The Australian Journal of
Agricultural and Resource Economics

AARES
AUSTRALASIAN AGRICULTURAL & RESOURCE ECONOMICS SOCIETY

Australian Journal of Agricultural and Resource Economics, 66, pp. 713–728

Models and muddles: comment on 'Calibration of agricultural risk programming models using positive mathematical programming'

Athanasios Petsakos* and Stelios Rozakis†

There is an emerging strand in the agricultural economics literature which examines the calibration of risk programming models using the principles of Positive Mathematical Programming (PMP). In a recent contribution to this journal, Liu et al. (2020) compare three different PMP approaches and attempt to find the 'most practical' method for calibrating risk programming models to be used in policy analysis. In this article, we argue that the comparison design by Liu et al. (2020) is problematic, as it is based on inappropriate metrics and it ignores recent advancements in PMP. This word of caution intends to provide constructive criticism and aims at contributing to the use of risk programming models in policy analysis.

Key words: agricultural policy, mathematical programming, risk & uncertainty.

1. Introduction

Positive Mathematical Programming (PMP) is arguably the most popular method for calibrating mathematical programming models of agricultural supply. Despite the wide acceptance of the method by the modelling community, many authors have pointed out important shortcomings and have not only proposed various improvements to address them, but also extended the use of the PMP approach to other types of models, like risk programming models with a mean-variance (E-V) objective function. Drawing on these developments, Liu *et al.* (2020) – henceforth LvKD – in a recent contribution to this journal, compared three risk programming farm models calibrated with PMP attempting to find the most practical model for use in policy analysis. While studies comparing alternative modelling approaches undoubtedly constitute an important reference for the applied modeller, they need to be meticulously designed. With this note, we show that this is not the case for LvKD and we express serious concerns about the

* Athanasios Petsakos (email: Athanasios.PETSAKOS@ec.europa.eu) is a Researcher at the European Commission, Joint Research Centre, Directorate D – Sustainable Resources, Unit D.4 – Economics of Agriculture, Edificio EXPO, C/ Inca Garcilaso 3E-41092 Seville, Spain.
† Stelios Rozakis is an Associate Professor at the School of Environmental Engineering, Bioeconomy and Biosystems Economics Lab, Technical University of Crete, Kounoupidiana Campus, 73100 Chania, Greece.

doi: 10.1111/1467-8489.12407

design and the metrics used to compare the different models. The comparison does not take into account the advantages and the disadvantages of each approach, it does not consider recent advancements in PMP, and eventually leads to results that we consider misleading.

We lay out our arguments by first providing a brief description of the three models compared by LvKD. The first model assumes a logarithmic utility function, it exhibits decreasing absolute risk aversion (DARA) and was developed by Petsakos and Rozakis (2015). The P&R-DARA model is applied to a single farm only and it aims at recovering profit expectations and covariances with a maximum entropy approach. Contrary to the other two models, it only employs a linear cost function to account for other non-observed (implicit) costs. The second model, developed by Arata *et al.* (2017), assumes an exponential utility function and therefore exhibits constant absolute risk aversion (CARA). The Arata-CARA model uses a cross section of farms to recover farm-specific CARA coefficients and nonlinear (quadratic) implicit costs that are common for all farms in the same farm type. The third model (FSSIM-ME-CARA) was introduced by Jeder *et al.* (2011)[1], it also assumes CARA preferences, it is applied to a single farm only, and its calibration algorithm includes two steps. The first step performs an iteration of the CARA coefficient in order to select the value that leads to an optimal model solution which best approximates the reference (observed) activity levels. Any remaining differences between the observed activity levels and the activity levels resulting from the model with the selected CARA coefficient are corrected during the second step by introducing an implicit quadratic cost function that is estimated with maximum entropy.

2. Model purpose, assumptions and data

Our first concern about the comparison exercise is that the concept itself of the ‘most practical model for policy analysis’, as perceived by LvKD, is ambiguous and rather narrow because it does not consider the actual policy question and the data available to answer that question. For meaningful impact assessment analyses, these issues should be the main drivers guiding the choice of an appropriate model. A closer examination of the three models compared by LvKD reveals that they were developed for different purposes and assume different data availability. The P&R-DARA model, for example, is better suited when farm-level information is insufficient to estimate farm-specific profit expectations and profit covariance matrices, as required by E-V models. In this

[1] We would like to correct LvKD regarding the references they provide for the FSSIM-ME-CARA model in Section 2 of their paper. The correct reference here is Jeder *et al.* (2011). The first cited paper (Louhichi *et al.*, 2010) explains that FSSIM uses two alternative objective functions, one with quadratic costs (PMP but no risk) and one with a mean–standard deviation formulation. This latter model version is calibrated by adjusting the CARA coefficient, but without introducing additional implicit costs (non-PMP calibration). The second cited paper, by Jeder *et al.* (2014), uses a non-risk version of the FSSIM model.

case, analysts often use average regional, or even national, statistics to calculate the needed parameters, although this practice may lead to erroneous model results. The problem of using inappropriate information for characterising the distribution of activity profits is not new (e.g. Gómez-Limón *et al.* 2003) and the re-estimation of the covariance matrix, which is the main idea behind the P&R-DARA model, has also been adopted by other authors (Louhichi *et al.* 2018). It suffices to say that if the analyst has reliable data to calculate expectations and variances of activity profits there would be no reason to re-estimate them. Although LvKD mention the intuition behind the P&R-DARA model, they nonetheless use average regional data in their comparison without taking into account the impact of data availability on model selection.

