AGGREGATION WITHOUT SEPARABILITY: TESTS OF U.S. AND MEXICAN AGRICULTURAL PRODUCTION DATA

Ni Lin

George C. Davis

C. Richard Shumway

Ni Lin is a consultant in marketing science at Chase Manhattan Bank (and formerly graduate assistant at Texas A&M University), George C. Davis is an assistant professor in the Department of Agricultural Economics at Texas A&M University, and C. Richard Shumway is a professor and chair in the Department of Agricultural Economics at Washington State University (and formerly a professor at Texas A&M University). Senior authorship is equally shared.

Contact author is George C. Davis, 306 Blocker Building, Texas A&M University, College Station, TX 77843-2124. PH:409.845.3788 e-mail:gdavis@tamu.edu.

AGGREGATION WITHOUT SEPARABILITY: TESTS OF U.S. AND MEXICAN AGRICULTURAL PRODUCTION DATA

Abstract

The generalized composite commodity theorem (Lewbel 1996) is used to test for consistent aggregation of U.S. and Mexican agricultural production data in each of the categories for which earlier tests rejected homothetic separability. All U.S. agricultural outputs can be justifiably aggregated into as few as four categories. All Mexican agricultural outputs can be aggregated into as few as five categories. The aggregation of all outputs into a single output cannot be supported in either country by sufficient conditions provided by the generalized composite commodity theorem and/or a homothetically separable technology.

Keywords: aggregation, separability, generalized composite commodity theorem

Note to Discussant: The companion paper from which this paper was carved is about twice as long as this paper. The companion paper describes in more detail the theory and testing procedures and presents more of the empirical results in the body of the paper. Because of space restrictions, this meeting version emphasizes the problem statement and final results without going into minor details regarding the theory, testing procedures, and empirical results. However, for your convenience, we have provided a discussant's appendix with the empirical results from which the conclusions were drawn.

AGGREGATION WITHOUT SEPARABILITY: TESTS OF U.S. AND MEXICAN AGRICULTURAL PRODUCTION DATA

Introduction

A common assumption found in the agricultural economics literature is the existence of a single aggregate agricultural output (e.g., Capalbo and Denny; Clark and Youngblood; Fernandez-Cornejo and Shumway; Luh and Stefanou). As is known, the existence of a single output can be justified if either the Hicks-Leontief composite commodity theorem is satisfied or the aggregate production function is output separable. Recently, Williams and Shumway (1998a, 1998b) conducted extensive nonparametric tests for homothetic separability in U.S. and Mexican agricultural production data. They found no support was found for aggregating all outputs into a single output as is frequently assumed in production analysis.

The Williams and Shumway results are troubling for two reasons. First, if the assumption of a single agricultural output is not satisfied, then many other issues of importance cannot be accurately analyzed within an aggregate modeling framework (e.g., dynamics and expectations (Luh and Stefanou), technical change (Clark and Youngblood), and productivity (Capalbo and Denny)). Second, more disaggregated modeling frameworks are also burdened by major difficulties, such as limited degrees of freedom. Indeed, a viable aggregation alternative to separability or the composite commodity theorem would be welcomed.

Fortunately, Lewbel has identified a third sufficient alternative for consistent aggregation, which he called the generalized composite commodity theorem (GCCT). The importance of Lewbel's theorem is that even if commodity aggregation is not justified by separability or by the Hicks-Leontief theorem, it may be justified by the GCCT. Unlike separability, the GCCT imposes no restrictions on the technology set.

This paper revisits the aggregate output categories considered by Williams and Shumway. The GCCT is used to test for consistent aggregation in each of the categories for which Williams and Shumway rejected homothetic separability. For comparison purposes, several categories are also tested for which Williams and Shumway failed to reject homothetic separability. The hope is that the GCCT will justify a higher degree of aggregation than the separability conditions considered by Williams and Shumway.

