70551

AGGREGATION IN MATHEMATICAL PROGRAMMING SECTOR MODELS AND MODEL STABILITY

1991

by

Hayri Onal and Bruce A. McCarl

Simulation methods

Hayri Onal, Asst. Professor, Department of Agricultural Economics, University of Illinois at Urbana-Champaign.

Bruce A. McCarl, Professor, Department of Agricultural Economics, Texas A&M University, College Station.

AAEA 1991.

#20551

ABSTRACT

This paper discusses the accuracy and stability of an aggregation procedure based on extreme point generation. Numerical results obtained from an empirical model show that the aggregation procedure is highly satisfactory in terms of aggregation errors and the aggregate model is very stable under objective function parameter changes.

AGGREGATION IN MATHEMATICAL PROGRAMMING SECTOR MODELS AND MODEL STABILITY

Introduction

Aggregation is important when modeling the agricultural sector using mathematical programming. Serious errors, such as extreme product specialization and excessive resource exchange between firms, may arise if firms are not properly aggregated (Baker and McCarl; Egbert and Kim). For this reason exact aggregation has attracted a considerable amount of theoretical interest (Day; Miller; Lee; Buckwell and Hazell; Paris and Rausser; Paris). However, the sufficient conditions of exact aggregation have been restrictive and empirically impractical. Consequently, modelers have tried to reduce aggregation bias by grouping the firms according to their characteristics (Frick and Andrews; Sheehy and McAlexander; Buckwell and Hazell; Thompson and Buckwell; Kennedy), introducing flexibility constraints (Sahi and Craddock; Sharples and Schaller) and multiple-crop production activities (Hildreth and Rieter, Heady and Srivastava). The latter approaches, however, are not theoretically supported, rather they are based on subjective judgement or empirical observation.

Aggregation becomes especially difficult when the disaggregate information is hard to compile. The existing literature is mostly based on the assumption of full information about the firm technologies which is again restrictive. Given that full disaggregate information is usually unavailable, a practical approach should only require some sort of aggregate information about the firms rather than individual (micro level) firm data.

The third issue in empirical aggregation is related to the utilization of the aggregate model for policy analysis. Sector models are developed to analyse the impacts of policy instruments wherein model parameters such as prices and costs are changed. Therefore, an important question is whether the aggregate model developed using one set of data still reflects the aggregate behavior of the system after policy changes are made. The aggregate model is required to be stable in the face of relevant policy changes.

The first two issues have been addressed extensively in literature. This paper addresses the third issue, namely the stability of the aggregate model in the context of policy analysis using the aggregation methodology proposed by McCarl. An empirical study will be done on the performance of this aggregation approach in regard with model stability.

The Exact Aggregation Problem in Sector Modeling

The aggregation problem in the mathematical programming context involves development of an aggregate model which generates identical solutions that would be obtained from a large number of independent firm models. Typically this is posed in terms of a set of structurally identical firm models with potentially differing data (see Day). This definition does not well represent the aggregation problem in agricultural sector modeling. First, the firm models are not usually of the same size. Second, the sector analyst is usually interested in the aggregate solution rather than the disaggregate firm solutions. Third, the firm models are not static, rather the objective function coefficients (prices) are determined endogenously through the aggregated firm responses. Finally, the technical coefficients and structures of the firm models are not usually available on a broad scale. Thus an alternative definition is needed.

In agricultural sector modeling the aggregation problem is to develop a model which i) is smaller than the disaggregate model, ii) uses aggregate data depicting firm performance across groups of firms, and iii) generates the same aggregate information obtained from the individual firm responses under the class of policy instruments to be studied.

