Aggregation and Calibration of Agricultural Sector Models Through Crop Mix Restrictions and Marginal Profit Adjustments

Torben Wiborg, Bruce A. McCarl, Svend Rasmussen, and Uwe A. Schneider

For inquiries, please contact:

Torben Wiborg
The Danish Agricultural Advisory Centre
Udkaersvej 15, DK–8200 Aarhus N, Denmark
Email: tw@lr.dk

Professor Bruce A. McCarl
Department of Agricultural Economics, Texas A&M University,
College Station, TX 77843-2124, USA
Email: mccarl@dust.tamu.edu

Associate professor, Dr. Svend Rasmussen
KVL, Department of Economics and Natural Resources
Rolighedsvej 23, DK–1958 Frederiksberg C, Denmark
Email: sr@kvl.dk Author Affiliation and Contact Information

Uwe A. Schneider (referring author)
Assistant Professor
Research Unit Sustainability and Global Change, Hamburg University
Bundesstrasse 55, D-20146 Hamburg, Germany
Email: schneider@dkrz.de

Paper prepared for presentation at the PMP, Extensions and Alternative Methods Organised Session of the XIth EAAE Congress (European Association of Agricultural Economists) The Future of Rural Europe in the Global Agri-Food System Copenhagen, Denmark, August 23-27, 2005

AGGREGATION AND CALIBRATION OF AGRICULTURAL SECTOR MODELS THROUGH CROP MIX RESTRICTIONS AND MARGINAL PROFIT ADJUSTMENTS

Abstract

All agricultural sector models must deal with aggregation and calibration somehow. The aggregation problem involves treating a group of producers as if they all responded in the same way as a single representative unit. The calibration problem concerns making a model reproduce as closely as possible an empirically observed set of decision maker actions. This paper shows how both calibration and aggregation are addressed through crop mix restrictions combined with marginal profit adjustments.

Keywords

Mathematical programming, aggregation, calibration, crop mix, marginal cost, agricultural sector model

JEL Classification

C6 - Mathematical Methods and Programming, C61 - Optimization Techniques; Programming Models; Dynamic Analysis, Q1 – Agriculture, Q11 - Aggregate Supply and Demand Analysis; Prices, Q17 - Agriculture in International Trade, Q18 - Agricultural Policy; Food Policy R12 - Size and Spatial Distributions of Regional Economic Activity; Interregional Trade, R13 - General Equilibrium and Welfare Economic Analysis of Regional Economies, R14 - Land Use Patterns

Why do we need to Aggregate and Calibrate?

There are several reasons why modellers face aggregation and calibration problems. In practice every model contains simplifications of reality, omitting important information, which is taken into consideration by farmers when choosing what crops to plant.

Suppose we formulate an aggregate, uncalibrated programming model that represents a group of farms in a region, as in Model 1.

$$\begin{aligned} &\underset{x}{MAX}\ (\mathrm{r} - \mathrm{c})x \\ &ST \\ &(1) \qquad Ax \le b \\ &(2) \qquad x \ge 0 \end{aligned}$$

Model 1

Also suppose the true real-world situation on one of the farms represented in Model 1 is actually described as shown in Model 2.

$$\begin{aligned} &\underset{x}{MAX}\ (\mathrm{r} - \mathrm{c} - \mathrm{d} - \mathrm{e}x)x \\ &ST \\ &(3) \qquad Ax \le b \\ &(4) \qquad Dx \le f \\ &(5) \qquad x \ge 0 \end{aligned}$$

Model 2

The notation in both models is:

x	An aggregate vector of crop acreages.
r	A vector of average revenue.
c	Variable costs related to the crop acreage, which are *included* in both models. These costs normally include the costs reported in accounting statistics, such as fertiliser, seed, tilling, pesticide, energy and labour costs.
d	Variable costs taken into consideration by the farmer, which are *not included* in Model 1. These may include marketing costs, for example.
e	Variable costs that increase with increasing area of the crop and or revenue declines associated with the rate of decline in the yield with increasing area of the crop. These items occur due to increased disease pressure, marginally decreasing soil quality, etc. These items are taken into consideration by the farmer but are *not included* in Model 1.
A	A matrix of technical coefficients, which are *included* in both models.
b	A resource vector corresponding to matrix A.
D	A matrix with technical coefficients, which are *not included* in Model 1.
f	A resource vector corresponding to matrix D.