Along the same line of reasoning, the theoretical assumptions about the behaviour of economic agents (CARA vs. DARA preferences) also reveal the type of policy questions that are best addressed by each model. We note that all three models were developed with the European Union's Common Agricultural Policy (CAP) in mind as the overarching policy setting. As of 2006, a big part of subsidies given under the CAP take the form of decoupled payments. Among the three models examined, only the P&R-DARA model can capture the effect of the variation in decoupled payments on farm production choices, because of its nonlinear E-V objective function. Conversely, if one expects the wealth effects from decoupled payments to be negligible, or for policy questions not involving changes in decoupled payments, like the adoption of agri-environmental schemes for risk management purposes examined by Arata *et al.* (2017), then a model with a DARA objective function may not be required.

Another caveat of the model selection process, related to the theoretical underpinnings of the FSSIM-ME-CARA model, which LvKD have overlooked, is that the two-step process implies that risk considerations are the main source of calibration problems. More precisely, nonlinearities in PMP models are motivated by many reasons like heterogeneous land quality, risk or restricted management capacity (Howitt, 1995). By treating risk separately, and by choosing a CARA coefficient which leads to an optimal solution that is close to the observed activity allocation, the information presumably contained in the quadratic cost function about these other factors is limited. In theory, a CARA coefficient derived from this first step could also lead to perfect model calibration and thus additional quadratic costs would not be needed. On the contrary, the P&R-DARA and Arata-CARA models estimate the risk premium and their implicit cost functions simultaneously without making any prior assumptions about the relative importance of risk, which is in line with other PMP-risk models (Cortignani and Severini, 2012; Louhichi *et al.* 2018; Britz and Arata, 2019).

One of the anonymous reviewers also suggested that the two-step approach of the FSSIM-ME-CARA model is inconsistent altogether because the CARA coefficient derives from a model with linear costs, but it is ultimately used in the second step to calibrate a model with a nonlinear cost function. Since the correct model is assumed to be the latter, the recovered CARA coefficient, and eventually the recovered nonlinear costs are biased because they were estimated using the wrong model. Although we agree, we examine the FSSIM-ME-CARA as a calibration approach and not from an estimation viewpoint, where this critique would be more relevant. Even in an estimation context, however, the induced bias may be small if the first step achieves a close approximation of the observed activity allocation, in the sense that the quadratic cost function will not significantly affect the response of the model (Jeder *et al*. 2011).

The above criticism is not meant to invalidate the two-step process in the FSSIM-ME-CARA model. It simply highlights the fact that the underlying assumptions in a real-case modelling exercise, which may depend on the policy question, the data available, or even the modeller's subjective perceptions about the behaviour of economic agents, will ultimately determine the choice of the appropriate model. For example, if the analyst has reasons to believe that risk considerations are more important than other factors affecting model calibration, and a CARA coefficient that allows to closely approximate the observed activity allocation can be recovered, then the two-step process in the FSSIM-ME-CARA model can be indeed an acceptable calibration approach.

3. Model comparison

To compare the three models, LvKD assumed to know with accuracy the values of all the parameters estimated during PMP calibration that are, in reality, unknown to the analyst. They called them the 'true' parameter set. For the Arata-CARA and FSSIM-ME models, these parameters include the CARA coefficients, and the linear and nonlinear terms of the implicit quadratic cost functions, while for P&R-DARA, they include the linear implicit costs, the profit expectations and the covariance matrix of activity profits. Next, for each model they performed an initial run using these true parameters and considered the resulting optimal solutions as their reference (or observed) values against which all models had to be calibrated. The selection of the most practical model was finally based on (a) whether the calibration process can lead to parameter estimates which closely approximate the true parameter set (the accuracy criterion) and (b) how reasonable the response of each calibrated model is when activity profits change (the response criterion).

3.1 The accuracy criterion

We express our concern about the choice of accuracy as a criterion in the model comparison exercise because accuracy, as defined by LvKD, depends primarily on the estimation method (entropy and least squares) that accompanies each model. LvKD explicitly state in section 4 that they examine the ability of each model to calibrate the true values of the unobserved parameters. This implies that the three PMP models are treated as estimation methods and that they are assessed according to their ability to recover the unknown parameters in their objective function. However, Garnache *et al.* (2017) criticise this practice, insisting that PMP is a calibration tool that should not be judged according to econometric criteria. Building on this critique, we also want to emphasise that PMP is not an estimation method in a statistical sense and should not be assessed as such. It is a calibration principle that allows for reproducing some reference activity allocation in a mathematical programming model. PMP achieves its calibration objective by recovering unknown parameters in the objective function – not necessarily through statistical estimation – so that the model's first-order conditions evaluated at this reference allocation are satisfied. In other words, accuracy, as defined by LvKD, does not seem to be a relevant criterion for PMP models because the goal of a PMP modeller is not to recover the underlying data generation process related to an implicit quadratic cost function or other unknown parameters. In a mathematical programming context, where models are used ex ante to simulate the potential impact from changes in one or more parameters, a more relevant interpretation of accuracy is that of forecasting performance, as an indicator of model response, which can only be assessed ex post (e.g. Kanellopoulos *et al.* 2010).

We believe that the choice of estimation method that is used to recover the unknown parameters in PMP calibration should be seen independently from the individual assumptions of the PMP model itself, namely the implied behaviour of economic agents, the algorithmic steps, the specification of the objective function, the variables that are considered stochastic and the parameters that need to be recovered.