Testing Overview

Following Lewbel, let p_i be prices of individual commodities i = 1,2,…,n, where i can be either an input or output. Dropping the i subscript gives the corresponding vector $\boldsymbol{p}$. Let $\boldsymbol{P}$ be vectors of group price indices P_I, where I indexes groups of commodities. Define $r_i = \ln(p_i)$, $R_I = \ln(P_I)$, $\rho_i = \ln(p_i / P_I) = r_i - R_I$, and let $\boldsymbol{r}$, $\boldsymbol{\rho}$, and $\boldsymbol{R}$ be the vectors with elements r_i, ρ_i, and R_I, respectively. Furthermore define $g_i(\boldsymbol{r})$ and $G_I(\boldsymbol{R})$ to be the disaggregate and aggregate netput share demand (supply) functions respectively.

Lewbel proves that if (i) the netput share functions $g_i(\boldsymbol{r})$ for all i are rational, and (ii) the distribution of the vector of relative prices $\boldsymbol{\rho}$ is independent of $\boldsymbol{R}$, then the disaggregate commodities can be consistently aggregated (i.e., $G_I(\boldsymbol{R})$ satisfies all the normal properties). The first assumption is equivalent to profit maximization. The key assumption is the second assumption and in order to justify aggregation by the GCCT, this assumption must be tested.

Lewbel's approach to testing assumption two is straightforward: test if $\boldsymbol{\rho}$ and $\boldsymbol{R}$ are independent. If the variables are stationary, then a correlation test for independence is appropriate. However, as Lewbel discusses, if the variables are nonstationary, a cointegration test is needed. Therefore, the nonstationarity or stationarity of the $\boldsymbol{\rho}$'s and $\boldsymbol{R}$'s must first be tested. Following Lewbel, two stationarity tests are conducted on these variables: the Dickey-Fuller test with a null of nonstationarity and the Kwiatkowski et al. test with a null of stationarity. However, having two tests introduces the possibility of conflicting results.

Let I(0) be the null of a stationary process and I(1) the null of a nonstationary process. There are then three possible tests conclusions: (i) stationarity is rejected if I(0) is rejected but I(1) is not rejected; (ii) stationarity is not rejected if I(0) is not rejected but I(1) is rejected; (iii) the results are indeterminate if both I(1) and I(0) are rejected or not rejected. Because there are three possible tests conclusions and the tests are applied to two variables (ρ_i and R_I), there are a total of nine possible conclusions.

Table 1 summarizes the appropriate test for independence based on the conclusions from the stationarity/nonstationarity tests. Lewbel discusses the first two rows. The third row shows that no testing is required if one series is stationary and the other is nonstationary. This is a direct result of the algebra of cointegration (Granger and Hallman), which says that two series cannot be cointegrated if one is stationary and the other is nonstationary. No cointegration is interpreted here as in Lewbel as suggesting the series are independent. The fourth row indicates that if either ρ_i or R_I is stationary and the other is indeterminate, then a correlation test is appropriate. This follows from the algebra of cointegration. If the indeterminate series is actually nonstationary, then the stationary series cannot be cointegrated with the nonstationary series. However, if the indeterminate series is actually stationary, then the correlation test is appropriate. The fifth row indicates that if either ρ_i or R_I is nonstationary and the other is indeterminate, the appropriate test is a cointegration test. The logic in this case is similar to that for the fourth row. Finally, row six indicates that if both ρ_i and R_I are indeterminate, then both a correlation test and a cointegration test must be conducted. This result combines the logic and results of rows four and five.

Data and Empirical Results

The data source for U.S. agriculture was Ball's (1996) agricultural data set that included annual output prices and quantities 68 outputs for the period 1949-1991. The aggregate group price indices were constructed as Tornqvist indices using 1982 as the base period. The data set for Mexican agriculture was compiled by Williams (1997). It contains 52 output quantities and prices for the period 1966-1991. Price indices were again constructed as Tornqvist indices using 1978 as the base period. Following Lewbel, the tests were conducted for both countries on the nominal and the deflated values. The deflator was a Tornqvist price index over all commodities. As stated, the test for nonstationary is the augmented Dickey-Fuller test and the test for stationarity is the Kwiatkowski et al. test. The cointegration test is the Engle-Granger test and the correlation test is the Spearman rank test (Mendenhall, Scheaffer, and Wackerly).

