,962 (226)	3,407 (476)
$\cos(2\pi t/5)$	8,847 (779)	428 (213)	−1,975 (437)
LU	1,090 (15)	282 (15)	385 (4)
TotUAA	1,051 (21)	256 (7)	237 (2)
Cows	−	556 (16)	−
TF (1410)	29,751 (4,647)	−	−
TF (1420)	16,505 (2,788)	−	−
TF (1443)	6,960 (2,763)	−	−
TF (4220)	−	−	18,396 (207)
TF (4320)	−	−	7,605 (211)
Residual s.e. ($\hat{\sigma}_\epsilon$)	11,534	8,003	3,971
AR factor ($\hat{\rho}$)	0.21	0.48	0.37
Shape par. ($\hat{\Delta}$)	> 100	−45	> 100

Table 1 Estimation results for explanatory variables (non-significant effects have been removed from the model)

5 Evaluation of the proposed EU income stabilisation tool

5.1 Description of the mechanism

Among the EU's legislative decisions for the 2014-2020 CAP is a new risk management tool which has been put forward as a response to the current climate and food price variability. Under an income stabilization tool (IST), farmers would receive up to 70% of any income loss from a mutual fund if their income falls by at least 30%, below a reference value: (A) a 3-year average figure; (B) a 5-year average figure; or (C) a 3-year average based on the preceding 5-year period, excluding the highest entry and the lowest entry (WTO definition, also called olympic average).

These 30% and 70% thresholds are set to ensure that this new tool is classified as green box compatible. As explained in the introduction, this new tool is in line with modern risk management in agriculture focussing on whole farm income insurance programs. In that setting, a single policy is provided, which covers global risk of jointly produced farm crop and livestock enterprises. See, e.g., Turvey [22].

Turning the IST mechanism into formulas, the target revenue for farm i in calendar year t is

$$\begin{aligned}\bar{Y}_{it}^A &= \frac{Y_{i,t-1} + Y_{i,t-2} + Y_{i,t-3}}{3} \\ \bar{Y}_{it}^B &= \frac{Y_{i,t-1} + Y_{i,t-2} + Y_{i,t-3} + Y_{i,t-4} + Y_{i,t-5}}{5} \\ \bar{Y}_{it}^C &= \frac{Y_{i,t-1} + Y_{i,t-2} + Y_{i,t-3} + Y_{i,t-4} + Y_{i,t-5} - M_{it} - m_{it}}{3} \\ M_{it} &= \max\{Y_{i,t-1}, Y_{i,t-2}, Y_{i,t-3}, Y_{i,t-4}, Y_{i,t-5}\} \\ m_{it} &= \min\{Y_{i,t-1}, Y_{i,t-2}, Y_{i,t-3}, Y_{i,t-4}, Y_{i,t-5}\}\end{aligned}$$

and the corresponding IST benefit is

$$B_{it}^m = \begin{cases} 0.7(\bar{Y}_{it}^m - Y_{it}), & Y_{it} \leq 0.7 \times \bar{Y}_{it}^m \\ 0, & \text{otherwise} \end{cases}$$

for $m \in \{A, B, C\}$. That the target income $\bar{Y}_{it}$ is set equal to a 3-year or 5-year moving average has important consequences when incomes are subject to cycles. In the descending phase of the cycle, the IST is likely to indemnify farmers but the benefits decrease over time as the target income diminishes. The safety net provided by the IST thus delays shocks due to cycles but does not totally protect farmers against them. It is also important to point out that the total amount of benefits paid in execution of the IST is not stable at all over time but strongly depends on the farms position in the cycle.

On an individual level, this feature may nevertheless be regarded as desirable since inefficient farms with decreasing incomes are not artificially supported by the IST. Actually, the 3-year or 5-year moving average defining the IST target income can be considered as a form of experience rating, or bonus-malus system letting the sum insured depend on individual characteristics. The last section of this paper comes back to this issue.