The use of information-theoretic estimation methods in PMP became common practice only after the seminal paper by Paris and Howitt (1998) who used the maximum entropy criterion to calibrate a model with a fully specified **Q** matrix (the nonlinear part of the implicit quadratic cost function). Information-theoretic estimation methods rely on exogenous information and their application in PMP was motivated by the need to address an important criticism, which has been at the very core of the PMP literature ever since Howitt's (1995) seminal paper[2] and concerns the use of a single observation on input–output decisions to recover nonlinear terms in the model's objective function. This under-determinacy of the calibration system in PMP models (more unknown parameters than equations), and the parameter identification problems that arise thereof, forbid the use of traditional estimation techniques like ordinary least squares. Calibration of single-observation PMP models thus constitutes an ill-posed problem in the sense that there is no unique solution to the equation system of first-order conditions.[3] For this reason, early day PMP applications relied on various ad

[2] LvKD (p.3) mention that PMP was introduced by Howitt (2005). However, the PMP approach was first formally introduced by Howitt (1995).

[3] It is beyond the scope of this comment note to offer a detailed discussion on the criticism of the PMP method. We refer the interested reader to the two excellent papers by Garnarche et al. (2017) and Heckelei and Wolff (2003).

hoc methods for calculating the parameters of the quadratic cost function, for example equating average and marginal costs, or setting the linear elements equal to zero (Heckelei, 2002). In this context, maximum entropy, or any other information-theoretic estimation method, can provide a solution insofar as appropriate priors can be found for all the unknown parameters. However, if the selection of priors is arbitrary, the calibration problem remains ill-posed because the model would also calibrate with any other feasible set of priors. The paper by Paris and Howitt (1998) eventually marked the beginning of a new strand in the literature that sought to combine econometric techniques with PMP principles (using first-order conditions) for estimating – not calibrating – mathematical programming models when multi-year observations are available (e.g. Buysse *et al.* 2007; Jansson and Heckelei, 2011; Britz and Arata, 2019). However, most PMP-related studies, including the models examined by LvKD, focus on calibration and clearly lie outside the literature domain on model estimation because multiple data points on farm-specific activity allocations are not always available.

We would like to note at this point that, although we strongly argue that PMP models should be viewed and assessed within a calibration and not an estimation context, it is also true that many of the assumptions invoked by PMP models do rely on econometric studies. A typical example is the use of exogenous information on supply elasticities – further elaborated in the next section – which originate from studies that estimate primal or dual models of agricultural supply (e.g. Guyomard *et al.* 1996; Heckelei and Britz, 2000; Jansson and Heckelei, 2011). Similarly, decision-making under risk and the types of risk preferences often exhibited by farmers have been studied extensively using econometric models (e.g. Chavas and Holt, 1990; Bar-Shira *et al.* 1997; Sckokai and Moro, 2006; Koundouri *et al.* 2009).

To further motivate the argument that the selection of the estimation method is only tangent to a PMP model and not uniquely linked to it, we note that the P&R-DARA and FSSIM-ME-CARA models can be combined with any other estimation approach that relies on prior information, like cross-entropy or the higher posterior density (HPD) estimator. The HPD, in particular, is used for the calibration of the European Commission's CAPRI and IFM-CAP models (Louhichi *et al.* 2018)[4] and it is considered by Heckelei *et al.* (2008) to be computationally simpler than entropy methods. Similarly, the Arata-CARA model could be potentially calibrated with HPD or some entropy method, assuming of course that suitable priors and their

[4] IFM-CAP is the only recent EU-wide farm model that has been used to inform policy decisions (European Commission, 2018) and adopts concepts and assumptions from both the P&R-DARA and the Arata-CARA models, namely a common **Q** matrix and the re-estimation of the covariance matrix. However, LvKD chose not to consider its calibration approach due to lack of necessary data. Although we acknowledge the large data requirements of IFM-CAP, the calibration approach of each model is irrelevant to the actual data needed to run the model itself, especially under a fictitious setting where all unknown parameters are (assumed to be) known a priori.

distributions are available, or adequately justified, for all of the model's unknown parameters. However, from an econometric and estimation viewpoint, the least squares approach in the Arata-CARA model is theoretically preferred to maximum entropy (or other similar method thereof) because it does not rely on prior information, it uses all sample information and it may even allow for drawing inferences on the estimated parameters if the standard conditions for least squares hold. Yet, even if all three models employed the same estimation technique, they would still be difficult to compare with respect to their accuracy, as defined by LvKD, because each model aims at recovering a different set of unknown parameters under different data assumptions. The Arata-CARA model uses cross-sectional information to estimate farm-specific linear PMP terms and a common fully specified **Q** matrix over 16 farms, FSSIM-ME-CARA recovers a quadratic cost function with a diagonal **Q** matrix for a single farm, assuming a single data point (observed activity allocation for an individual farm in a single year), whereas the P&R-DARA model attempts to recover linear PMP costs, profit expectations and the covariance matrix, also for a single farm and a single data point.

The argument that the accuracy criterion largely depends on the selected estimation method can be further explained by analysing the calibration approach behind the FSSIM-ME-CARA model, which LvKD selected as the most practical for policy analysis. In order to obtain a unique global solution under the maximum entropy formalism, prior information should be used for *all* unknown parameters, otherwise the calibration problem remains ill-posed. This condition is clearly not satisfied in the case of the FSSIM-ME-CARA model since the calibration program includes $2K$ equations and $3K+1$ unknown parameters, of which only $2K$ are assigned prior information. To better explain this point, we present equations 20 and 21 in LvKD using the same notation:

$$p_k y_k - \sum_{z=1}^{Z} z_{k,z}^{a_k} \pi_{k,z}^{a_k} - \sum_{z=1}^{Z} z_{k,z}^{Q_{kk}} \pi_{k,z}^{Q_{kk}} \times x_k^0 - \psi - \varphi \sum_{i=1}^{K} S_{k,i} \times x_i^0 = 0 \tag{1}$$