Table 2 identifies the 13 groups of U.S. commodities tested for consistency with the GCCT. The common letters in each column indicate which commodities are hypothesized to be grouped together. Group A - Group G includes several commonly used output groups in agricultural economics research. Group H - Group L are some alternative groupings considered in Williams and Shumway's (1998b) study. Group M hypothesizes that all outputs can be aggregated into a single group.

Table 3 identifies the 17 groups of Mexican commodities tested for consistency with the GCCT. Group A - Group G exhaustively index the output commodity set. Group Q considers aggregation of all outputs into a single group. The other groups are other alternatives considered by Williams and Shumway (1998a) for possible aggregation. Cucumbers and squash were not considered in the analysis of the Mexican data due to incomplete data.

Table 4 gives a summary of our findings coupled with those of Williams and Shumway. [Note to discussant's. Because of space limitations, the discussant's appendix gives the detailed tables for all the analysis that is summarized in table 4]. Based on their homothetic separability tests, Williams and Shumway (1998a, 1998b) found support for aggregating all agricultural outputs in each country into as few as 11 categories, some of which only included a single commodity. When combined with the additional test results for generalized composite commodity theorem, the empirical evidence supports exhaustive aggregation of all agricultural outputs into as few as four categories in the U.S. and five categories in Mexico. The smallest set of justified aggregates in the U.S. consists of livestock-feed-food grains, vegetables, fruits-nuts, and oilseeds-other field crops. In Mexico, the smallest set consists of livestock, grains-oilseeds, annuals-vegetables, fruits, and other perennials. Options exist for aggregating outputs into a larger number of categories in each country, but no support is provided by these combined test results for aggregating outputs into a smaller number of categories.

Conclusions

Lewbel's (1996) generalized composite commodity theorem was used to test for 13 aggregate U.S. output groups and 17 aggregate Mexican output groups. These groups were the same as had previously been tested, and many rejected, by Williams and Shumway (1998a, 1998b) for a homothetically separable technology. Either property constitutes a sufficient condition for commodity-wise aggregation and consistent two-stage choice modeling.

Empirical support was found for five composite commodities in the U.S. and six in Mexico. When combined with output groups for which Williams and Shumway failed to reject homothetic separability, all U.S. agricultural outputs can be justifiably aggregated into as few as four categories. All Mexican agricultural outputs can be aggregated into as few as five categories.

Given that the existence of a single output aggregate production function is the most commonly maintained hypothesis in primal and dual specifications of production, the rejection of this hypothesis brings into question the results of all studies based on this underlying assumption. Buccola and Sil have shown that measures of productivity can suffer significant bias when output is incorrectly assumed to be separable and hence aggregable. Because of the errors of inference that can occur with misspecified models, it is clear that empirical testing is warranted with other data sets before glibly aggregating all outputs into a single category.

Table 1. Appropriate Test for Independence

Stationary/Nonstationary Results	Test
ρ_i and R_I are both stationary	Correlation
ρ_i and R_I are both nonstationary	Cointegration
ρ_i or R_I is stationary and the other is nonstationary	None
ρ_i or R_I is stationary and the other is indeterminate	Correlation
ρ_i or R_I is nonstationary and the other is indeterminate	Cointegration
ρ_i and R_I are both indeterminate	Correlation & Cointegration