5.2 Individual income trajectories

The Skew Normal linear mixed model can be used for predicting future income trajectories for each farm. To this end, we use the predictive distribution of Y_{T+1} given past observed revenues $\mathbf{Y} = (Y_1, \dots, Y_T)$ whose probability density function is given by

$$f_{Y_{T+1}|\mathbf{Y}}(y_{T+1}|\mathbf{y}) = \frac{f(\mathbf{Y}, Y_{T+1})((\mathbf{y}, y_{T+1}); \tilde{\mathbf{X}}\boldsymbol{\beta}, \tilde{\boldsymbol{\Sigma}}^*, \tilde{\boldsymbol{\Delta}})}{f_{\mathbf{Y}}(\mathbf{y}; \mathbf{X}\boldsymbol{\beta}, \boldsymbol{\Sigma}^*, \boldsymbol{\Delta})},$$

where $\tilde{\mathbf{X}}\boldsymbol{\beta}$ is a $(T+1 \times 1)$ vector and $\tilde{\boldsymbol{\Sigma}}^*$ and $\tilde{\boldsymbol{\Delta}}$ are $(T+1 \times T+1)$ matrices. Fitted values for calendar year t are obtained from $\mathbb{E}[Y_{it}|Y_{i,t-1}, Y_{i,t-2}, \dots]$ conditioning on all the past observations contained in the FADN databasis. Individual trajectories are easily simulated using the following proposition (the proof is given in Lin [14]).

Proposition 3 *Let $\mathbf{Y}$ be a T -dimensional random vector obeying the multivariate Skew Normal distribution with location parameter $\mathbf{X}\boldsymbol{\beta}$, scale parameter $\boldsymbol{\Sigma}^*$ and asymmetry parameter $\boldsymbol{\Delta}$. Then, $\mathbf{Y}$ is distributed as $\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\Delta}|\mathbf{Z}| + \mathbf{U}$, where the T -dimensional random vector $\mathbf{Z}$ is multivariate Normal with mean vector $\mathbf{0}$ and variance-covariance matrix $\mathbf{I}_T$ and where the T -dimensional random vector $\mathbf{U}$ is multivariate Normal with mean vector $\mathbf{0}$ and variance-covariance matrix $\boldsymbol{\Sigma}^*$.*

5.3 Validation of the model

We validate the model by verifying whether the observed benefits B_{it}^A from the IST fall inside prediction intervals around adjusted values provided by the model. We obtain these intervals by simulating IST benefits for the last three years of observations. We obtain 10,000 random draws from the *a posteriori*, or predictive distribution of the annual income from which we calculate corresponding benefits for mechanism $m = A$.

Table 2 displays the following information for each calendar year 2005-2007:

- the observed proportion of farms being indemnified, i.e., those such that $Y_{it} < 0.7 \times \bar{Y}_{it}^A$, and a 99% confidence interval around the fitted value of that proportion;
- the observed average amount of indemnity for those farms such that $Y_{it} < 0.7 \times \bar{Y}_{it}^A$ and a 99% confidence interval around the fitted value $\mathbb{E}[B_{it}^A | Y_{it} < 0.7 \times \bar{Y}_{it}^A]$; and
- the observed total amount of indemnity for all the farms and a 99% confidence interval around the fitted value $\sum_i \mathbb{E}[B_{it}^A | Y_{it} < 0.7 \times \bar{Y}_{it}^A]$.

For each year and farm category, the observed value falls within the confidence interval constructed around the adjusted value. We conclude that the model can provide reasonable predictions for the farm-specific net value added trajectories. One should note that confidence intervals are sometimes wide and this variability reflects the complexity of the underlying process. In the current analysis, using smaller confidence levels (e.g., 90% or 95%) do not appreciably reduce the size of these intervals.

5.4 Prediction

We use the estimated model and characteristics of farms for past years to simulate 1,000 replications of IST benefits for year 2008 using results in subsection 5.2. Table 3 displays for calendar year 2008 the prediction for the quantities displayed in Table 2. We limit ourselves to predictions for calendar year 2008 because predicting results for the following years 2009, 2010, etc. involves partly unobserved income trajectories and unknown farm characteristics. For 2008, we use the same covariates as those observed in 2007.

