Simultaneous Equation Market Models: A New Approach to the Problem of Multicollinearity

by

Cheng-few Lee*

and

Joseph D. Vinso*

Working Paper No. 11-76

RODNEY L. WHITE CENTER
FOR FINANCIAL RESEARCH

University of Pennsylvania

The Wharton School

Philadelphia, Pa. 19174

Simultaneous Equation Market Models: A New Approach to the Problem of Multicollinearity

I. Introduction

The security market model of Sharpe [1964], Lintner [1965] and Mossin [1966] relates returns on an asset to the variations in a market index and is known as the capital asset pricing model (CAPM). While this approach has been the foundation for much of the research in finance, it has been suggested that changes in the single-index model could be made so as to improve the power of the model. Cohen and Pogue [1967] have suggested that an industry factor (I_t) could be included in Sharpe's model to increase the explanatory power of that model. Likewise, Beaver [1972] and Downes and Dyckman [1973] argue that certain types of accounting information are taken into account in security pricing and, thus, should be included in a model of capital asset pricing.

These studies culminated in the recent work by Simkowitz and Logue [1973] (S-L) who derive a simultaneous equation security market model. This model contains industry data as well as certaim accounting information as follows:

$$
\begin{aligned}
\text{(i)}\quad R_{1t} &= \alpha_1 - \gamma_{12} R_{2t} - \gamma_{13} R_{3t} - \cdots\cdots - \gamma_{1n} R_{nt} + b_{11} X_{11t} + b_{12} X_{12t} \\
&\quad + b_{13} X_{13t} + \beta_{1m} R_{mt} + E_{1t} \\
\vdots\quad & \quad\vdots \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (1) \\
\text{(vii)}\quad R_{nt} &= \alpha_n - \gamma_{n1} R_{1t} - \gamma_{n2} R_{2t} - \cdots\cdots - \gamma_{nn-1} R_{n-1} + b_{n1} X_{n1t} + b_{n2} X_{n2t} \\
&\quad + b_{n3} X_{n3t} + \beta_{nm} R_{mt} + E_{nt}
\end{aligned}
$$

Where R_{jt} = the return on the j^{th} security over time interval t in a group classified by a reasonable classification scheme, (j = 1, 2, ..., n)

R_{mt} = the return on a market index over time interval t,

X_{j1t} = the profitability index of j^{th} firm over time interval t, (j = 1, 2, ..., n)

X_{j2t} = the leverage index of j^{th} firm over time period t, (j = 1, 2, ..., n)

X_{j3t} = the dividend policy index of j^{th} firm over time period t, (j = 1, 2, ..., n)

b_{jk} = the coefficient of the k^{th} firm related variable in the j^{th} equation, (k = 1, 2, 3)

γ_{ji} = the coefficient of the i^{th} endogenous variable in j^{th} equation, (i = 1, 2, ..., n; j = 1, 2, ..., n)

β_{mj} = the coefficient of market rate of return in the j^{th} equation,

E_{jt} = the disturbance term for j^{th} equation.

Equation (1) is a simultaneous equation system with n endogenous and 3n + 1 exogenous variables.[1] It can easily be shown that every equation in the system is over-identified.

S-L point out that the standard CAPM is expressed by equation (1) with the following constraints:

$$
\begin{aligned}
&\text{(i)}\quad \begin{pmatrix} E_1 \\ E_2 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ E \end{pmatrix}_{nx1} (E_1, E_2, \ldots, E_n)_{1xn} = \begin{pmatrix} \sigma^2 & 0000\ldots\ldots 0 \\ \cdot\ \ \sigma^2 & 0 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ 0 & \cdot \\ 0 & 00\ldots\ldots\ldots\sigma^2 \end{pmatrix}_{nxn} \\
&\text{(ii)}\quad b_{jk} = 0 \qquad (j = 1, 2, \ldots, n) \\
&\qquad\qquad\qquad\quad\ (k = 1, 2, 3) \\
&\text{(iii)}\quad \gamma_{ji} = 0 \qquad \text{if } i \neq j \qquad (i = 1, 2, \ldots, n) \\
&\qquad\quad\ \ = 1 \qquad \text{if } i = j \qquad (j = 1, 2, \ldots, n)
\end{aligned}
\tag{2}
$$

After imposing these constraints on equation (1), S-L show that Sharpe's model can be defined as:

$$
R_{jt} = \alpha_j + \beta_j R_{mt} + E_{jt} \qquad (j = 1, 2, \ldots, n) \tag{3}
$$

where α_j and β_j are regression parameters and E_{jt} is the disturbance term. While the parameters in equation (3) are estimated using ordinary least squares (OLS), S-L use two stage least squares (2SLS) to estimate the parameters of the simultaneous equation system shown in equation (1).