and

$$c_k + \lambda_k = \sum_{z=1}^{Z} z_{k,z}^{a_k} \pi_{k,z}^{a_k} + \sum_{z=1}^{Z} z_{k,z}^{Q_{kk}} \pi_{k,z}^{Q_{kk}} \times x_k^0 \tag{2}$$

where k and i are the indices for crops, p denotes prices, y are yields, z^a and z^Q are the support intervals (priors) for the linear PMP term a and the quadratic PMP term Q, respectively, π is the probability assigned to each $\{z, k\}$ element of these priors, x_k^0 is the observed crop allocation, ψ is the shadow price of land, φ is the CARA coefficient, $S_{k,i}$ is the covariance matrix of crop profits, and λ denotes crop shadow prices. The calibration procedure aims at recovering a, Q,ψ and λ by using priors only for a and Q, which are themselves estimated – through maximum entropy – as the expected value of the prior information z^a and z^Q over the discrete probability distributions π^a and z^Q (φ was recovered in the first step of the process).

An initial remark is that equation (2) is redundant because it introduces K additional unconstrained unknown parameters (λ_k) which do not affect in any way the first-order conditions represented by equation (1). This means that the number of unknown parameters which are relevant for calibration is reduced to $2K+1$ (namely a_k, Q_{kk} and ψ). Second, although the maximum entropy program will return optimal values for a and Q that lie within the bounds defined by their support intervals (the prior information), ψ is free to take any non-negative value. Therefore, equation (1) will always be satisfied as long as ψ can be identified. Identification here refers to the numerical feasibility of the optimisation model and not to the uniqueness of the solution for ψ. In other words, the model will solve if there exists at least one value of ψ that satisfies all individual K equations for given priors a and Q. However, since a and Q are defined over intervals, and also different support intervals for a and Q lead to different ψ values that satisfy equation (1), the solution for ψ is not unique and the calibration problem for the FSSIM-ME-CARA model remains ill-posed.

The implications of this ill-posedness of the FSSIM-ME-CARA model for the accuracy criterion become apparent when we examine more closely the selection of the support intervals for a and Q. The issue of selecting appropriate priors for the unknown parameters of an entropy model is a critical issue in applied modelling, but it has received almost no attention by LvKD. It is widely accepted that priors are an important source of subjectivity from the analyst's part and therefore their values need to be adequately justified. However, selecting priors is not a trivial task because relevant exogenous information may not always exist, and inappropriate priors may bias calibration results, or even lead to numerical infeasibilities if the prior information is not compatible with the observed activity choices.

Although LvKD do not elaborate on the support intervals they used in the FSSIM-ME-CARA model, the accompanying GAMS code and the solution of the maximum entropy program reveal that the priors for the unknown parameters a and Q are defined around their assumed true values. However, since the calibration problem for the FSSIM-ME-CARA model is ill-posed, calibration is possible even when priors are defined over support intervals that do not contain these true values. In this case, the FSSIM-ME-CARA model will still be able to reproduce the reference activity allocation, but with parameter estimates far from their assumed true values, thus completely failing to satisfy the accuracy criterion.

A situation similar to the one we have just described is very likely to occur in a real-case modelling exercise because matrix $\mathbf{Q}$ does not correspond to a tangible and measurable quantity, and therefore, no exogenous information is available. This means that the use of specific priors, which were defined around the true parameter values, was the main reason for the high accuracy of the FSSIM-ME-CARA model in the comparison exercise. However, outside an experimental setting accuracy is not guaranteed because prior information on the model's unknown parameters is not readily available. This problem, which relates to the data availability issues discussed in section 2, contradicts the concept of a practical model that motivated LvKD's comparison exercise. Interestingly, LvKD mention the problem when they justify the use of a diagonal **Q** matrix in FSSIM-ME-CARA, as compared to the fully specified matrix they use for the Arata-CARA model. Specifically, they claim that a full matrix would be inconsistent with only one observation, especially given the lack of information on the possible values of the elements of a fully specified **Q** in a real-case modelling scenario. Nevertheless, LvKD dismiss the Arata-CARA model in favour of FSSIM-ME-CARA because the former cannot accurately recover the true parameter set.

Finally, we need to note that, although the above discussion about the role of priors applies also for the accuracy of the P&R-DARA model, support intervals for its unknown parameters are generally easier to justify than in FSSIM-ME-CARA. As explained by Petsakos and Rozakis (2015), the unknown farm-specific, yield and price distributions are close to the observed regional distributions, while the implicit linear costs are set close to accounting farm profits. Nevertheless, a numerical advantage of the FSSIM-ME-CARA model over P&R-DARA is that the former is less likely to prove infeasible as a result of inappropriate priors because its calibration is an ill-posed problem and the feasible range of prior values is larger than in the case of a well-posed program.

3.2 The response criterion

The response criterion indeed corresponds to a necessary requirement for a calibrated model and therefore appears to be more relevant for comparing optimisation models than its accuracy counterpart. However, the way it is used by LvKD is rather problematic. The problem concerns the true values of the unknown parameters, which were assumed to be known with accuracy in the first step of the comparison. By predefining the implicit cost functions and producing estimates (calibrated parameters) that closely approximated these cost functions, LvKD also predetermined the response of each model because the **Q** matrix largely defines the final model's price elasticity of supply. Please note that we focus here on the **Q** matrix and not on the covariance matrix, although both affect the calibrated model's supply elasticity. In a real-case modelling scenario, and assuming that the analyst has enough data to calculate profit covariances (as in the case of LvKD), the model's elasticities will only depend on the estimation of matrix **Q**.