We can also study the impact of the choice of the reference value ($\bar{Y}_{it}^A$, $\bar{Y}_{it}^B$ or $\bar{Y}_{it}^C$ defined in subsection 5.1) and of the franchise percentage of loss needed to

Year	Information	Farm type		
		Crop (C.I.)	Dairy (C.I.)	Cattle (C.I.)
2005	Proportion	0.17 (0.00 – 0.26)	0.04 (0.02 – 0.10)	0.18 (0.03 – 0.19)
	Average	22,342 (6,992 – 38,292)	17,618 (14,356 – 27,414)	16,363 (12,204 – 31,408)
	Total	134,051 (0 – 196,944)	105,709 (70,243 – 342,687)	196,357 (47,415 – 221,504)
2006	Proportion	0.03 (0.00 – 0.16)	0.10 (0.02 – 0.10)	0.09 (0.09 – 0.27)
	Average	14,722 (3,142 – 46,724)	21,995 (13,927 – 29,768)	12,757 (12,526 – 24,393)
	Total	14,722 (0 – 106,776)	329,932 (58,794 – 348,179)	76,543 (68,147 – 333,388)
2007	Proportion	0.00 (0.00 – 0.07)	0.02 (0.01 – 0.09)	0.16 (0.04 – 0.21)
	Average	0.00 (0 – 43,404)	24,750 (13,221 – 35,865)	12,274 (11,560 – 24,198)
	Total	0.00 (0 – 43,404)	74,251 (30,920 – 282,984)	135,017 (49,048 – 236,111)

Table 2 Observed values and 99% confidence intervals obtained from the model for net value added in current euros following mechanism *A*

Year	Information	Farm type		
		Crop (C.I.)	Dairy (C.I.)	Cattle (C.I.)
2008	Proportion	0.04 (0.02 – 0.05)	0.02 (0.00 – 0.04)	0.05 (0.02 – 0.08)
	Average	24,973 (1,777 – 77,298)	22,637 (4,359 – 94,082)	13,463 (3,908 – 29,215)
	Total	84,159 (20,049 – 128,074)	136,005 (44,707 – 283,954)	100,219 (36,661 – 171,206)

Table 3 Mean predicted values and 95% confidence intervals obtained from the model for net value added in current euros following mechanism *A*

activate the mechanism, i.e., the value b entering the following formula:

$$B_{it}^m = \begin{cases} 0.7(\bar{Y}_{it}^m - Y_{it}), & Y_{it} \leq b \times \bar{Y}_{it}^m \\ 0, & \text{otherwise} \end{cases}$$

for $m \in \{A, B, C\}$ and $b \in (0.7, 1.0)$. Figure 6 displays results for calendar year 2008. We can see there the point estimates listed in Table 3 obtained when b increases from 70% to 100%. Increasing b relaxes the condition to be indemnified by the IST so that proportions of farms being indemnified clearly increase with b . As expected, indemnifying smaller losses reduce the average claim amount so that the second panels reveal a decreasing behaviour. Overall, increasing b results in higher total amounts paid in execution of the IST. Considering the total amounts paid, the value of b has an important impact on the aggregate IST cost so that this parameter should be selected with great care. The type m of averaging used

to determine the reference revenue has almost no effect for cattle farms. For dairy farms, the proportion of the farms being indemnified is reduced when replacing the 3-year average with the 5-year average or the Olympic average. The average claim amount increases, especially in the range of b between 70 and 85%. The effect on the total IST cost depends on the value of b : for b in excess of 80% to 85%, this cost decreases when the 3-year average is replaced with the 5-year average or the Olympic average. For crop farms, different effects depending on the value of b and no firm conclusions can be drawn.

Fig. 6 Prediction results for crop farm (top), dairy farm (middle) and cattle farm (bottom) net value added for the quantities displayed in Table 2

In our model, predictive farm income distribution is computed using the farm-specific risk factors comprised in the random vectors Θ_i that accumulate past experience. The *a posteriori* distribution of Y_{i,T_i+1} given observed past farm incomes allows for revising predictions of future farm incomes based on available observations. Figure 7 illustrates the impact of including past farm incomes for

crop farm incomes. Using the predictive income distribution or the unconditional one does not influence the proportion of farms being indemnified but the average claim cost. In the latter case, the inclusion of past observed incomes in the predictive distribution lowers the mean payment to farms benefitting from the IST, especially for b smaller than 85%. The resulting effect for the total IST cost in 2008 is moderate and depends on the value of b . The same conclusions apply to dairy and cattle farms (not shown here).