As the models are described here, there are important differences. The underlying cause for the omissions in Model 1 is the lack of full information about farm resources and costs, transaction costs, incentives and market conditions. These differences make calibration and aggregation necessary. Adding relevant information to Model 1 could reduce the problem. One could imagine that if all relevant information were added, the model would not only calibrate correctly to current production, but also to all counter factual scenarios, since the incentives and production possibility representation are correct. However, when a model is expanded from describing a few farms to sector or society level, the costs of gathering and maintaining an adequate amount of information for all the included cases are enormous.

The differences between the sector model in Model 1 and the real world as illustrated in Model 2 may be due to the following reasons:

Sector Models Typically Depict Groups of Farms within a Single Sub-model. Usually, a large number of farms of a particular type in a geographical region such as a state or province are represented as a single typical farm. Thus, models contain for example a single sub-model representing all corn-soybean farms in Iowa, or all dairy operations in Sweden, even though there may be hundreds or even thousands of such types of farms. The construction of such typical farms introduces aggregation, and is done as a result of data and model size considerations**.**

Resource Availability and Availability of Details of Potential Production are Typically much less in Sector Model Sub-models than in Individual Farm Models. Typically, there are sub-models within sector models, and these are highly aggregated representations of the operations they depict. They involve annual land, labour and water availability without considering many, or sometimes any, of a large variety of farm specific factors such as crop rotation, quality of labour, land type, implementation and tractor time constraints. For example, in modelling an Indiana corn-soybean-wheat-silage farm, the farm level model employed by McCarl *et al.* (1977) had more than 200 production possibilities and more than 175 resource-related constraints. In contrast, the aggregate ASM model (Chang *et al*., 1993) represents all of Indiana production, including livestock production potential, with less than two dozen production possibilities and a dozen constraints. Data and model size considerations force such an aggregate depiction**.**

Sector Models Typically Ignore Market Factors. Typically, more aggregate models depict regional producers and consumers as if they traded set of homogeneous commodities at a single commodity-specific price. However, it is commonly observed that prices for a single commodity such as hay vary within any region across the places in the region as well as by time of sale and commodity quality characteristics.

The Data Available for use in a Sector Model Force Aggregation on the Modeller. Typically, sector-modelling exercises requires the use (and possibly the development) of consistent data on a national basis. When trying to use or develop such data, one usually finds that such crucial items as crop production budgets are available for average or "typical" regional operations, but not for a large number of possible alternative enterprises. Confidentiality and the costs of finer data development generally preclude more detailed data sets. Price data are also typically averages over regions, days, sale contract terms and grades, as are consumption data.

Differences Exist Between Farmers' and the Model's Objective Function/Constraint Sets. Models often depict profit maximisation subject to resource constraints excluding other potentially important factors. Examples of relevant excluded considerations include risk aversion, financial reserves and personal expectations concerning yields and prices. These are all difficult to measure and depict on an aggregate scale.

Budgets and a Lack of Depiction of Production Possibilities. Budgets give statistically based data describing production practices carried out at one point in time on average, not how it could have been. So budgets do not give a full spectrum of possible responses. Also lags in budget availability (i.e. with those available being one or two years old) and geographical averaging may bias the model response.

Specialisation and Mathematical Programming Solutions. Mathematical programming solutions, particularly those from linear programming models, tend to produce extremely specialised solutions (corner solutions) since the number of production possibilities employed is influenced by the richness of the constraint set and the embodied production possibilities. Thus, model solutions may be generated which give regions as producing only part of the crop and livestock potential production which is actually common within the boundaries of the region.