Mathematically, let the k^{th} firm problem be represented by the linear program (LP):

$$\text{[FM.k]} \qquad \max \ (pY^k - wR^k - vB^k - c^k)\, x^k \tag{1.1}$$

$$\text{s.t.} \qquad A^k x^k \le b^k ,\ x^k \ge 0 \tag{1.2}$$

where p, w, v are vectors of prices for products, fixed price and price endogenous inputs respectively; x^k is a vector of production activity levels; c^k is a vector of exogenous costs per unit production activity; Y^k, B^k, R^k are matrices of yields and input requirements per unit x^k; A^k is a firm level technology matrix; and b^k is a vector of firm level resources. At the firm level p, w, and v are all exogenous. However, at the sectoral level p and v are endogenous and the firm level optimal production activities and endogenous prices must be consistent. This can be achieved by solving the following mathematical program (McCarl and Spreen):

$$\text{[SM]} \qquad \max \{ Z=f(q) - g(s) - c^1x^1 - \dots - c^Kx^K \} \tag{2.1}$$

$$q \quad - Y^1x^1 - \dots - Y^Kx^K \quad = 0 \tag{2.2}$$

$$s \ - R^1x^1 - \dots - R^Kx^K \quad = 0 \tag{2.3}$$

$$A^kx^k \le b^k , \quad x^k \ge 0 \quad \text{for } k=1,..,K \tag{2.4.k}$$

where the new symbols q, s are vectors of aggregate output and input use, and f(q), g(s) are the sum of areas under the product demand and input supply functions, respectively.

In case the disaggregate technical data and economic data are available, the model [SM] is fully described. However, in practice the number of firms (K) is usually so large that the sector model [SM] may not be computationally manageable and/or the structural details may not be specifiable if all firms were explicitly modeled. Therefore, aggregation of firms is necessary but this has to be done without loss of aggregate information.

An Exact Aggregation Procedure

An approach to the solution of the above problem has been suggested by McCarl; and Onal and McCarl (1989). The crop-mix approach proposed by McCarl for modeling production activities in mathematical programming sector models is also an aggregation procedure. This approach is based on the extreme point representation of the firm level models embodied in a large system as used in Dantzig-Wolfe decomposition. Theoretically, a model which is an exact aggregate representation of the original disaggregate problem can be composed by using all of the extreme points from the firm models. However, a model generated in this way would include a very large number of columns and would require firm level data. But, in practice, all extreme points are not needed. A reasonably small number of extreme points usually suffice to depict the relevant portion of the feasible region and to form the desired aggregate model. The decomposition algorithm produces the extreme point information sequentially (Dantzig and Wolfe). Alternatively, a set of extreme points can be pregenerated by using a set of parameter values (Hamilton, McCarl and Adams). The latter method, however, is an approximation rather than being exact aggregation since the extreme points generated may not cover all of the extreme points needed in the exact aggregate model. Whether the full set or a subset of the extreme points are used, the end result of this procedure is as follows:

$$
\begin{array}{llll}
\text{[SM']} & \max & f(q) - g(s) - \Sigma_{i,g}\, c_{ig}\, \lambda_{ig} & (3.1) \\
 & \text{s.t.} & q - \Sigma_{i,g}\, y_{ig}\, \lambda_{ig} = 0 & (3.2) \\
 & & s - \Sigma_{i,g}\, r_{ig}\, \lambda_{ig} = 0 & (3.3) \\
 & & \Sigma_i\, \lambda_{ig} \leq 1 \quad \text{for all } g & (3.4)
\end{array}
$$

where g is a grouping of the firms, c_{ig}, y_{ig}, r_{ig} are the cost, output and input use coefficients aggregated over all firms in the group g at the i^{th} extreme point[1], and the λ variables are the weights associated with the extreme points. Onal and McCarl (1990) showed that [SM'] is equivalent to [SM] and if $(q^*, s^*, (\lambda^*_{ig}))$ is the optimal solution of [SM'] then (q^*, s^*, x^*) is the optimal solution of [SM], where x^* is the convex combination of the extreme points with the weights $\{\lambda^*_{ig}\}$.

This method can incorporate heterogenous firms of any size. Therefore, structural similarity of the firm models is not required. Further, this aggregation procedure requires overall response information on the items in the sector model rather than full disaggregate technical data. If full disaggregate information is not available, the stock of information about the extreme points can be built up by using the aggregate historical data involving total output, input use and associated total costs which are easier to find in official statistics. Such an annual data set for a region can be interpreted as the aggregate optimal solution

[1]namely, $c_{ig} = \Sigma_k c^k x_i^{k*}$, $y_{ig} = \Sigma_k Y^k x_i^{k*}$, $r_{ig} = \Sigma_k R^k x_i^{k*}$ for all k in the group g.

of the firms in the region and multiple year data sets can be used as a proxy to the required extreme point information.

