JOURNAL OF THE

Northeastern
Agricultural
Economics
Council
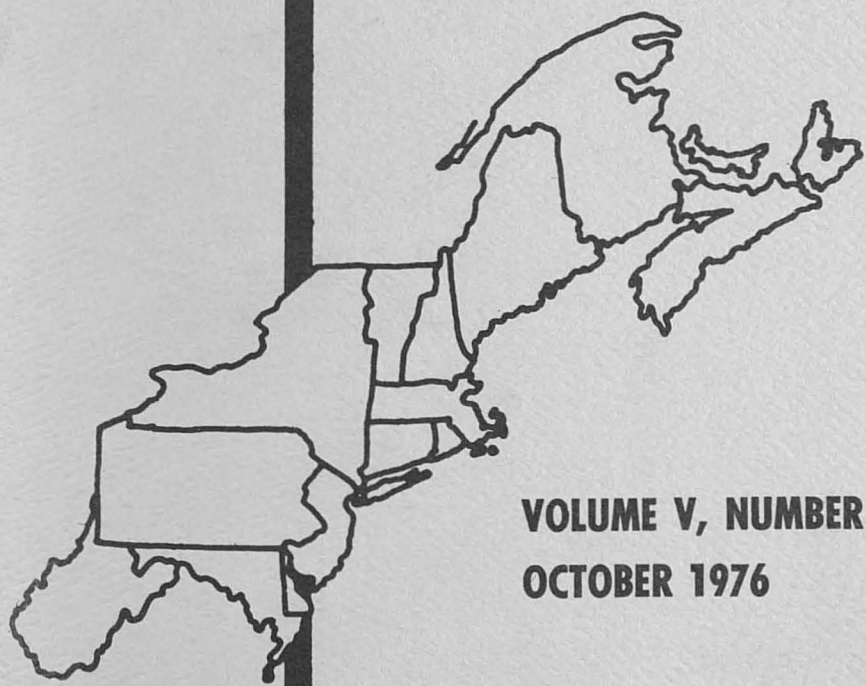
VOLUME V, NUMBER 2
OCTOBER 1976

THE SOIV ESTIMATOR: A COMPARISON WITH ALTERNATIVE
ESTIMATORS IN A MODEL OF THE U.S. APPLE INDUSTRY

Ronald R. Piggott
Lecturer in Economics
University of Newcastle
Newcastle, Australia

William G. Tomek
Professor of Agricultural Economics
Cornell University

Multicollinearity and a lack of degrees of freedom are problems in
simultaneous systems, particularly in large models where the number of
predetermined variables may exceed the number of observations. Such
problems are somewhat less obvious in the small and medium-sized models
typically estimated by agricultural economists. Thus, estimators de-
signed to cope with the degrees of freedom (undersized sample) problem
have received little attention in agricultural economics applications
relative to macroeconometric uses.

In estimating a model of the U.S. apple industry, we experimented
with an estimator suggested by Franklin Fisher [3, 4, 6], structurally
ordered instrumental variables (SOIV), as well as with two-stage least
squares based on principal components of the predetermined variables
(TSLSPC).[1] The simultaneous component of the model contains eight
equations with 10 predetermined variables, and the data consists of 20
annual observations for the crop years 1952-71. Some variables are
highly collinear. This model is perhaps a reasonable example of the
kind of simultaneous model often estimated in agricultural economics.

Since SOIV has received little attention in agricultural economics,
we think that a discussion of the method as well as a comparison with
alternative estimators in an empirical application to a small model
should be useful to others. Hence, this paper briefly describes SOIV
and some problems in its use, and the results obtained from SOIV esti-
mation are compared with those from TSLS and TSLSPC. The main question
is whether the added computational burden of SOIV or TSLSPC provides
benefits in terms of more precise estimates of the parameters. There
is also the question of how SOIV compares with the better known TSLSPC.
In addition, we mention some of the practical difficulties of applying
these procedures.

_____

[1] Another alternative, a three-step procedure attributed to Jorgenson
and Maddala, is not considered here [see 4, p. 27].