Fig. 7 Prediction results (mean and 95% confidence intervals) for crop farm net value added for the quantities displayed in Table 2 including past information (solid lines) and including no information (broken lines)

6 Discussion

The analysis conducted in the present paper allows for an evaluation of the cost of the implementation of the IST mechanism in the Walloon region of Belgium, as recently proposed by the European Commission. This analysis can be extended to any region of the EU since it relies on FADN statistics that are also collected for any other EU region. Even if the total expected amount of indemnities remains moderate for this region, confidence and prediction intervals are rather wide, suggesting considerable yearly variations. A fund operating such an IST should therefore hold a significant amount of capital as a buffer against unexpected losses. Over the years, this capital can be formed by accumulating past surpluses but this need raises some concerns during the first years of the fund's existence.

The IST proposed by the European Commission does not counteract moral hazard. Adverse selection refers to a system in which “opportunistic” individuals are more likely to be selected than “average” individuals due to partially asymmetric information. For the IST, adverse selection is a potential problem because the expected indemnities calculated from available data may underestimate the actual costs of the tool. To this end, the IST benefit could be adjusted. One possibility consists in adopting the following rule inspired from home windstorm insurance. In this case, for the policyholder to be indemnified for his damages, the wind speed as measured in the nearest meteorological station has to exceed a specified threshold, typically 100 km/h. Transposed to our context, the additional condition that could be imposed for obtaining IST benefits could be that the average loss

experienced by a reference group of farmers exceeds some threshold. For instance, we could impose that

$$\sum_i Y_{it} < 0.8 \sum_i \bar{Y}_{it}$$

that is, that the group of farmers belonging to the same region and farm category suffers from a loss of at least 20% compared to the aggregate target income, or we could impose that at least a given percentage of farmers suffered from a loss of at least 30% compared to the target income. Of course, both rules can be combined requiring for instance that at least a given percentage x of the farms experienced a loss exceeding $y\%$ with respect to their target incomes for some appropriate values of x and y . Note that this approach implicitly assumes that the farms have approximately the same economic size. In the limiting case where one farm dominates all the others, this rule could distort competition.

A positive side effect of such additional rules is that it can be expected to favour the emergence of private insurance contracts. Indeed, the systematic risk present in agricultural activities threatens the emergence of a viable insurance market. The additional conditions proposed above confine IST to events affecting the majority of farms. This leaves diversifiable risks that could be efficiently pooled by private insurance companies. Requiring compulsory participation in the system could also efficiently counteract adverse selection.

Rejesus et al [20] study the feasibility of implementing an experience-based premium rate discount system in crop insurance in the United States. In the IST, past experience enters the system in setting the target income, 70% of which playing the role of the sum insured. As discussed earlier, this may pose problems in the descending phase of the cycle but is certainly desirable to avoid maintaining inefficient farms in activity and delaying structural adjustment.

In our model, predictive income distribution is computed by means of the farm-specific risk factor Θ_i that accumulates past experience. Contributions to the mutual fund can be based on the predictive distribution of incomes and so integrate past income trajectory. To make the system more transparent to farmers, simplifications such as linear credibility formulas or bonus-malus scales, for instance, could be considered.

The analysis does not account for the feedback effect of the income stabilisation mutual fund on the choice of the farm activities, which in turn may affect the cost of the mutual fund [22]. Variability in farm net incomes may also increase further in the future as a result of the deregulation of the sector, in particular in the dairy and sugar sub-sectors.

As discussed earlier, the analysis conducted here only concerns amounts of indemnities paid to farmers, and not the system operating the IST. Depending on the rules adopted to determine the compensation to be paid in execution of the IST, running expenses may be considerable, especially to counteract moral hazard. Also, the cost of holding the capital acting as a buffer against unexpected losses or the cost of the guarantee obtained from financial institutions or public bodies needs to be taken into account to evaluate the total costs related to the IST.

To end with, let us mention that the present paper does not consider pricing revenue insurance products but confines to evaluating the amount of indemnities paid in execution of the public IST system. For issues related to pricing agricultural insurance products, we refer the interested reader to Stokes and Nayda [21], Myers

et al [17] and Chambers [5]. Estimates that are generated from our model are nevertheless useful for pricing this kind of insurance product.

Acknowledgements The authors wish to thank the financial support from the *Direction Générale opérationnelle de l'Agriculture, des Ressources naturelles et de l'Environnement* of the Belgian Ministry of Wallonia.