Once the aggregate sector model [SM'] is developed it is desirable that it be usable for repeated analyses without requiring new extreme point data whether generated historically or using submodels. The structure of [SM'] is suggestive of the conditions under which [SM'] would be an adequate representation of [SM], namely if c^k, Y^k and R^k are varied in such a way that the new optimum solution aggregated over firms falls into the convex hull of the extreme points covered or is in a close neighborhood of it. This occurs for sure when the policy changes involve objective function parameters. Minor alterations in the parameters of the firm constraints would also be permissible. This paper investigates the stability issue in the former case, namely when objective function parameters are altered. A case study is carried out for this purpose and the results of this study are presented below.

An Empirical Example

A moderate size sector model for Illinois agriculture is used to see how the methodology works in a real empirical application. The reason for using a real empirical model rather than a simple small size illustrative example is to show the performance of the aggregation procedure in a real modeling situation.

The case example involves a full disaggregate model with nine representative firms (for nine geographical subregions) and an aggregate model where the firms are grouped into three broader regions. The methodology described above is used during the aggregation. Illinois farmers traditionally produce three major crops, namely corn, soybeans and wheat. Alternative production methods are introduced for each region which differ from each other by the time of planting and harvesting operations as well as crop yields. Also these methods are assumed to be used in rotation with each other. The most common rotation practices which are currently applied by Illinois farmers (i.e. corn-corn, corn-corn-soybean, corn-soybean, corn-soybean-wheat) are included as rotation activities in each firm model. Soybean can also be produced as a second crop after wheat in the central and south subregions. With these specifications the model has 15 alternative cropping methods which are combined within 45 different crop rotation activities defined for each subregion. Labor and machinery constraints are defined for the critical biweekly time periods. Both corn and wheat are subject to government support, namely the producers receive an exogenous target price for these crops provided that a certain percentage of the land allocated to these crops is not planted. For simplicity it is assumed that all corn and wheat producing firms participate the government program. Soybean price is purely endogenous and a linear farmgate demand function is used for this crop. Therefore, the only nonlinear variable in the model is the soybean demand (or sale) variable. All inputs but land are assumed to have perfectly elastic supplies while the latter has an inelastic supply in each region. This leads to a sector model which is similar to [SM]. In all the disaggregate model has 328 constraints and 583 variables. This size is quite close to the limit of GAMS/MINOS (Brooke, Kendrick and Meeraus) on a 286 PC-AT computer.

Note that the model involves heterogenous firms, namely the production technologies and resource availabilities are different for each representative firm. Furthermore, the firm models are of different size since second crop soybean activity is not defined for northern firms.

The model is run using the real agronomic and economic data. The results are highly satisfactory in validating the actual production pattern in each region as well as in the state. Since aggregation is the main focus of this paper, only the results related to the aggregation aspect will be reported here.

The extreme points required to build the aggregate model are generated by explicit use of the regional firm models. First, five different price levels are specified for corn and soybean, ranging between $ 2-4 and $ 4-8 respectively, while wheat price is fixed at the observed level. The observed prices fall in the middle of these price ranges. Then each regional firm model (an LP similar to FM-k) is solved using all combinations of those five corn and five soybean prices. The optimum production levels of the three firm LP models included in each of the three aggregate regions are then summed to create a multifirm response. The direct production costs associated with the optimum firm solutions are also summed. This leads to 25 columns for each of the three aggregate regions in the context of [SM']. In all, the aggregate model [SM'] has 6 constraints and 78 variables. This is a much smaller model than the original disaggregate sector model and excludes many complexities that the firm models involve.

To investigate stability of the aggregation under price changes both the full sector model and the aggregate model are solved for twelve alternative corn target prices while soybean price is endogenous. The results obtained are presented in Table 1. For the sake of space, only the results for three target price levels are reported. The table also reports the average absolute deviations between the full sector model and the aggregate model. Both sets of results show very small aggregation error ranging between 0.6 and 4 % in total.