The Model[2/]

The apple crop-year extends approximately from July 1 through June 30. During the first six months (period 1), apples are sold from the orchard for current fresh use, canning (mainly sauce and slices) and juice products. Presumably, relative prices in each end use and production are the important determinants of the quantities allocated to each of these end uses. In turn, relative prices would be influenced by the quantities allocated to each end use. Hence, the model assumes that prices and quantities in period 1 are simultaneously determined.

Apples not sold in period 1 are either stored for fresh use in the second six-month period (period 2), or they are abandoned. The model assumes that the quantity placed into storage in period 1 is determined by period 1 prices as well as production. The price of fresh apples in period 2 would be determined, in part, by the fixed quantity available for use in that period. Hence, price is the only current endogenous variable in period 2.

The model is presented below, and variable definitions are given in Table 1. Linear functional forms are assumed throughout. Endogenous variables are denoted by Y's and predetermined variables by X's. The plus or minus sign indicates the expected sign of the relationship between right-hand-side (RHS) and left-hand-side (LHS) variables. The expected sign of Y_1 in equation (7) is uncertain [see 8, pp. 36-41].

(1) $Y_1 = f_1 \ (-Y_4, \ -Y_5, \ +X_2, \ -X_3, \ -X_9)$ Period 1 fresh demand

(2) $Y_2 = f_2 \ (-Y_4, \ -Y_5, \ +X_2, \ -X_5, \ -X_{10})$ Canning demand

(3) $Y_3 = f_3 \ (-Y_6, \ +X_1, \ +X_2, \ -X_4, \ -X_{10})$ Juice demand

(4) $Y_4 = f_4 \ (+Y_1, \ -Y_2, \ +X_6, \ +X_7, \ -X_8)$ Period 1 fresh allocation

(5) $Y_5 = f_5 \ (-Y_1, \ +Y_2, \ -Y_3, \ +X_6, \ +X_7, \ +X_8)$ Canning allocation

(6) $Y_6 = f_6 \ (-Y_2, \ +Y_3, \ +X_6, \ +X_7)$ Juice allocation

(7) $Y_7 = f_7 \ (\underline{+}Y_1, \ -Y_2, \ +X_6, \ +X_7, \ +X_8)$ Period 2 fresh allocation

(8) $10(Y_8) \equiv X_6 + X_7 - 10 \ (Y_4 + Y_5 + Y_6 + Y_7)$ Identity

(9) $Y_9 = f_9 \ (-X_9, \ -X_{11}, \ +X_{12}, \ -X_{13}, \ +X_{14})$ Period 2 fresh demand

---

[2/] For a detailed model description see [8, ch. 2].

Table 1.
Variable Definitions

Variable	Definition
Y_1, Y_2, Y_3, Y_9	Nominal prices of fresh apples period 1, canning apples, juice apples, fresh apples period 2, respectively, $'s per ton.
Y_4, Y_5, Y_6, Y_7, Y_8	Quantities of apples for fresh use period 1, canning, juice, fresh use period 2, and abandonment, respectively, tons per 10,000 persons.
X_1	Dummy variable equal 0 for 1952-1961 crop years and 1 for 1962-1971 crop years.
X_2, X_{12}	Nominal personal disposable income periods 1 and 2, respectively, $100 per person.
X_3, X_{13}	Consumption fresh oranges in periods 1 and 2, respectively, tons per 10,000 persons.
X_4	Consumption processed oranges, apple crop year, tons per 1,000 persons.
X_5	Apple processors' stocks processed apple products September 1, tons per 10,000 persons.
X_6	Apple production "fresh regions," tons per 1,000 persons.
X_7	Apple production "processing regions," tons per 1,000 persons.
X_8	1 in 1952, 2 in 1953, ..., 20 in 1971.
X_9, X_{10}	Marketing margin indexes, fresh and processing apples, respectively, 1967=100, units of 10 percent.
X_{11}	Identically equal to Y_7.
X_{14}	Ratio of controlled atmosphere to regular storage stocks of fresh apples, November 30, percent.