Table 2. U.S. Generalized Composite Commodity Test Groups

Number	Output	Aggregation Test Groups[a]				
1	cattle	B	H		K	M
2	eggs	B	H		K	M
3	hogs	B	H		K	M
4	sheep(composite of sheep and lamb)	B	H		K	M
5	milk sold directly to consumer	A	H		K	M
6	milk sold to plant and dealer	A	H		K	M
7	milk utilized on farm	A	H		K	M
8	miscellaneous livestock	B	H		K	M
9	poultry	B	H		K	M
10	barley	E	I	J	K	M
11	corn	E	I	J	K	M
12	cotton		I		L	M
13	flaxseed	C	I			M
14	hay	E	I	J	K	M
15	miscellaneous crops		I		L	M
16	oats	E	I	J	K	M
17	peanuts	C	I			M
18	rice	D	I	J		M
19	rye	D	I	J		M
20	sorghum	E	I	J	K	M
21	soybeans	C	I			M
22	tobacco		I		L	M
23	wheat	D	I	J		M
24	asparagus	F	I		L	M
25	broccoli	F	I		L	M
26	carrots	F	I		L	M
27	cauliflower	F	I		L	M
28	celery	F	I		L	M
29	cucumbers	F	I		L	M
30	dry beans	F	I		L	M
31	fresh sweet corn	F	I		L	M
32	fresh tomatoes	F	I		L	M
33	honeydew melons	F	I		L	M
34	lettuce	F	I		L	M
35	onions	F	I		L	M
36	peas	F	I		L	M
37	potatoes	F	I		L	M
38	processed sweet corn	F	I		L	M
39	processed tomatoes	F	I		L	M
40	snap beans	F	I		L	M
41	sweet potatoes	F	I		L	M
42	almonds	G	I		L	M
43	apples	G	I		L	M
44	apricots	G	I		L	M
45	avocados	G	I		L	M
46	cranberries	G	I		L	M
47	dates	G	I		L	M
48	figs	G	I		L	M
49	filberts	G	I		L	M

50	grapes	G	I		L	M
51	grapefruit	G	I		L	M
52	lemons	G	I		L	M
53	limes	G	I		L	M
54	macadamia nuts	G	I		L	M
55	nectarines	G	I		L	M
56	olives	G	I		L	M
57	oranges	G	I		L	M
58	peaches	G	I		L	M
59	pears	G	I		L	M
60	pecans	G	I		L	M
61	plums	G	I		L	M
62	prunes	G	I		L	M
63	strawberries	G	I		L	M
64	sweet cherries	G	I		L	M
65	tangelos	G	I		L	M
66	tangerines	G	I		L	M
67	tart cherries	G	I		L	M
68	walnuts	G	I		L	M

[a] Group Codes: A dairy, B other livestock, C oilseeds, D food grains, E feed grains-hay, F vegetables, G fruits-nuts, H livestock, J grains-hay, K livestock-feed, L other crops, M all outputs.

Table 3. Mexican Generalized Composite Commodity Test Groups

Number	Output	Aggregation Test Groups[a]				
1	corn	A	H		O	Q
2	beans	A	H		P	Q
3	wheat	A	H		P	Q
4	rice	A	H		P	Q
5	sorghum	A	H		O	Q
6	barley	A	H		O	Q
7	cottonseed	B	I	M	N	Q
8	safflower	B	H	M	N	Q
9	sesame	B	H	M	N	Q
10	soybean	B	H	M	N	Q
11	alfalfa	D	J		O	Q
12	henequen	D	J	M	N	Q
13	green chili	E	I			Q
14	potatoes	E	I			Q
15	tomatoes	E	I			Q
16	copra	D	J	M	N	Q
17	cucumbers					
18	squash					
19	onions	E	I			Q
20	strawberries	E	I			Q
21	avocados	F	J			Q
22	bananas	F	J			Q
23	oranges	F	J			Q
24	lemons	F	J			Q
25	cacao	D	J	M	N	Q
26	apples	F	J			Q
27	grapes	F	J			Q
28	coffee	D	J	M		Q
29	sugar cane	C	I		N	Q
30	watermelons	E	I			Q
31	tobacco	C	I		N	Q
32	oats	A	H		O	Q
33	pineapple	F	J			Q
34	peaches	F	J			Q
35	mangos	F	J			Q
36	encarcelada nuts	D	J	M	N	Q
37	peanuts	C	H		N	Q
38	dry chili	E	I			Q
39	papaya	F	J			Q
40	cotton lint	C	I		N	Q
41	chickpeas	E	I			Q
42	beef	G	K		O	Q
43	goat	G	K		O	Q
44	hogs	G	K		O	Q
45	sheep	G	K		O	Q
46	chickens	G	L		O	Q
47	turkeys	G	L		O	Q
48	cow milk	G	L		O	Q
49	goat milk	G	L		O	Q