Transaction Costs are Often Omitted. Many models are built on the basis of farm budgets, but then use consumer or regional-level average prices. There are costs accruing in the marketing channel for handling and transport that are frequently not present in budgets, so there are often price differentials between farm level prices received and prices paid by consumers which are not fully captured in models.

The following section reviews different approaches to these problems, and compares them in the context of the theoretical model.

Convex Combinations of Historical Crop Mixes

McCarl (1982) proposed a method to aggregate over regions in sector models by restricting the crop mix to the space spanned by a convex combination of historical crop mixes. This was followed up in McCarl, Hamilton and Adams (1985) and Önal and McCarl (1989). The model proposed by McCarl (1982) is:

$$\begin{aligned}
&\max_{x,\varphi} \quad (r-c)x \\
&\text{Subject to} \\
&\mathrm{A}x \leq \mathrm{b} \\
&\varphi\hat{X} = x \\
&\sum_i \varphi_i = 1 \\
&\varphi_i \geq 0 \text{ for all } i
\end{aligned} \qquad \textbf{Model 3}$$

Here the vector φ contains the convex combination weights φ_i, the matrix $\hat{X}$ contains exogenous crop mix observations while *i* is an index for the crop mix observations. The main assumption is that there is a duality between solving an aggregate model with the full detail of all the farm firm models included on the one hand, and on the other building an aggregate model without the farm firm models which is constrained to the production possibility set spanned by a convex combination of all possible optimal solutions of the farm firm models. For all practical purposes one could never construct all the detailed farm firm models that would be required in order to find $\hat{X}$, but rather one can use empirical observations on observed crop mixes. This approach has been implemented in a number of settings, for example Chang *et al*., (1993) and Adams *et al*., (1996).

There are some basic problems with this approach. First, the use of historical crop mixes does not constitute as rich a production possibility set, as one would have with the full detail in a model, which more completely represented individual units. Historical crop mixes are reflections of producer decisions in the face of prevailing prices. Thus the crop mixes will not be an accurate representation either if the expected prices confronted by the model are well outside the historical range or if the situation to be examined substantially revises the production possibilities. Önal and McCarl (1989) found that when the prices and product mixes were not substantially different, the historical mixes gave a solution very close to that produced by more disaggregated modelling schemes. Others have attempted to correct such problems by augmenting the historical crop mix information with expert opinion or survey information. Tanyeri-Abur *et al*., (1993) added additional crop mixes containing much less sugar when examining U.S. sugar import policy. Apland and Jonasson (1992) followed a similar approach in eliminating oil crops. Schiable *et al*., (1999) added additional mixes from a farm survey in a study of farm policy revision, with the mixes being based on survey questions about reactions to policy revisions.

Another problem with this approach is that it does not take account of changes in production costs, inputs and yields when crop mixes change. Any farmer knows that crop yields depend very much on the land use the previous year. Furthermore, costs may also vary with the previous crop. Some crops assist in avoiding diseases, thereby reducing the need for pesticides or other crop protection, while other crops have the opposite effect. These cost and yield changes are not taken into account.

Calibration - Closing the Gap between the Marginal Cost of Production and Marginal Revenue

Fajardo *et al*., (1981) calibrate a model of Nicaraguan agriculture by imposing the production economic optimality criterion marginal revenue (*MR*) = marginal cost (*MC*) for all activities. A calibration restriction restricted the model to the empirically observed crop mix, and the shadow price vector from this equation, λ, was interpreted as the unexplained difference between *MR* and *MC*. If perfect competition is assumed farmers must produce such that *MR*=*MC*. The difference between MR and MC in the model must therefore by definition represent the difference between the true optimisation problem and the model.

This difference was placed into a miscellaneous cost category, and in the calibrated model this cost is subtracted. This cost must by definition be omitted costs, such as transaction costs, marketing costs or un-modelled resource costs. Chang *et al*., (1993) and McCarl *et al*., (1998) use this approach in the ASM model in conjunction with convex combinations of optimal productions as described above.