A price-quantity scatter diagram for juice suggests that juice demand underwent a permanent shift to the right in 1962, perhaps reflecting a shift in consumer preferences. A dummy variable (X_1) is included to capture this shift. Because one group of states produce mainly for the fresh market and another mainly for the processing market, total production is divided into two parts (X_6 and X_7) designating the location of production. Linear trend variables are included in the fresh and canning allocation functions as surrogates for a number of influences, such as the rapid increases in costs associated with growing fruit for the fresh outlet and the growth of storage facilities, which have had the net effect of decreasing the allocation to the period 1 fresh market and increasing the allocation both to the canning and the period 2 fresh markets. Finally, fruit stored in controlled atmosphere facilities tends to maintain its quality better than regular storage fruit, and hence, commands a price premium; X_{14} in (9) is intended to capture the effect of increased controlled atmosphere storage on price.

Estimation

The instrumental variables used to estimate a particular structural equation may be defined as the predetermined variables in the equation and as computed values of the RHS endogenous variables (the $\hat{Y}$'s). In normal TSLS, the $\hat{Y}$'s are a function of all the predetermined variables in the system, but the number of observations and/or multicollinearity may preclude using the entire list of variables. A common practice has been to delete some predetermined variables from the first stage on an ad hoc basis. Fisher [3] suggests that the internal logic of the model should play a role in defining the instrumental variables, hence the name SOIV. However, like any IV estimator, the instruments must be defined so as to preserve the identification of the structural equation and the consistency of the estimator [see 4, pp. 20-22].

The first step in SOIV is to establish a preference ordering of variables in the model relative to the RHS current endogenous variables in each equation. To do this, the system of equations is written with each endogenous variable appearing only once on the LHS (somewhat analogous to a recursive order). Then, the ordering is established by using the relationships among endogenous and predetermined variables. For example, the endogenous variable Y_4 in equation (1) is a variable to be explained, and the explanatory variables in equation (4), where Y_4 is on the LHS, are first causal order variables since the model implies that they have a direct effect on Y_4.

Space limitations do not permit a detailed description of the establishment of the preference ordering of the variables [see 3]. The ordering, in essence, reflects the closeness of the relationship, implied by the structure, among predetermined variables and the RHS Y's. Preference orderings for several variants of the model of the apple industry are shown in Table 2.

Table 2.
Preference Orderings of Instruments
Relative to RHS Endogenous Variables[a]

Predetermined Variable	RHS Endogenous Variable					
	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6
X_1	7*	7*	2(3)	9*	4	4
X_2	1	1*	1*	4(4)	4	3*
X_3	2*	7	9*	5*(6)	4*	9
X_4	7*	7*	2(3)	9*	4	4*
X_5	7	2(3)	6(7)	5*(6)	4	4*
X_6	4	4	4(5)	1	1	1
X_7	4	4	4(5)	1	1	1
X_8	6*	6	8*	3	1	8
X_9	2*	7	9	5*(6)	4	9
X_{10}	7	2	6*(2)	5*(4)	4	4

[a] Where the preference orderings differed between the original and revised model specifications, the original preference ordering appears in parenthesis. The asterisk indicates that the variable was eliminated in the SOIV procedure. For instance, for Y_4, X_6 and X_7 have the first preference order, and X_1, X_3, X_4, X_5, X_9, and X_{10} were deleted in defining $\hat{Y}_4$.

In establishing the preference order for the predetermined variables, the entire list is retained. The intent, however, is to delete certain predetermined variables from the list in order to reduce multicollinearity and increase the degrees of freedom. Variables might be dropped on the basis of the preference ordering alone, but Fisher [4, p. 28] suggests the use of supplementary regressions to assist in dropping X's. The RHS endogenous variable is first regressed, if degrees of freedom permit, against all the X's. Commencing with the least preferred variable, a decision is made on whether to retain the variable by examining its t value.[3/] The process is repeated until the X's which remain have coefficients which meet the t criterion. The $\hat{Y}$'s obtained from the reduced list of X's may be viewed as the instrumental variables for the corresponding RHS Y's.

The procedure is rather mechanical, and indeed instruments could be constructed such that the necessary conditions for the equation to be identified and the estimator to be consistent are not met. If the predetermined variables which remain after the procedure has been used are insufficient for an equation to be identified, then in our judgment the original structure should be checked for specification error.[4/] Consistency of the SOIV estimator can be assured by following an approach suggested by Mitchell and Fisher [see 6, pp. 229-232].

Johnston [5, pp. 393-395] discusses the important choices that have to be made in using the TSLSPC estimator. In the present study, the principal components are based on all 10 predetermined variables in the model. Up to the first four principal components are used in estimating each equation. These accounted for 93 percent of the variation in the X's. When less than four principal components were used, the authors experimented with choosing those principal components which were least correlated with the predetermined variables in the equation.