References

1. Anton J, Kimura S, Martini R (2011) Risk management in agriculture in Canada. Agriculture and Fisheries Working Papers, OECD
2. Antonio K, Beirlant J (2007) Actuarial statistics with generalized linear mixed models. *Insurance: Mathematics and Economics* 40(1):58–76
3. Arellano-Valle R, Bolfarine H, Lachos V (2005) Skew-normal linear mixed models. *Journal of Data Science* 3:415–438
4. Bolance C, Guillen M, Pelican E, Vernic R (2008) Skewed bivariate models and nonparametric estimation for the CTE risk measure. *Insurance: Mathematics and Economics* 43:386–393
5. Chambers R (2007) Valuing agricultural insurance. *American Journal of Agricultural Economics* 89:596–606
6. Coble K, Dismukes R, Thomas S (2007) Policy implications of crop yield and revenue variability at differing levels of disaggregation. Paper presented at the American Agricultural Economics Association Annual Meeting, Portland, Oregon, United States
7. DGARNE (2011) Évolution de l'économie agricole et horticole de la Région wallonne 2009-2010, Direction Générale opérationnelle de l'Agriculture, des Ressources naturelles et de l'Environnement, Service public de Wallonie, Namur, Belgium
8. Diaz-Caneja M, Garrido A (2009) Evaluating the potential of whole-farm insurance over crop-specific insurance. *Spanish Journal of Agricultural Research* 7:3–11
9. Eling M (2012) Fitting insurance claims to skewed distributions: Are the skew normal and skew-student good models? *Insurance: Mathematics and Economics* 51:239–248
10. European Commission (2009) Income variability and potential cost of income insurance for EU, Directorate-General for Agriculture and Rural Development, European Commission, Brussels, Belgium
11. European Commission (2011) Proposal for a Regulation of the European Parliament and of the Council on support for rural development by the European Agricultural Fund for Rural Development (EAFRD), European Commission, Brussels, Belgium
12. Greene W (2011) *Econometric Analysis*. Prentice Hall, New Jersey, US
13. Henry de Frahan B, Saegerman C, Denuit M, Dubuisson B, Ledoux O, Pigeon M, Vandeputte S, Weynants S (2011) Étude de la possibilité et proposition de mise en place de mécanismes assurantiers ou de mutualisation des risques dans le secteur agricole en région wallonne. Louvain-la-Neuve, Belgique: Université catholique de Louvain et Université de Liège

14. Lin T (2009) Maximum likelihood estimation for multivariate skew normal mixture models. *Journal of Multivariate Analysis* 100:257–265
15. Mahul O, Wright B (2003) Designing optimal crop revenue insurance. *Agricultural Journal of Agricultural Economics* 85:580–589
16. Meuwissen M, Huirne R, Skees J (2003) Income insurance in european agriculture. *EuroChoices* 2:12–17
17. Myers R, Liu Y, Hanson S (2005) How should we value agricultural insurance contracts? Paper presented at the American Agricultural Economics Association Annual Meeting, Providence, Rhode Island, United States
18. OECD (2011) Managing risk in agriculture: policy assessment and design, OECD Publishing, Paris, France
19. Pigeon M, Antonio K, Denuit M (2013) Individual loss reserving with the multivariate skew normal framework. *ASTIN Bulletin* 43:399–428
20. Rejesus R, Coble K, Knight R, Jin Y (2006) Developing experience-based premium rate discounts in crop insurance. *American Journal of Agricultural Economics* 88:409–419
21. Stokes J, Nayda W (2003) The pricing of revenue assurance: Reply. *American Journal of Agricultural Economics* 85:1066–1069
22. Turvey C (2012) Whole farm income insurance. *Journal of Risk and Insurance* 79:515–540
23. WTO (1994) Final Act of the 1986 - 1994 Uruguay round of trade negotiations, World Trade Organization, Geneva, Switzerland

A Appendix

A.1 Preliminary results

Lemma 1 Let $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Then, for $\mathbf{a}$ ($k \times 1$) and $\mathbf{B}$ ($k \times n$), we have

$$E[\Phi_k(\mathbf{a} + \mathbf{B}\mathbf{X}; \boldsymbol{\nu}, \boldsymbol{\Omega})] = \Phi_k(\mathbf{a}; \boldsymbol{\nu} - \mathbf{B}\boldsymbol{\mu}, \boldsymbol{\Omega} + \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}').$$