Calibrating against exogenous information on supply elasticities is a contemporary topic in the literature and a state-of-the-art method in PMP calibration (Mérel and Bucaram, 2010; Garnache and Mérel, 2015). The approach addresses the under-determinacy criticism in PMP which implies that an infinite combination of calibrating parameter values exists, each corresponding to a different model response. By using exogenous information on price elasticities of supply, the range of feasible values for the model's unknown parameters is constrained in such a way so that the model exhibits behaviour which is consistent with this exogenous information. Using supply elasticities is a more elegant way of addressing the under-determinacy criticism than applying prior values directly to a model's unknown parameters. The reason is that (a) parameter-specific information may not always exist, as is the case for the **Q** matrix; and (b) the calibration process simultaneously controls the model's second order properties. This does not mean, however, that the two approaches are mutually exclusive. Using maximum entropy, or equivalent information-theoretic methods, the analyst can allow the model's implied response to deviate from the elasticity priors to ensure that calibration is numerically feasible (Mérel and Bucaram, 2010; Petsakos and Rozakis, 2015).

For the two CARA models, an explicit expression of supply elasticities can be found in Louhichi *et al.* (2018 eq.9), while equivalent elasticity formulas for standard PMP-quadratic cost models can be found in various other papers (e.g. Heckelei and Wolff, 2003 eq.14; Mérel and Bucaram, 2010 eq.4) where it becomes clear how the **Q** matrix affects the models' response to price changes. Similarly, one can derive the price supply elasticity of the P&R-DARA model with equations (18) -(21) in Petsakos and Rozakis (2015).

Although LvKD dismissed the option of using supply elasticity priors in their analysis on the premise that such information does not exist for their study region, they nonetheless imposed a certain supply response to each of the models they examined through their pre-defined true parameter set. This response turned out to be reasonable for the two CARA models but was obviously unreasonable for the P&R-DARA model.

Besides the impact of the true parameter set on how the examined models respond to parameter changes, the behaviour of the P&R-DARA model is also the direct result of the selection of initial wealth, interpreted by LvKD as the average net worth of farms in their study region. On the contrary, Petsakos and Rozakis (2015) had interpreted initial wealth in their model as the sum of decoupled payments, not as the total value of the farm. High initial wealth in the P&R-DARA model leads a to very low absolute risk aversion coefficient, it effectively negates the role of risk in the objective function and thus forces it to behave almost like a linear (risk-neutral) model. Even with the quadratic cost function suggested by LvKD, a large initial wealth would still zero out the risk premium associated with farmers' choices and it would reduce P&R-DARA to a standard PMP-quadratic cost model.[5]

[5] LvKD did not test the quadratic cost function because Petsakos and Rozakis (2015) 'fail to provide technical specifications' for it. Specifying the P&R-DARA model with a nonlinear function is, however, straightforward and mathematically trivial. The interested reader can find the related first-order conditions and the formulas for the calculation of the model's supply elasticities in the Appendix.

This risk-neutral behaviour of the P&R-DARA model should not be confused with the zero CARA coefficient estimated for 10 out of 16 farms in the Arata-CARA model, which suggests that these farms are risk-neutral with respect to *income* risk. The objective function of the P&R-DARA model, as specified by LvKD, refers to *wealth* risk, it expresses the certainty equivalent wealth, and risk neutrality is imposed to the P&R-DARA model because the stochastic part of wealth (crop profits) is very small compared to total farm wealth. LvKD justify their choice of high initial wealth by claiming that the DARA coefficient would otherwise be very sensitive to the choice of the crop mix, resulting in possible calibration failure. This statement is not clear because high initial wealth is not mathematically necessary since the first-order conditions included in the calibration program ensure that the reference activity allocation will always be recovered.[6]

We need to emphasise at this point that the use of total farm value to characterise initial wealth is not by itself theoretically wrong; in fact, it is common practice in most econometric studies on decision-making under risk, especially when investment decisions are dealt with (e.g. Sckokai and Moro, 2009). As we have shown above, the problem instead lies in the different maximisation objective of the three models compared by LvKD, which has important implications for the response of the models in the simulation exercise.

The question that arises from LvKD's treatment of the P&R-DARA model is whether there is scope for a complex DARA objective function in an optimisation model when (a) the research question does not require a DARA model (no decoupled payments); and/or (b) the impact of risk in decision-making is negligible due to the choice of high initial wealth. The first point refers to the policy context and to the questions that each model is designed to answer. LvKD have overlooked both issues in their model comparison. However, the second point seems to be more difficult to answer. Although, in principle, an assumed risk-neutral behaviour does not call for an E-V model, setting initial wealth equal to decoupled payments, instead of total farm value, will admittedly result in high wealth elasticities that may contradict the small wealth effects found in most risk-related econometric studies (Moro and Sckokai, 2013). It is beyond the scope of this comment note to discuss which assumption regarding wealth elasticities is more meaningful for farm-level optimisation models of agricultural supply. We believe, however, that LvKD have serendipitously brought an interesting issue in the applied modelling fore which merits further investigation and has probably not been obvious until now because of the simplicity and dominance of linear E-V models positing CARA preferences and zero wealth effects.

[6] We suspect that the argument of calibration failure that prompted the choice of a high initial wealth parameter in the P&R-DARA model may in fact be a misinterpreted numerical infeasibility problem.