The TSLSPC and SOIV estimates may be more precise than the TSLS estimates, but this need not be the case. The estimated variance of each regression coefficient is the product of the estimated variance of regression and an appropriate element from the principal diagonal of the inverse of a matrix of sums of squares and cross products [e.g., see 5, p. 384]. The variance of regression depends on both the correlation of the explanatory variables with the dependent variable and

---

[3/] Fisher [3, p. 268] suggests the use of R^2 corrected for degrees of freedom as the criterion for eliminating instruments. This is equivalent to dropping variables if $|t| < 1$, the criterion used here [see 1, pp. 91-92].

[4/] The structural equations can meet identification requirements in an accounting sense, but not in a real or causal sense. The SOIV method can be used as an empirical "test" of whether a model specification is consistent with identification requirements in a causal sense.

the intercorrelation of the explanatory variables [5, p. 60]. The RHS $\hat{Y}$'s defined by regular TSLS have a higher correlation with the LHS Y than the $\hat{Y}$'s defined by TSLSPC or SOIV, but the RHS $\hat{Y}$'s defined by TSLSPC and SOIV have smaller correlations with the RHS X's. Thus, the question of which estimator has the smallest variance of regression must be answered empirically. Likewise, the elements of the diagonal of the inverse matrix may, or may not, be smaller for SOIV and TSLSPC than for TSLS. Clearly, the multicollinearity related to predetermined variables within a particular equation is not reduced by TSLSPC or SOIV. Fisher [4, p. 29] also speculates that "SOIV... is equivalent to finding the optimal number of principal components in terms of ... consistency and efficiency."

Both SOIV and TSLSPC are computationally more expensive than TSLS or an ad hoc IV procedure. Hence, given the uncertainty about the improved precision of SOIV and TSLSPC, questions obviously exist about the relative benefits and costs of the alternatives. With this background, we turn to the empirical application.

Results

TSLS, TSLSPC, and SOIV all gave poor initial results for equations (1), (3), and (5). The signs associated with X_3 in (1), X_{10} in (3), and X_6 in (5) were opposite to those expected. Also, the OLS estimate of equation (9) resulted in the sign of the coefficient associated with X_9 being illogical. These results may be partly due to multicollinearity. The model was revised slightly in an attempt to overcome these problems: X_{10} was omitted from equation (3), X_6 from equation (5), and X_9 from equation (9). Attempts to revise the specification of equation (1), however, resulted in illogical relative magnitudes for the coefficients associated with Y_4 and Y_5.

It is not uncommon to revise a model specification to obtain "acceptable" estimates. For SOIV, however, revising the model can result in a large amount of additional computation. In the present case, the deletion of X_6 and X_{10} from equations (5) and (3), respectively, did not alter the set of predetermined variables in the system. Hence, the TSLS and TSLSPC estimates of the remaining equations were unaffected, but the deletion of X_{10} from equation (3) did affect the preference orderings of instruments relative to some RHS Y's for SOIV estimation.

The preference orderings of predetermined variables for the revised model specification are given in Table 2. Those instances where the preference orderings under the revised specification differ from the preference orderings under the original specification are indicated by including the latter in parentheses. Table 2 also indicates the results of the regression analysis used to delete X's.

The results for the revised model are shown in Table 3. OLS estimates are included for comparative purposes. In general, the results compare favorably with those from previous studies (e.g. [2, 7, 9]). Although the estimates of some coefficients are sensitive to the choice of estimation method, it is not immediately apparent from Table 3 that any one estimation method generally gave better results than others.

Of the simultaneous estimators, the estimated variance of regression is usually smallest for TSLS. In five of seven cases the estimated variance of regression for SOIV lies between the TSLS and TSLSPC estimates. As explained previously, TSLSPC and SOIV might provide smaller elements on the principal diagonal of the inverse of the sums of squares and cross products matrix. To facilitate comparison, these elements for the TSLSPC and SOIV estimates are divided by the corresponding elements for TSLS (Table 4). In all cases except one, the element from the inverse of the sums of squares and cross products matrix is smallest for TSLS. In most cases, the element for the SOIV estimate is only slightly bigger than the element for TSLS (the ratio is very close to one). On the other hand, the elements for TSLSPC are usually larger than the elements for TSLS and SOIV.