Proof. Let $\mathbf{Z} \sim N_k(\boldsymbol{\nu}, \boldsymbol{\Omega})$ and

$$\begin{aligned} E[\Phi_k(\mathbf{a} + \mathbf{B}\mathbf{X}; \boldsymbol{\nu}, \boldsymbol{\Omega})] &= E[P(\mathbf{Z} \leq \mathbf{a} + \mathbf{B}\mathbf{x} | \mathbf{X} = \mathbf{x})] \\ &= E[P(\mathbf{Z} - \mathbf{B}\mathbf{x} \leq \mathbf{a} | \mathbf{X} = \mathbf{x})] \\ &= E[P(\mathbf{U} \leq \mathbf{a} | \mathbf{X} = \mathbf{x})] \\ &= P(\mathbf{U} \leq \mathbf{a}), \end{aligned}$$

where $\mathbf{U} \sim N_k(\boldsymbol{\nu} - \mathbf{B}\boldsymbol{\mu}, \boldsymbol{\Omega} + \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}')$.

Lemma 2 Let $\mathbf{X} \sim N_q(\boldsymbol{\nu}, \boldsymbol{\Omega})$ and $\mathbf{Y} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Then,

$$\begin{aligned} \phi_p(\mathbf{y}; \boldsymbol{\mu} + \mathbf{A}\mathbf{x}, \boldsymbol{\Sigma}) \phi_q(\mathbf{x}; \boldsymbol{\nu}, \boldsymbol{\Omega}) &= \phi_p(\mathbf{y}; \boldsymbol{\mu} + \mathbf{A}\boldsymbol{\nu}, \boldsymbol{\Sigma} + \mathbf{A}\boldsymbol{\Omega}\mathbf{A}') \\ &\quad \times \phi_q(\mathbf{x}; \boldsymbol{\nu} + \mathbf{A}\mathbf{A}'\boldsymbol{\Sigma}^{-1}(\mathbf{y} - \boldsymbol{\mu} - \mathbf{A}\boldsymbol{\nu}), \boldsymbol{\Lambda}), \end{aligned}$$

where $\boldsymbol{\Lambda} = (\boldsymbol{\Omega}^{-1} + \mathbf{A}'\boldsymbol{\Sigma}^{-1}\mathbf{A})^{-1}$.

Proof. Define $\mathbf{z} = \mathbf{y} - \boldsymbol{\mu} - \mathbf{A}\boldsymbol{\nu}$ and $\mathbf{w} = \mathbf{x} - \boldsymbol{\nu}$, we have

$$\begin{aligned} &(\mathbf{z} - \mathbf{A}\mathbf{w})'\boldsymbol{\Sigma}^{-1}(\mathbf{z} - \mathbf{A}\mathbf{w}) + \mathbf{w}'\boldsymbol{\Omega}^{-1}\mathbf{w} \\ &= \mathbf{z}'(\boldsymbol{\Sigma} + \mathbf{A}\boldsymbol{\Omega}\mathbf{A}')^{-1}\mathbf{z} + (\mathbf{w} - \mathbf{A}\mathbf{A}'\boldsymbol{\Sigma}^{-1}\mathbf{z})'\boldsymbol{\Lambda}^{-1}(\mathbf{w} - \mathbf{A}\mathbf{A}'\boldsymbol{\Sigma}^{-1}\mathbf{z}), \end{aligned}$$

and $|\boldsymbol{\Sigma} + \mathbf{A}\boldsymbol{\Omega}\mathbf{A}'||\boldsymbol{\Lambda}| = |\boldsymbol{\Sigma}||\boldsymbol{\Omega}|$.

A.2 Proof of Proposition 1

(i) We have, with $\mathbf{x} = \mathbf{y} - \mathbf{X}\beta - \Omega\mathbf{t}$,