4. Concluding remarks

In this comment, we take a critical look at the findings of the comparison of three agricultural risk PMP models undertaken by LvKD. We argue that the one-size-fits-all comparison method ignores peculiarities of all three models and that the choice of the most practical model is directly tied to the policy question and the availability of data. We then analyse our concerns about (a) the very choice of the accuracy criterion; and (b) the application mode of the response criterion. Regarding the former, we emphasise the fine thread that distinguishes the calibration from the estimation concept, pointing out that models under comparison are meant to be calibration tools. As regards the response criterion, we remind recent advancements in PMP that need to be considered and we examine the specificities of the objective functions of the models under scrutiny in order to shed more light on the results of the comparison exercise. Questions raised about the compatibility of farmers' behaviour with the objective function specification necessitate further research, that along with the discussion per se, can be credited in the merits of the LvKD paper. We do believe that careful experimental design in comparative modelling exercises requires to be aware of the advantages and disadvantages of each model plotted against policy context and data resources. Thus, the applied modeller can consciously select the most appropriate modelling approach for the case under investigation.

Disclaimer

The authors are solely responsible for the content of the paper. The views expressed are purely those of the authors and may not in any circumstances be regarded as stating an official position of the European Commission.

Data availability statement

Data sharing is not applicable to this article as no new data were created or analysed in this study.

References

Arata, L., Donati, M., Sckokai, P. and Arfini, F. (2017). Incorporating risk in a positive mathematical programming framework: a dual approach, *Australian Journal of Agricultural and Resource Economics* 61, 265–284.

Bar-Shira, Z., Just, R.E. and Zilberman, D. (1997). Estimation of farmers' risk attitude: an econometric approach, *Agricultural Economics* 17, 211–222.

Britz, W. and Arata, L. (2019). Econometric mathematical programming: an application to the estimation of costs and risk preferences at farm level, *Agricultural Economics* 50, 191–206.

Buysse, J., Fernagut, B., Harmignie, O., de Frahan, B.H., Lauwers, L., Polomé, P., Van Huylenbroeck, G. and Van Meensel, J. (2007). Farm-based modelling of the EU sugar reform: impact on Belgian sugar beet suppliers, *European Review of Agricultural Economics* 34, 21–52.

Chavas, J. and Holt, M.T. (1990). Acreage decisions under risk: the case of corn and soybeans, *American Journal of Agricultural Economics* 72, 529–538.

Cortignani, R. and Severini, S. (2012). Modelling farmer participation to a revenue insurance scheme by the means of the Positive Mathematical Programming, *Agricultural Economics - Czech* 58, 324–331.

European Commission (2018). Impact assessment. Commission Staff Working Document 301, final (Part 1/3). Available from https://ec.europa.eu/commission/sites/beta-political/files/budget-may2018-cap-swd-part1_en.pdf. [accessed on March 20, 2020].

Garnache, C. and Mérel, P.R. (2015). What can acreage allocations say about crop supply elasticities? A convex programming approach to supply response disaggregation, *Journal of Agricultural Economics* 66, 236–256.

Garnache, C., Mérel, P., Howitt, R. and Lee, J. (2017). Calibration of shadow values in constrained optimisation models of agricultural supply, *European Review of Agricultural Economics* 44, 363–397.

Gómez-Limón, J.A., Arriaza, M. and Riesgo, L. (2003). An MCDM analysis of agricultural risk aversion, *European Journal of Operational Research* 151, 569–585.

Guyomard, H., Baudry, M. and Carpentier, A. (1996). Estimating crop supply response in the presence of farm programmes: application to the CAP, *European Review of Agricultural Economics* 23, 401–420.

Heckelei, T. (2002). Calibration and estimation of programming models for agricultural supply analysis [University of Bonn]. Available from https://www.ilr.uni-bonn.de/agpo/staff/heckelei/heckelei_hab.pdf. [accessed on November 15, 2019].

Heckelei, T. and Britz, W. (2000). Positive mathematical programming with multiple data points: a cross-sectional estimation procedure, *Cahiers d'Economie et Sociologie Rurales* 57, 28–50.

Heckelei, T., Mittelhammer, R. and Jansson, T. (2008). A Bayesian alternative to Generalized Cross Entropy Solutions for Underdetermined Econometric Models (Discussion Paper 2008:2). Institute for Food and Resource Economics, University of Bonn. Available from http://www.ilr.uni-bonn.de/agpo/publ/dispap/download/dispap08_02.pdf. [accessed on April 05, 2020].

Heckelei, T. and Wolff, H. (2003). Estimation of constrained optimisation models for agricultural supply analysis based on generalised maximum entropy, *European Review of Agriculture Economics* 30, 27–50.

Howitt, R.E. (1995). Positive Mathematical Programming, *American Journal of Agricultural Economics* 77, 329–342.

Howitt, R.E. (2005). *Agricultural and Environmental Policy Models: Calibration, Estimation and Optimization*. https://core.ac.uk/download/pdf/48034443.pdf.

Jansson, T. and Heckelei, T. (2011). Estimating a Primal Model of Regional Crop Supply in the European Union, *Journal of Agricultural Economics* 62, 137–152.

Jeder, H., Sghaier, M. and Louhichi, K. (2011). Impact of water pricing policy on the sustainability of the irrigated farming systems: a case study in the South-east of Tunisia, *New Medit* 10, 50–57.

Jeder, H., Sghaier, M., Louhichi, K. and Reidsma, P. (2014). Bio-economic modelling to assess the impact of water pricing policies at the farm level in the Oum Zessar watershed, southern Tunisia, *Agricultural Economics Review* 15, 29–47.

Kanellopoulos, A., Berentsen, P., Heckelei, T., Van Ittersum, M. and Lansink, A.O. (2010). Assessing the forecasting performance of a generic bio-economic farm model calibrated with two different PMP Variants, *Journal of Agricultural Economics* 61, 274–294.