Based on the evidence (Tables 3 and 4), the performance of TSLSPC and SOIV relative to TSLS is disappointing. TSLS provides the largest ratios of the slope coefficients to their standard errors for 16 cases, TSLSPC for 11 cases, and SOIV for five cases. However, in 12 of the 16 cases where TSLS is largest, the SOIV ratios are similar in magnitude (implying that SOIV is more precise than TSLSPC in these cases). Those instances in which the TSLSPC estimate resulted in the largest ratios are generally attributable to a larger absolute value for the regression coefficient rather than a smaller standard error. In three of the five cases in which SOIV has the largest ratio, the absolute size of the estimated coefficient is similar to the TSLS estimate.

Also, based on logic, the SOIV estimates would seem preferable in a few instances. The price flexibilities of demand are shown in Table 5. In the case of the demand for fresh apples in period 1, the flexibilities computed from the SOIV coefficients are the most logical in terms of relative magnitudes and absolute sizes. Although the TSLSPC estimate of the cross flexibility with respect to oranges has the correct sign, the TSLSPC estimate of the cross flexibility with respect to canning apples exceeds the own flexibility, which is illogical. The TSLS estimate of the own flexibility implies a price elasticity of demand of at least -2.9, while the SOIV estimate implies a more reasonable lower limit of -1.2.[5]

_____
[5] The price elasticity of demand is at least as great as the reciprocal of the price flexibility [10, p. 357].

Table 3.
Estimates of Model

Equation		Net Regression Coefficients (and Ratio of Regression Coefficient to Standard Error)			
		OLS	TSLS	TSLSPC	SOIV
(1) Fresh Demand period 1, Y_1	Intercept	72.974 (0.987)	129.144 (1.258)	154.230	186.027 (1.405)
	Y_4	0.242 (0.251)	-0.805 (-0.475)	-0.394 (-0.204)	-1.911 (-0.836)
	Y_5	-0.493 (-0.570)	-0.670 (-0.708)	-1.682 (-1.222)	-0.802 (-0.747)
	X_2	8.518 (1.916)	6.037 (1.077)	6.441 (1.078)	3.448 (0.496)
	X_3	-0.056 (0.092)	0.355 (0.405)	-0.183 (-0.163)	0.809 (0.729)
	X_9	-5.997 (-0.826)	-4.606 (-0.595)	-3.789 (-0.474)	-3.207 (-0.364)
	s^2		186.86	197.19	234.12
(2) Canning Demand, Y_2	Intercept	271.748 (3.071)	300.973 (3.102)	357.073	316.725 (3.167)
	Y_4	-0.342 (-0.621)	-0.451 (-6.625)	-0.902 (-1.105)	-0.602 (-0.790)
	Y_5	-0.862 (-1.520)	-1.168 (-1.864)	-1.635 (-2.309)	-1.287 (-1.988)
	X_2	6.146 (2.944)	6.406 (2.756)	5.977 (2.392)	6.191 (2.602)
	X_5	-5.813 (-3.670)	-5.915 (-3.616)	-6.264 (-3.596)	-6.029 (-3.633)
	X_{10}	-20.997 (-2.559)	-22.565 (-2.698)	-23.737 (-2.676)	-22.707 (-2.689)
	s^2		83.37	93.07	84.93

Table 3 (continued)