If this problem is viewed in the context of the real and the model version of the optimisation problem (Model 1 and 2), the uncalibrated model corresponds to Model 1, while the calibration model is Model 4 and the calibrated model is Model 5 below:

$$
\begin{array}{ll}
\underset{x}{MAX}\ (\mathrm{r} - \mathrm{c})x & \\
ST & \\
(1) \quad Ax \le b & \\
(7) \quad x \le \bar{x} \quad (\lambda) & \\
(2) \quad x \ge 0 &
\end{array}
\qquad \textbf{Model 4}
$$

$$
\begin{array}{ll}
\underset{x}{MAX}\ (\mathrm{r} - \mathrm{c} - \lambda)x & \\
ST & \\
(1) \quad Ax \le b & \\
(2) \quad x \ge 0 &
\end{array}
\qquad \textbf{Model 5}
$$

If the true problem is Model 2, then the shadow price of the calibration restriction (7), λ, evaluated at $x^* = \bar{x}$ will be $\lambda = d + ex + \gamma D$, where γ is the shadow price from equation (4) in Model 2.

Fajardo *et al*.'s approach is designed to compensate for missing marketing costs and omitted constraints. These arise due to: a) the existence of transactions costs encountered to move goods through the marketing channel which are not reported in the crop budgets or differ by farming location; or b) differences in the quality of and timing of sale for the goods, which are assumed to be homogeneous in the models. However, as demonstrated in (2), other costs, of which some vary in *x*, are also parts of the miscellaneous cost category λ.

A programming model with endogenous prices as used in Fajardo *et al*. (1981) and Chang *et al*. (1993) will calibrate exactly using this technique, but a model with fixed (exogenous) prices will not. This is illustrated in Figure 1, where *S* is the uncalibrated indirect supply function for both models and *S'* is the calibrated indirect supply function for both models. The price endogenous model calibrates uniquely to x^*, while the price exogenous model has a range of optimal points around x^*.

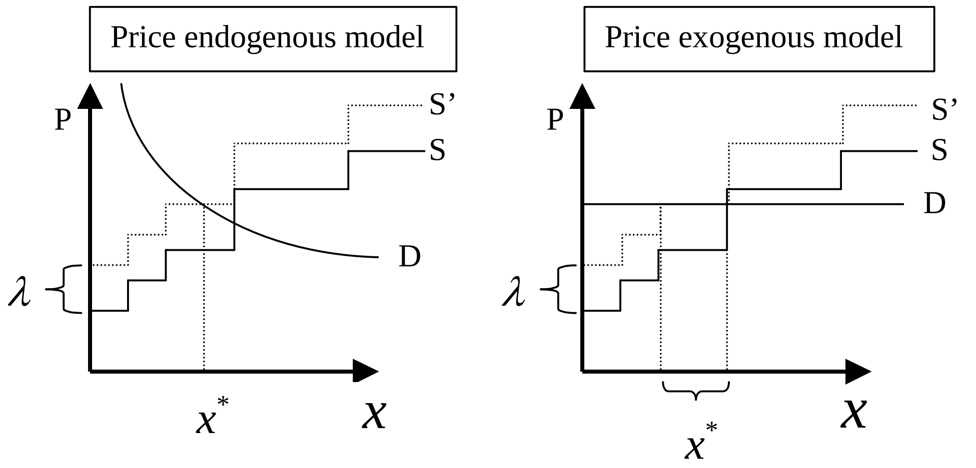

Figure 1: Calibration using Fajardo *et al*'s proposal in price endogenous and price exogenous models.

Conclusion

To evaluate the explained technique, it is necessary to formulate some criteria representing what a good aggregation and calibration procedure should achieve. The obvious calibration criterion is to *reproduce the crop mix* in a given period, and also to *reproduce adjustments under policy shocks*. But the underlying and main criterion is to *distort the production possibilities set to the least degree while providing correct incentives*. If this criterion is fulfilled, the model is likely to reproduce most policy shocks with an acceptable margin of error. Furthermore, the method should be *based in economic theory*, and be *empirically applicable*. Unfortunately, these criteria often conflict (see the discussion in Önal and McCarl (1989)). For example, the necessary conditions for correct aggregation as formulated by Day (1963) are impossible to meet in an empirical model. The perfect method for aggregation and calibration still remains to be discovered, and given the nature of the problem, modellers must resign themselves to using less than perfect solutions.