50	eggs	G	L	O	Q
51	honey	G	L	O	Q
52	wool	G	L	O	Q

[a] Group Code: A grains, B oilseeds, C other annual crops, D other perennial crops, E vegetables, F fruit, G livestock, H grains and oilseeds, I annuals and vegetables, J perennials and fruits, K meat animals, L other livestock, M oilseeds and other field crops-A, N oilseeds and other field crops-B, O livestock and feed, P food grains, Q all outputs.

Table 4. Output Categories Consistent with Nonparametric Homothetic Separability and/or Generalized Composite Commodity Theorem Tests

Country	Number of Categories	Aggregation Groups[a]
U.S.	4	D, G, O, P
	5	D, F, G, K, O
		F, G, H, J, O
	6	D, F, G, H, J, N
	7	A, B, D, F, G, J, N
Mexico	5	D, F, G, H, I
	6	D, F, H, I, K, L

a. See tables 2 and 3 for commodities in aggregate groups.

References

Ball, V. E. "U.S. Agricultural Data Set." Unpublished data, U.S. Department of Agriculture, Washington DC, 1996.

Buccola, S. and Sil, J. "Productivity in the Agricultural Marketing Sector." *American Journal of Agricultural Economics* 78(December 1996): 1366-1370.

Capalbo, S. and Denny, M. "Testing Long-Run Productivity Models for the Canadian and U.S. Agricultural Sectors." *American Journal of Agricultural Economics* 68(August 1986):615-625.

Clark, S. and Youngblood, C. "Estimating Duality Models with Biased Technical Change: A Time Series Approach." *American Journal of Agricultural Economics* 74(May 1992): 353-360.

Dickey, D.A., and W.A. Fuller. "Distribution of the Estimators for Autoregressive Time Series with a Unit Root." *Journal of the American Statistical Association* 74(July 1979): 427-31.

Engle, R. F., and C.W. J. Granger. "Cointegration and Error Correction: Representation, Estimation, and Testing." *Econometrica* 55(March 1987): 251-76.

Granger, C. W. J. and Hallman, J. "The algebra of I(1)." *Finance and Economics Discussion Series, paper 45,* Division of Research and Statistics, Federal Reserve Board, Washington DC. 1989.

Fernandez-Cornejo, J. and Shumway, R. "Research and Productivity in Mexican Agriculture." *American Journal of Agricultural Economics* 79(August 1997): 738-753.

Kwiatkowski, D., P. C. B. Phillips, P. Schmidt, and Y. Shin. "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root." *Journal of Econometrics* 54(Oct./Dec. 1992):159-78.

Lewbel, A. "Aggregation Without Separability: A Generalized Composite Commodity Theorem." *American Economic Review* 86(June 1996): 524-43.

Luh, Y. and Stefanou, S. "Estimating Dynamic Dual Models Under Nonstatic Expectations."

American Journal of Agricultural Economics 78(November 1996): 991-1003.

Mendenhall, W., R. L. Scheaffer, and D. D. Wackerly. *Mathematical Statistics with Applications*. Boston: PWS-Kent Publishing Company, 1990.

Williams, S.P. “Trade Liberalization Impacts on U.S. and Mexican Agriculture.” Ph.D. dissertation, Texas A&M University, 1997.

Williams, S. P., and C. R. Shumway. “Aggregation of Data and Profit Maximization in Mexican Agriculture.” *Applied Economics* 30(1998a):forthcoming.