Equation		Net Regression Coefficients (and Ratio of Regression Coefficient to Standard Error)			
		OLS	TSLS	TSLSPC[a/]	SOIV
(3) Juice Demand, Y_3	Intercept	31.603 (5.752)	32.671 (5.838)	33.439	32.129 (5.756)
	Y_6	-1.021 (-3.670)	-1.132 (-3.838)	-1.211 (-3.645)	-1.076 (-3.641)
	X_1	7.934 (1.974)	8.147 (2.014)	8.301 (2.027)	8.039 (1.995)
	X_2	1.861 (1.727)	2.004 (1.838)	2.107 (1.885)	1.932 (1.777)
	X_4	-0.277 (-0.627)	-0.302 (-0.679)	-0.320 (-0.710)	-0.289 (-0.653)
	s^2		18.81	19.19	18.66
(4) Fresh allocation period 1, Y_4	Intercept	-1.876 (-0.092)	6.806 (0.271)	1.384	8.933 (0.341)
	Y_1	0.362 (3.562)	0.256 (1.763)	0.329 (1.689)	0.244 (1.570)
	Y_2	-0.161 (-1.778)	-0.060 (-0.460)	-0.133 (-0.723)	-0.056 (-0.401)
	X_6	3.185 (2.365)	2.526 (1.557)	2.956 (1.527)	2.405 (1.427)
	X_7	2.642 (2.566)	2.615 (2.261)	2.606 (1.962)	2.553 (2.165)
	X_8	-2.532 (-9.551)	-2.302 (-6.430)	-2.459 (-5.309)	-2.273 (-5.982)
	s^2		7.42	6.80	7.53

Table 3 (continued)

Equation		Net Regression Coefficients (and Ratio of Regression Coefficient to Standard Error)			
		OLS	TSLS	TSLSPC[a]	SOIV
(5) Canning Allo-cation, Y_5	Intercept	15.873 (1.315)	18.496 (1.347)	7.128	16.746 (1.179)
	Y_1	-0.262 (-3.174)	-0.349 (-3.234)	-0.474 (-2.533)	-0.370 (-3.259)
	Y_2	0.430 (3.221)	0.581 (3.276)	0.652 (2.179)	0.590 (2.931)
	Y_3	-0.284 (-0.976)	-0.398 (-1.112)	-0.031 (-0.054)	-0.333 (-0.842)
	X_7	2.575 (2.221)	2.523 (1.917)	3.877 (1.728)	2.741 (2.014)
	X_8	1.302 (4.742)	1.555 (4.578)	1.520 (2.895)	1.547 (4.133)
	s^2		8.70	13.67	9.09
(6) Juice Allo-cation, Y_6	Intercept	-43.945 (-2.780)	-38.615 (-2.156)	-54.161	-36.888 (-2.008)
	Y_2	-0.156 (-1.411)	-0.236 (-1.807)	-0.111 (-0.725)	-0.240 (-1.734)
	Y_3	0.572 (3.626)	0.674 (3.996)	0.620 (3.493)	0.657 (3.768)
	X_6	2.616 (2.216)	2.347 (1.883)	2.995 (2.279)	2.288 (1.815)
	X_7	4.658 (4.359)	4.386 (3.812)	5.123 (4.134)	4.309 (3.689)
	s^2		10.12	10.26	10.12

-12-

Table 3 (continued)

Equation		Net Regression Coefficients (and Ratio of Regression Coefficient to Standard Error)			
		OLS	TSLS	TSLSPC[a/]	SOIV
(7) Fresh Allocation period 2, Y_7	Intercept	5.191 (0.280)	6.253 (0.282)	3.927	4.969 (0.213)
	Y_1	-0.055 (-0.600)	-0.020 (-0.153)	0.065 (0.322)	0.0002 (0.001)
	Y_2	-0.108 (-1.307)	-0.168 (-1.459)	-0.277 (-1.647)	-0.189 (-1.527)
	X_6	3.931 (3.205)	3.991 (2.783)	4.327 (2.195)	4.100 (2.732)
	X_7	1.266 (1.349)	1.070 (1.047)	0.857 (0.686)	1.057 (1.007)
	X_8	0.841 (3.483)	0.775 (2.450)	0.606 (1.266)	0.733 (2.167)
	s^2		5.80	7.26	5.97
(9) Fresh Demand period 2, Y_9	Intercept	237.338 (7.473)			
	X_{11}	-3.579 (-5.796)			
	X_{12}	2.389 (0.803)	[Not applicable]		
	X_{13}	-0.527 (-1.340)			
	X_{14}	1.274 (2.098)			

a/ The estimate of equation (1) is based on the second, third and fourth principal components and the estimate of equation (4) is based on the first, second and fourth principal components. The estimates of the other equations are based on the first four principal components.