Koundouri, P., Laukkanen, M., Myyra, S. and Nauges, C. (2009). The effects of EU agricultural policy changes on farmers' risk attitudes, *European Review of Agricultural Economics* 36, 53–77.

Liu, X., Kooten, G.C. and Duan, J. (2020). Calibration of agricultural risk programming models using positive mathematical programming, *Australian Journal of Agricultural and Resource Economics* 64, 795–817.

Louhichi, K., Ciaian, P., Espinosa, M., Perni, A. and Gomez y Paloma, S. (2018). Economic impacts of CAP greening: application of an EU-wide individual farm model for CAP analysis (IFM-CAP), *European Review of Agricultural Economics* 45, 205–238.

Louhichi, K., Kanellopoulos, A., Janssen, S., Flichman, G., Blanco, M., Hengsdijk, H., Heckelei, T., Berentsen, P., Lansink, A.O. and Ittersum, M.Van (2010). FSSIM, a bio-economic farm model for simulating the response of EU farming systems to agricultural and environmental policies, *Agricultural Systems* 103, 585–597.

Mérel, P. and Bucaram, S. (2010). Exact calibration of programming models of agricultural supply against exogenous supply elasticities, *European Review of Agricultural Economics* 37, 395–418.

Moro, D. and Sckokai, P. (2013). The impact of decoupled payments on farm choices: Conceptual and methodological challenges, *Food Policy* 41, 28–38.

Paris, Q. and Howitt, R.E. (1998). An analysis of Ill-posed production problems using maximum entropy, *American Journal of Agricultural Economics* 80, 124–138.

Petsakos, A. and Rozakis, S. (2015). Calibration of agricultural risk programming models, *European Journal of Operational Research* 242, 536–545.

Sckokai, P. and Moro, D. (2006). Modeling the Reforms of the Common Agricultural Policy for Arable Crops under Uncertainty, *American Journal of Agricultural Economics* 88, 43–56.

Sckokai, P. and Moro, D. (2009). Modelling the impact of the CAP Single Farm Payment on farm investment and output, *European Review of Agricultural Economics* 36, 395–423.

Appendix

In the case of a quadratic implicit cost function, the maximisation of the certainty equivalent for the P&R-DARA model is given by the following program:

$$\begin{aligned}\max_x \mathrm{CE} &= W_0 + (\mathbf{p}\cdot\mathbf{y})'\mathbf{x} - \mathbf{c}'\mathbf{x} - \mathbf{a}'\mathbf{x} - \frac{1}{2}\mathbf{x}'\mathbf{Q}\mathbf{x} \\ &-\frac{1}{2}\frac{\mathbf{x}'\mathbf{S}\mathbf{x}}{\left(W_0 + (\mathbf{p}\cdot\mathbf{y})'\mathbf{x} - \mathbf{c}'\mathbf{x} - \mathbf{a}'\mathbf{x} - \frac{1}{2}\mathbf{x}'\mathbf{Q}\mathbf{x}\right)}\end{aligned}$$

$$\text{s.t.}\ \mathbf{A}\mathbf{x} \le \mathbf{b}\ [\boldsymbol{\theta}]$$

where W_0 denotes initial wealth, $\mathbf{x}$ is the vector for the activity levels (model variable), $\mathbf{p}$ is the vector of the final estimated output prices, $\mathbf{y}$ is the final estimated activity yield vector, $\mathbf{c}$ is the vector of variable costs, $\mathbf{a}$ is the linear part of the quadratic cost function (all vectors of $I\times 1$ dimension), '•' is the element-wise operator, $\mathbf{Q}$ is the $I\times I$ matrix related to the nonlinear part of the quadratic cost function (can be diagonal with positive elements, or fully specified and positive definite), $\mathbf{S}$ is the final estimated $I\times I$ covariance matrix of activity profits, $\mathbf{A}$ is the $I\times M$ matrix of resource and policy constraints, $\mathbf{b}$ is the $M\times 1$ vector of resource endowments, and $\boldsymbol{\theta}$ is the vector of shadow prices associated with the constraints (also $M\times 1$).

The first-order conditions with respect to $\mathbf{x}$ are given by:

$$\begin{aligned}&\mathbf{p}\cdot\mathbf{y} - \mathbf{c} - \mathbf{a} - \mathbf{Q}\mathbf{x} - \frac{\mathbf{S}\mathbf{x}}{\left(W_0 + (\mathbf{p}\cdot\mathbf{y})'\mathbf{x} - \mathbf{c}'\mathbf{x} - \mathbf{a}'\mathbf{x} - \frac{1}{2}\mathbf{x}'\mathbf{Q}\mathbf{x}\right)} \\ &+\frac{1}{2}\frac{\mathbf{x}'\mathbf{S}\mathbf{x}(\mathbf{p}\cdot\mathbf{y} - \mathbf{c} - \mathbf{a} - \mathbf{Q}\mathbf{x})}{\left(W_0 + (\mathbf{p}\cdot\mathbf{y})'\mathbf{x} - \mathbf{c}'\mathbf{x} - \mathbf{a}'\mathbf{x} - \frac{1}{2}\mathbf{x}'\mathbf{Q}\mathbf{x}\right)^2} - \mathbf{A}'\boldsymbol{\theta} = 0\end{aligned} \tag{S1}$$

Since supply is defined as $\mathbf{q} = \mathbf{x}\bullet\mathbf{y}$, and assuming yields $\mathbf{y}$ are constant, the matrix of price elasticities of supply can be written as

$$\mathbf{E} = \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}\mathbf{p}}\bullet\frac{\mathbf{p}}{\mathbf{x}}$$