A number of different techniques deal with aggregation and calibration problems in sector models including the use of flexibility constraints (Miller, 1972) and Positive Mathematical Programming (PMP, Howitt 1995). Perfect aggregation and calibration is not possible, due to the existence of a large amount of missing information that is impossible to obtain and maintain for a larger area. The most interesting calibration candidates are the positive mathematical programming method and Fajardo *et al*.'s method combined with the aggregation approaches involving convex combinations of crop mixes and convex combinations of full-scale farm observations. Both the PMP and Fajardo *et al*.'s methods are biased by their assumption that the missing costs are either only linear in x or only non-linear in x. A better knowledge on the relative size of the linear and non-linear costs would improve the representation of the true underlying optimisation problem. The PMP structure keeps the model very close to the calibration point, but the curvature of the PMP cost functions outside $\bar{x}$ is arbitrary and biases the model response to policy shocks significantly. Convex combinations of crop mixes may be applied simultaneously with PMP or the approach of Fajardo *et al*. However, the methods with added new costs (PMP, Fajardo *et al*.) cannot be combined.

References

Adams, D. M., R. J. Alig, J. M. Callaway, B. A. McCarl and S. M. Winnett (1996) The Forest and Agricultural Sector Optimization Model (FASOM): Model Structure and Policy Applications. PNW-RP-495. U.S. Department of Agriculture, Pacific Northwest Research Station, Portland, Oregon.

Apland, J. and L. Jonasson (1992) The Conceptual Background and Structure of SASM: A Swedish Agricultural Sector Model. Report 45. Swedish University of Agricultural Sciences, Department of Economics, Uppsala.

Chang, C.-C., B. A. McCarl, J. D. Atwood and W. I. Nayda (1993) Resource Policy Analysis. Documentation of the ASM: The U.S. Agricultural Sector Model. (Unpublished)

Day, R. H. (1963) On Aggregating Linear Programming Models of Production. *Journal of Farm Economics* 45:797-813.

Fajardo, D., B. A. McCarl and R. L. Thompson (1981) A Multicommodity Analysis of Trade Policy Effects: The Case of Nicaraguan Agriculture. *American Journal of Agricultural Economics* February. 23-31.

Hamilton, S. A., B. A. McCarl and R. M. Adams (1985) The Effect of Aggregate Response Assumptions on Environmental Impact Analyses. *American Journal of Agricultural Economics* 67:407-413.

Howitt, R. E. (1995) Positive mathematical programming. *American Journal of Agricultural Economics* 2. 77:329-342.

McCarl, B. A. (1982) Cropping Activities in Agricultural Sector Models: A Methodological Proposal. *American Journal of Agricultural Economics* 4. 64:768-772.

McCarl, B. A., W. V. Candler, D. H. Doster and P. R. Robbins (1977) Experiences with Farmer Oriented Linear Programming for Crop Planning. *Canadian Journal of Agricultural Economics* 25:17-30.

Miller, T. A. (1972) Evaluation of Flexibility Restraint Procedures For Recursive Programming Models Used for Prediction. *American Journal of Agricultural Economics* 1. 54:68-76.

Schiable, G. D., B. A. McCarl and R. D. Lacewell (1999) The Edwards Aquifer Water Resource Conflict: USDA Farm Program Resource-Use Incentives. *Water Resources Research* 10. 35:3171-3183.

Tanyeri-Abur, A., B. A. McCarl, C.-C. Chang, R. D. Knutson, W. F. Peterson and K. Koble (1993) An Analysis of Possible U.S. Sugar Import Policy Revisions. *Review of Agricultural Economics* 2. 15:255-268.

Önal, H. and B. A. McCarl (1989) Aggregation of heterogeneous firms in mathematical programming models. *European Review of Agricultural Economics* 4. 16:499-513.