Table 4.
Ratios of Elements From Diagonal of the
Inverse of the Sums of Squares and
Cross Products Matrix for TSLSPC
and SOIV to the Elements for TSLS

Equation	Estimator	Explanatory Variable				
		Y_4	Y_5	X_2	X_3	X_9
(1)	TSLSPC	1.227	2.000	1.075	1.561	1.011
	SOIV	1.448	1.021	1.226	1.268	1.035
		Y_4	Y_5	X_2	X_5	X_{10}
(2)	TSLSPC	1.161	1.149	1.035	1.015	1.007
	SOIV	1.097	1.042	1.028	1.009	1.000
		Y_6	X_1	X_2	X_4	
(3)	TSLSPC	1.261	1.005	1.030	1.009	
	SOIV	1.022	1.000	1.002	1.000	
		Y_1	Y_2	Y_6	Y_7	X_8
(4)	TSLSPC	1.931	2.174	1.556	1.441	1.827
	SOIV	1.103	1.130	1.063	1.025	1.110
		Y_1	Y_2	Y_3	X_7	X_8
(5)	TSLSPC	2.000	1.806	1.655	1.849	1.530
	SOIV	1.077	1.278	1.162	1.023	1.167
		Y_2	Y_3	X_6	X_7	
(6)	TSLSPC	1.647	1.107	1.096	1.144	
	SOIV	1.118	1.071	1.022	1.031	
		Y_1	Y_2	X_6	X_7	X_8
(7)	TSLSPC	1.862	1.393	1.509	1.191	1.827
	SOIV	1.103	0.929	1.063	1.025	1.110

Table 5.
Price Flexibilities of Demand Computed at
the Point of Means

Price Flexibility	Estimation Technique			
	OLS	TSLS	TSLSPC	SOIV
Fresh apples period 1				
Own	a/	-0.340	-0.166	-0.805
Cross (canning apples)	-0.160	-0.217	-0.545	-0.260
Cross (fresh oranges)	-0.019	a/	-0.061	a/
Income	0.882	0.625	0.667	0.357
Canning apples				
Own	-0.576	-0.781	-1.093	-0.860
Cross (fresh apples)	-0.297	-0.393	-0.784	-0.524
Income	1.313	1.369	1.277	1.323
Juice apples				
Own	-0.702	-0.778	-0.833	-0.740
Cross (proc. oranges)	-0.197	-0.215	-0.228	-0.206
Income	0.710	0.764	0.804	0.737
Fresh apples period 2				
Own	-1.194	Equation not estimated		
Cross (fresh oranges)	-0.243	by simultaneous equation		
Income	0.234	estimators.		

a/ Indicates that the estimated coefficient was positive.

If the SOIV estimate of the own-price flexibility for fresh apples
in period 1 is accepted as the most reasonable, then the symmetry condi-
tion for price flexibilities suggests rejecting the TSLS estimate of the
canning demand function in favor of either the SOIV or TSLSPC estimate.[6]
The authors feel more confident about the SOIV estimate of the canning
demand function than the TSLSPC estimate. The latter estimate was highly
sensitive to the choices made concerning the number and mix of principal
components used in the estimation process (Table 6). Adding principal

---

[6] The approximation to the symmetry relation used in this study is F_{ij}
= (E_i/E_j) (F_{ji}), where F_{ji} = price flexibility of j with respect to
the quantity of i and E_i, E_j are the expenditures on i and j. The
expenditures were taken as the mean value of P_iQ_i for the crop years
1952 through 1971 [for a discussion of the symmetry relation, see 11,
p. 84].

Table 6.
TSLSPC Estimates of the Canning Demand
Function Using Different
Principal Components[a]

Princi-pal Compon-ents	Estimated Regression Coefficients					
	Intercept	Y_4	Y_5	X_2	X_5	X_{10}
1 2 3 -	483.580	$-2.126^{[b]}$	-2.584	$4.198^{[b]}$	-7.185	-24.775
- 2 3 4	365.343	$-0.630^{[b]}$	-1.874	7.224	-6.099	-26.505
1 2 3 4	357.073	$-0.902^{[b]}$	-1.635	5.977	-6.264	-23.737

[a] LHS variable is Y_2.

[b] Absolute size of regression coefficient is less than 1.5 times its standard error.

components appears to increase the coefficient of X_2 and to reduce the absolute size of the other coefficients, and interestingly the SOIV estimate of the coefficient of X_2 is larger and the other coefficients smaller than those obtained using four principal components. This result is consistent with Fisher's view that SOIV estimates should be similar to TSLSPC with the optimal number of principal components.