To test the model stability with respect to the nonlinear objective parameters, the price intercept of the soybean demand function is changed by $\pm$ 10 % in both the full sector model and the aggregate model (a demand shift). The results are shown in Table 2. Again the aggregate model performs extremely well. The aggregation errors at regional level as well as the overall state level are negligibly small. These experiments show that i) the aggregation procedure used here is highly satisfactory in regard with the aggregation error involved, ii) in this case the aggregate model is stable under the objective function parameter changes and could be used safely for policy analysis involving price variations once it is developed using the base data.

Conclusion

This paper demonstrated the performance of an aggregation procedure proposed for mathematical programming agricultural sector models and the stability of the aggregate model under the changes in the objective function parameters of the sector model. The aggregation procedure is based on extreme point representation of the individual firm model constraints. It is shown that the aggregation errors involved are negligibly small even with a modest number of extreme points and that the aggregate model is remarkably stable under the mentioned parameter changes. Given that policy analysis usually involves changes in the economic data, the above results demonstrate that this aggregation methodology can be fruitfully used in policy analysis with mathematical programming. The approach can be used for both linear and nonlinear programs with linear constraints and practically any number of firms can be included in the analysis without increasing the size of the aggregate model as long as the aggregate regions remain same.

Table 1: The results of the full sector model and the aggregate model with alternative corn target prices

	Sector Model				Aggregate Model			
	AR1	AR2	AR3	Total[3/]	AR1	AR2	AR3	Total
<u>Corn</u> (million bushels)								
A[1/]	433.6	538.5	347.8	1324.9	433.6	543.5	327.6	1304.7
B	401.3	515.9	306.0	1223.3	410.4	511.2	306.0	1227.6
C	284.3	328.0	189.0	801.4	273.5	360.4	189.0	822.8
MAD (%)[2/]					1.3	2.3	1.7	0.9
<u>Soybean</u> (million bushels)								
A	99.5	113.6	62.9	276.0	99.5	113.6	62.6	275.7
B	112.0	116.3	71.3	299.6	108.6	119.4	71.3	299.3
C	87.1	128.5	90.4	305.9	84.2	130.8	90.4	305.3
MAD (%)					1.8	2.4	1.4	0.6
<u>Wheat</u> (million bushels)								
A	0.	16.5	13.2	29.7	0.	16.5	14.4	30.9
B	0.	28.5	29.4	57.9	0.	26.3	29.4	55.7
C	71.9	106.5	66.1	244.4	77.5	87.1	66.1	230.7
MAD (%)					6.1	4.9	3.4	4.0

1/ A, B and C correspond to the experiments with corn target prices $3.50, 3.00 and 2.50, respectively.

2/ MAD is the percentage mean absolute deviation defined as $100 * \Sigma_k |AGG_k - SEC_k| / K * SEC_k$ where AGG_k and SEC_k correspond to the aggregate and full sector model results with the k^{th} corn target price and $K=12$.

3/ AR1, AR2 and AR3 denote the aggregate north, central and south regions, respectively.

Table 2: The sector model and aggregate model results with alternative demand specifications.

	Sector Model				Aggregate Model			
	AR1	AR2	AR3	Total	AR1	AR2	AR3	Total
Corn (million bushels)								
D[1/]	410.4	511.1	306.0	1227.5	401.3	515.9	306.0	1223.3
E	384.1	491.6	255.5	1131.2	384.1	491.6	257.4	1133.1
F	443.6	565.2	306.0	1314.8	458.1	548.4	306.0	1312.5
MAD					1.8	1.3	0.2	0.2
Soybean (million bushels)								
D	108.5	119.4	71.3	299.3	112.0	116.3	71.3	299.6
E	118.2	126.9	89.5	334.6	118.2	126.9	88.8	334.0
F	95.5	96.6	71.3	263.4	89.6	103.5	71.3	264.5
MAD					3.1	3.2	0.3	0.2
Wheat (million bushels)								
D	0.	26.3	29.4	55.7	0.	28.5	29.4	57.9
E	0.	26.1	29.4	55.5	0.	26.1	29.4	55.5
F	0.	28.5	29.4	57.9	0.	28.5	29.4	57.9
MAD					0.	2.8	0.	1.3

1/ D, E and F are the runs with +10 %, 0%, and -10% changes in the intercept of the soybean demand function.