$$\begin{aligned} M_{\mathbf{Y}}(\mathbf{t}) &= E[e^{\mathbf{t}'\mathbf{Y}}] \\ &= \int e^{\mathbf{t}'\mathbf{y}} 2^T \phi_T(\mathbf{y}; \mathbf{X}\beta, \Omega^*) \Phi_T(\Delta'(\Omega^*)^{-1}(\mathbf{y} - \mathbf{X}\beta); \mathbf{0}, \mathbf{I}_T - \Delta'(\Omega^*)^{-1}\Delta) d\mathbf{y} \\ &= 2^T e^{\mathbf{t}'\mathbf{X}\beta + 0.5\mathbf{t}'\Omega^*\mathbf{t}} \int \phi_T(\mathbf{x}; \mathbf{0}, \Omega^*) \Phi_T(\Delta'(\Omega^*)^{-1}(\mathbf{x} + \Omega^*\mathbf{t}); \mathbf{0}, \mathbf{I}_T - \Delta'(\Omega^*)^{-1}\Delta) d\mathbf{x} \\ &= 2^T e^{\mathbf{t}'\mathbf{X}\beta + 0.5\mathbf{t}'\Omega^*\mathbf{t}} E\left[\Phi_T(\Delta'(\Omega^*)^{-1}(\mathbf{X} + \Omega^*\mathbf{t}); \mathbf{0}, \mathbf{I}_T - \Delta'(\Omega^*)^{-1}\Delta)\right], \end{aligned}$$

where $\mathbf{X} \sim N_T(\mathbf{0}, \Omega^*)$. By using Lemma 1, we obtain

$$= 2^T e^{\mathbf{t}'\mathbf{X}\beta + 0.5\mathbf{t}'\Omega^*\mathbf{t}} \Phi_T(\Delta'\mathbf{t}).$$

(ii) The proof is direct from result (i).

A.3 Proof of Proposition 2

The marginal probability density function is given by (with $\Omega = \Sigma + \Delta\Delta'$)

$$\begin{aligned} f_{\mathbf{Y}}(\mathbf{y}) &= \int_{\mathbb{R}^T} MSN(\mathbf{y}|\Theta; \mathbf{X}\beta + \mathbf{Z}\Theta, \Sigma, \Delta) N_T(\Theta; \mathbf{0}, \Psi) d\Theta \\ &= \int_{\mathbb{R}^T} 2^T \phi_T(\mathbf{y}; \mathbf{X}\beta + \mathbf{Z}\Theta, \Omega) \phi_T(\Theta; \mathbf{0}, \Psi) \\ &\quad \times \Phi_T(\Delta'\Omega^{-1}(\mathbf{y} - \mathbf{X}\beta - \mathbf{Z}\Theta); \mathbf{0}, \mathbf{I}_T - \Delta'\Omega^{-1}\Delta) d\Theta. \end{aligned}$$

By using Lemma 2, we have

$$\phi_T(\mathbf{y}; \mathbf{X}\beta + \mathbf{Z}\Theta, \Omega) \phi_T(\Theta; \mathbf{0}, \Psi) = \phi_T(\mathbf{y}; \mathbf{X}\beta, \Omega + \mathbf{Z}\Psi\mathbf{Z}') \phi_T(\Theta; \Lambda\mathbf{Z}'\Omega^{-1}(\mathbf{y} - \mathbf{X}\beta), \Lambda),$$

where $\Lambda = (\Psi^{-1} + \mathbf{Z}'\mathbf{D}^{-1}\Omega)^{-1}$. So, the marginal probability density function can be written as

$$f_{\mathbf{Y}}(\mathbf{y}) = 2^T \phi_T(\mathbf{y}; \mathbf{X}\beta, \Omega + \mathbf{Z}\Psi\mathbf{Z}') E[\Phi_T(\Delta'\Omega^{-1}(\mathbf{y} - \mathbf{X}\beta - \mathbf{Z}\mathbf{W}); \mathbf{0}, \mathbf{I}_T - \Delta'\Omega^{-1}\Delta)],$$

where $\mathbf{W} \sim N_T(\Lambda\mathbf{Z}'\Omega^{-1}(\mathbf{y} - \mathbf{X}\beta), \Lambda)$. By using Lemma 1, we obtain

$$\begin{aligned} f_{\mathbf{Y}}(\mathbf{y}) &= 2^T \phi_T(\mathbf{y}; \mathbf{X}\beta, \Omega + \mathbf{Z}\Psi\mathbf{Z}') \\ &\quad \times \Phi_T(\Delta'(\Omega + \mathbf{Z}\Psi\mathbf{Z}')^{-1}(\mathbf{y} - \mathbf{X}\beta); \mathbf{0}, \mathbf{I}_T - \Delta'(\Omega + \mathbf{Z}\Psi\mathbf{Z}')^{-1}\Delta). \end{aligned}$$