While the SOIV estimates of the first two equations seem preferable, no one estimator gave better results than the others for the remaining equations. Apart from the estimated coefficient associated with Y_1 in equation (7),[7] the estimates of equations (3) through (7) are not very sensitive to the choice of estimator.

Conclusions

It is, of course, difficult to reach general conclusions from one empirical application. However, the results do add to the relatively small number of empirical contrasts of TSLSPC, SOIV, and TSLS in the agricultural economics literature. The TSLSPC and SOIV estimators are

---

[7] As mentioned, rationalizations exist for both a negative or a positive coefficient for Y_1 in equation (7). The various estimates of this equation suggest that in fact Y_7 may not depend upon Y_1 and Y_2. This is further supported by the following quite acceptable estimate of the equation by deleting Y_1 and Y_2 and using OLS:

$$\hat{Y}_7 = 25.222 + 5.338X_6 + 2.601X_7 + 0.643X_8.$$
$$\phantom{\hat{Y}_7 = } (-3.073) \quad (6.013) \quad (3.364) \quad (6.090)$$

$$\overline{R}^2 = .86.$$

computationally more expensive than TSLS, and for our model of the U.S. apple industry, TSLSPC and SOIV generally did not provide more precise estimates of the parameters. However, based on the magnitudes of the coefficients, the SOIV estimates seemed preferable to the alternatives for equations (1) and (2), but not for the other five structural equations. Thus, our experience in using SOIV and TSLSPC in a moderate-sized model is rather discouraging. These estimators still may be useful, of course, when models are larger or multicollinearity more prevalent than it is in our model.

In choosing between TSLSPC or SOIV, researchers should give attention to the sensitivity of the TSLSPC results to the number and mix of principal components. The TSLSPC estimates can be highly sensitive to these choices, and SOIV estimates may more nearly approximate the results from using the optimal, but unknown, number of principal components. Also, in the application reported here, the SOIV estimates of the parameters tended to be more precise than the TSLSPC estimates. On the other hand, the SOIV estimator can be especially costly to recompute in the face of model revisions. However, on balance, our experience suggests that SOIV may be preferable to TSLSPC when multicollinearity and lack of degrees of freedom are problems.

References

1. Aigner, D. J. Basic Econometrics. Englewood Cliffs, N.J., Prentice-Hall, Inc., 1971.

2. Brandow, G. E. A Statistical Analysis of Apple Supply and Demand. Penn. State Univ., Agric. Econ. and Rur. Soc. No. 2., 1956.

3. Fisher, F. M. "The Choice of Instrumental Variables in the Estimation of Economy-Wide Models", Intern. Econ. Rev., 6(1965):245-274.

4. Fisher, F. M. Simultaneous Equations Estimation: The State of the Art. Arlington, Virginia, Institute for Defense Analysis Economic Papers, July 1970.

5. Johnston, J. Econometric Methods. 2nd ed., New York, McGraw-Hill Book Co., 1972.

6. Mitchell, B. M. and F. M. Fisher. "The Choice of Instrumental Variables in the Estimation of Economy-Wide Econometric Models: Some Further Thoughts", Intern. Econ. Rev., 11(1970):226-234.

7. Pasour, E. C. and R. L. Gustafson. Intraseasonal Supply and Demand Functions for Apples. Michigan State Res. Bul. 10, 1966.

8. Piggott, R. R. An Econometric Model of the U.S. Apple Industry: Applications to Forecasting and Analyzing Policy Issues. Unpublished Ph.D. thesis, Cornell University, Jan. 1974. (Also see "Potential Gains from Controlling Distribution of the United States Apple Crop", Search: Agriculture, Cornell Univ. Agr. Exp. Sta., Vol. 6, No. 2, 1976.)

9. Tomek, W. G. Apples in the U.S.: Farm Prices and Uses, 1947-75. Cornell Agr. Exp. Sta. Bul. 1022, July 1968.

10. Tomek, W. G. and K. L. Robinson. Agricultural Product Prices. Ithaca, Cornell University Press, 1972.

11. Waugh, F. V. Demand and Price Analysis. USDA Tech. Bul. 1316, Nov. 1964.