References

Baker, T. and McCarl, B.A. "Representing Farm Resource Availability Over Time in Linear Programs: A Case Study". North Central Journal of Agricultural Economics, 3-1(1982): 7-21.

Brooke, A.,Kendrick, D. and A. Meeraus. GAMS, A User's Guide. The Scientific Press,1988.

Buckwell, A.E. and Hazell, P.B.R. "Implications of Aggregation Bias for the Construction of Static and Dynamic Linear Programming Supply Models". Journal of Agricultural Economics. 23(1972):119-34.

Dantzig, G.B. Linear Programming and Extensions, Princeton University Press, Princeton, New Jersey, 1963.

Dantzig, G.B. and Wolfe, P. "The Decomposition Algorithm for Linear Programming". Econometrica. 29-4(1961):767-78.

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

Egbert, A.C. and Kim, H.M. "Analysis of Aggregation Errors in Linear Programming Planning Models". American Journal of Agricultural Economics. 57(1975):292-301.

Frick, G.E. and Andrews, R.A. "Aggregation Bias and Four Methods of Summing Farm Supply Function". Journal of Farm Economics. 47(1965):696-700.

Hamilton, S.A., McCarl, B.A., and Adams, R.M. "Improving the Performance of Programming-Based Environmental Analyses of Agriculture". American Journal of Agricultural Economics. 67(1985):407-13.

Heady, E.O. and Srivastava, U.K. Spatial Sector Programming Models in Agriculture. Ames, Iowa: Iowa State University Press, 1975.

Hildreth, C.G. and Reiter, S. "On the Choice of a Crop Rotation Plan". In: Activity Analysis of Production and Allocation. T. J. Koopmans, ed. Cowles Commission Monograph 13, New York: John Wiley and Sons, 1951.

Kennedy, J.O.S. "Using Regression Analysis to Reduce Aggregation Bias to Linear Programming Supply Models". Australian Journal of Agricultural Economics. 19(1975):1-11.

Lee, J.E. "Exact Aggregation: A Discussion of Miller's Theorem". Agricultural Economic Research. 18(1966):58-61.

McCarl, B.A. "Cropping Activities in Agricultural Sector Models: A Methodological Approach". American Journal of Agricultural Economics, 64(1982):768-72.

McCarl, B.A. and Spreen, T. H. "Price Endogenous Mathematical Programming as a Tool for Sector Analysis". American Journal of Agricultural Economics. 62(1980):87-102.

Miller, T.A. "Sufficient Conditions for Exact Aggregation in Linear Programming Models". Agricultural Economics Research. 18(1966):52-57.

Onal, H. and McCarl, B.A. "Aggregation of Heterogeneous Firms in Mathematical Programming Models". European Review of Agricultural Economics. 16-4 (1989):499-513.

Onal, H. and McCarl, B.A. "Exact Aggregation in Mathematical Programming Sector Models". Submitted to Canadian Journal of Agricultural Economics, 1990

Paris, Q. "Perfect Aggregation and Disaggregation of Complementarity Problems". American Journal of Agricultural Economics. 62(1980):681-688.

Paris, Q. and Rausser, G.C. "Sufficient Conditions for Aggregation of Linear Programming Models". American Journal of Agricultural Economics. 55(1973):659-666.

Sahi, R. and Craddock, W. C. "Estimation of Flexibility Coefficients for Recursive Programming Models: Alternative Approaches". American Journal of Agricultural Economics. 56(1976):344-350.

Sharples, J. and Schaller, W. "Predictive Short Run Aggregate Adjustment to Policy Alternatives". Journal of Farm Economics. 50(1968):1523-37.

Sheehy, S. J. and McAlexander, R.H. "Selection of Representative Benchmark Farms for Supply Estimation". American Journal of Agricultural Economics. 47(1965):681-95.

Thomson, K. J. and Buckwell, A. E. "A Microeconomic Agricultural Supply Model". Journal of Agricultural Economics. 30(1979)(1):1-11.