____________. “Testing for Behavioral Objective and Aggregation Opportunities in U.S. Agricultural Data.” *American Journal of Agricultural Economics* 80(February 1998b):forthcoming.

Discussant's Appendix

Table D.1. U.S. Generalized Composite Commodity SummaryTest Results

Groups	Commodities in Group[a]	Cointegrated Commodities		Correlated Commodities		Composite Commodity	Separable Group[b]
		Nominal	Deflated	Nominal	Deflated		
A. Dairy	5,6,7	5,7				No	Yes
B. Other Livestock	1,2,3,4,8,9			1,2,3,4	1,2,3,4	No	Yes
C. Oilseeds	13,17,21					Yes	Yes
D. Food Grains	18,19,23					Yes	Yes
E. Feed Grains-Hay	10,11,14,16, 20				10,11,14,16	No	No
F. Vegetables	24-41					Yes	Yes
G. Fruits – Nuts	42-68					Yes	No
H. Livestock	1-9	2				No	Yes
I. Crops	10-68		17,41, 47,65			No	No
J. Grains – Hay	10,11,14,16, 18,19,20,23					Yes	No
K. Livestock – Feed	1-11,14,16, 20	10,16				No	Yes
L. Other Crops	12,15,22,24-68	22	41		24,28,38,40, 45,46,48, 52, 55,57,59,63	No	No
M. All Outputs	1-68	5,7,14, 16,56			1-5, 7,13-15, 19,21,23-30, 34,36-38, 40, 41,43-46,48, 50,52,53,55, 56,58,65,66, 68	No	No
N. Other Field Crops	12,15,22					c	Yes
O. Oilseeds - Other Field Crops	12,13,15,17,21,22					c	Yes
P. Livestock -Feed – Food Grains	1-11,14,16, 18-20,23					c	Yes

[a] Refer to commodity numbers in Table 2.
[b] Based on Williams and Shumway's (1998) nonparametric test conclusions.
[c] Not tested.

Table D.2. Mexican Generalized Composite Commodity SummaryTest Results

Groups	Commodities in Group[a]	Cointegrated Commodities		Correlated Commodities		Composite Commodity	Separable Group[b]
		Nominal	Deflated	Nominal	Deflated		
A. Grains	1,2,3,4,5,6,32		32	2,4,6,32	1,2,4,6	No	Yes
B. Oilseeds	7,8,9,10					Yes	No
C. Other Annual Crops	29,31,37,40			37	37	No	Yes
D. Other Perennial Crops	11,12,16,25,28, 36					Yes	Yes
E. Vegetables	13-15,19,20,30, 38,41			19	13,14,19	No	No
F. Fruits	21-24,26,27,33-35,39					Yes	No
G. Livestock	42-52					Yes	No
H. Grains and Oilseeds	1-6,8-10,32,37			2,4,6,8-10,32,37	1,2,4,10,37	No	Yes
I. Annuals and Vegetables	7,13-15,19,20, 29-31,38,40,41			7,14,15	7,20	No	Yes
J. Perennials and Fruits	11,12,16,21-28, 33-36,39				22,27,34,39	No	No
K. Meat Animals	42-45					Yes	Yes
L. Other Grains	46-52				49-52	No	Yes
M.Oilseeds and Other Field Crops-A	7-10,12,16,25, 28,36	16			7,8,12,16	No	Yes
N. Oilseeds and Other Field Crops-B	7-10,12,16,25, 29,31,36,37,40		9,12,37			No	Yes
O. Livestock and Feed	1,5,6,11,32,42-52		5	5,6,42-44, 49,50,52		No	No
P. Food Grains	2-4					Yes	Yes
Q. All Outputs	1-16,19-52			1-6,10,13, 14,19,20, 24,26,29, 34,36-38, 42,49,50	2,3,6,13, 14,19,20, 23,24,26, 29,34,36-38,42,49	No	No

[a] Refer to commodity numbers in Table 3.
[b] Based on Williams and Shumway's (1998a) nonparametric test conclusions using 13% measurement error as the rejection criterion.