We can then define the functions $\mathbf{x} = X(\mathbf{p},\boldsymbol{\theta})$ and $\boldsymbol{\theta} = \Theta(\mathbf{p})$ so that the total effect of $\mathbf{p}$ on $\mathbf{x}$, which also accounts for the effect of $\mathbf{p}$ on $\boldsymbol{\theta}$, is given by the total derivative:

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}\mathbf{p}} = \frac{\partial X}{\partial \mathbf{p}} + \frac{\partial X}{\partial \boldsymbol{\theta}}\frac{\partial \Theta}{\partial \mathbf{p}}$$

Since equation (S1) is highly nonlinear with respect to $\mathbf{x}$, the derivation of the elasticity formula requires the use of the implicit function theorem. Assuming that equation (S1) represents a vector function $F:\mathbb{R}_+^{2I+M}\to\mathbb{R}^I$ such that

$F(\mathbf{x},\mathbf{p},\boldsymbol{\theta})=0$ when $\mathbf{x},\mathbf{p}$ and $\boldsymbol{\theta}$ are evaluated very close to their assumed true values, then the partial derivatives appearing in the elasticity equation can be written as:

$$\frac{\partial \mathrm{X}}{\partial \mathbf{p}}=-\left(\frac{\partial F}{\partial \mathbf{x}}\right)^{-1}\frac{\partial F}{\partial \mathbf{p}}$$

$$\frac{\partial X}{\partial \boldsymbol{\theta}}=-\left(\frac{\partial F}{\partial \mathbf{x}}\right)^{-1}\frac{\partial F}{\partial \boldsymbol{\theta}}$$

$$\frac{\partial \Theta}{\partial \mathbf{p}}=-\left(\frac{\partial F}{\partial \boldsymbol{\theta}}\right)^{-1}\frac{\partial F}{\partial \mathbf{p}}$$

where

$$\begin{aligned}\frac{\partial F}{\partial \mathbf{x}} &= -\mathbf{Q}-\left[\frac{\mathbf{S}}{\left(W_0+(\mathbf{p}\cdot\mathbf{y})'\mathbf{x}-\mathbf{c}'\mathbf{x}-\mathbf{a}'\mathbf{x}-\frac{1}{2}\mathbf{x}'\mathbf{Q}\mathbf{x}\right)}\right]\\ &\quad+\frac{3}{2}\left[\frac{\mathbf{S}\mathbf{x}(\mathbf{p}\cdot\mathbf{y}-\mathbf{c}-\mathbf{a}-\mathbf{Q}\mathbf{x})'}{\left(W_0+(\mathbf{p}\cdot\mathbf{y})'\mathbf{x}-\mathbf{c}'\mathbf{x}-\mathbf{a}'\mathbf{x}-\frac{1}{2}\mathbf{x}'\mathbf{Q}\mathbf{x}\right)^2}\right]\\ &\quad-\frac{1}{2}\frac{\mathbf{x}'\mathbf{S}\mathbf{x}\mathbf{Q}}{\left(W_0+(\mathbf{p}\cdot\mathbf{y})'\mathbf{x}-\mathbf{c}'\mathbf{x}-\mathbf{a}'\mathbf{x}-\frac{1}{2}\mathbf{x}'\mathbf{Q}\mathbf{x}\right)^2}\\ &\quad-\frac{(\mathbf{p}\cdot\mathbf{y}-\mathbf{c}-\mathbf{a}-\mathbf{Q}\mathbf{x})\mathbf{x}'\mathbf{S}\mathbf{x}\mathbf{Q}(\mathbf{p}\cdot\mathbf{y}-\mathbf{c}-\mathbf{a}-\mathbf{Q}\mathbf{x})'}{\left(W_0+(\mathbf{p}\cdot\mathbf{y})'\mathbf{x}-\mathbf{c}'\mathbf{x}-\mathbf{a}'\mathbf{x}-\frac{1}{2}\mathbf{x}'\mathbf{Q}\mathbf{x}\right)^3}\end{aligned}$$

$$\begin{aligned}\frac{\partial F}{\partial \mathbf{p}} &= \mathbf{y}\mathbf{I}+\frac{\mathbf{S}\mathbf{x}(\mathbf{y}\cdot\mathbf{x})'}{\left(W_0+(\mathbf{p}\cdot\mathbf{y})'\mathbf{x}-\mathbf{c}'\mathbf{x}-\mathbf{a}'\mathbf{x}-\frac{1}{2}\mathbf{x}'\mathbf{Q}\mathbf{x}\right)^2}\\ &\quad+\frac{1}{2}\frac{\mathbf{x}'\mathbf{S}\mathbf{x}\,\mathbf{y}\mathbf{I}}{\left(W_0+(\mathbf{p}\cdot\mathbf{y})'\mathbf{x}-\mathbf{c}'\mathbf{x}-\mathbf{a}'\mathbf{x}-\frac{1}{2}\mathbf{x}'\mathbf{Q}\mathbf{x}\right)^2}\\ &\quad-\frac{\mathbf{x}'\mathbf{S}\mathbf{x}(\mathbf{p}\cdot\mathbf{y}-\mathbf{c}-\mathbf{a}-\mathbf{Q}\mathbf{x})(\mathbf{y}\cdot\mathbf{x})'}{\left(W_0+(\mathbf{p}\cdot\mathbf{y})'\mathbf{x}-\mathbf{c}'\mathbf{x}-\mathbf{a}'\mathbf{x}-\frac{1}{2}\mathbf{x}'\mathbf{Q}\mathbf{x}\right)^3}\end{aligned}$$

and

$$\frac{\partial F}{\partial F}=-\mathbf{A}'$$

where $\mathbf{I}$ is the identity $I\times I$ matrix.