Theoretically the S-L model specified in equation (1) has explicitly taken the interdependent relationship among the security returns within an industry into account. It is well-known, though, that the results associated with the 2SLS estimation procedure can be subject to the problems of multicollinearity and prediction errors. Aber [1973], for example, has pointed out that multi-index models are generally complicated by the problem of multicollinearity. In using 2SLS, S-L regressed R_j's on 22 (or 21) exogenous variables to obtain the estimated R_j's. They then used the estimated R_j's as regressors in the second stage regression. Their results indicate that market rate of return, R_m, is the most important explanatory variable for each first stage regression. However, R_m also appears in each second stage regression within the system. Since the correlation between R_m and the estimated R_j is very strong, it is possible that the problem of multicollinearity faced by S-L is not negligible.[2]

S-L have classified the interdependent relationship of security markets into two different cases, i.e., (a) only condition (2i) does not hold, and (b) all the conditions do not hold. They argue that case (a) belongs to Zellner's [1962] seemingly unrelated regression (SUR) problem. Generalized least squares (GLS) can then be employed to replace OLS to improve the efficiency of the estimated systematic risk. For the second case, they have employed the 2SLS to perform a robust test of the basic assumption of Sharpe's model and to examine the importance of market rate of return.

Because of the potential for multicollinearity, one must investigate whether the GLS can be employed in the first case to improve the efficiency of estimated systematic risk. Following Zellner [1962], the estimated systematic risk obtained by the GLS are:

$$\tilde{\beta}_m = (\gamma_m' \ \Omega^{-1} \ \gamma_m)^{-1} \ (\gamma_m' \ \Omega^{-1} \ \gamma_j), \tag{4}$$

where $\gamma_m' = (R_m, R_m, \ldots, R_m)_{1xn}$,

$\gamma_j' = (R_1, R_2, \ldots, R_n)_{1xn}$,

$$\Omega = \begin{pmatrix} E_1 \\ E_2 \\ \cdot \\ \cdot \\ \cdot \\ E_n \end{pmatrix} (E_1, E_2, \ldots, E_n) = \begin{pmatrix} S_{11} & S_{12} & \cdots & S_{1n} \\ S_{21} & S_{22} & & S_{2n} \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ S_{n1} & \cdots & \cdots & S_{nn} \end{pmatrix}$$

Kmenta [1971] shows that the GLS estimator, $\tilde{\beta}_m$, is identical to the OLS estimator of systematic risk, $\hat{\beta}_m$, as indicated in equation (5).

$$\hat{\beta}_m = (\gamma_m' \ \gamma_m)^{-1} \ (\gamma_m' \ \gamma_j) \tag{5}$$

In general, the relative efficiency between the $\tilde{\beta}_m$ and $\hat{\beta}_m$ is an increasing function of the degree of relationship among the residuals and a decreasing function of the degree of relationship among explanatory variables. However, there exists no gain in efficiency of the GLS estimator over the OLS estimator when the SUR involve exactly the same explanatory variables. This implies that the GLS cannot be employed to improve the efficiency of estimated systematic risk for the first case. The fact is not realized by S-L. Incidentally, if only the restriction (2iii) is held for equation (1), then Zellner's SUR may be used to improve the efficiency of the estimators of regression coefficients. Since the relationship among the residuals of the firms in a same industry will not be trivial, the gain in efficiency of the GLS estimator will be essentially dependent upon the relationship among the firm-related variables in the industry.

Besides the problem with multicollinearity, the 2SLS is also susceptible to prediction error. If the adjusted coefficient of determination ($\bar{R}^2$) for the first stage regression is low, then the estimated R_j's contain some prediction errors (or measurement errors in a wide sense). In such a case, the efficiency of the second stage estimators will be affected (for a complete of this problem see, for example, Cochran 1972).

Because of the multicollinearity and prediction error problems with 2SLS, the estimation procedure for the simultaneous equation system developed by S-L may be inappropriate. Thus, the results developed in their paper must be viewed with skepticism. The purpose of this paper is to develop a security market model which does not suffer these problems yet retains the desirable features of the simultaneous equation method by combining Sharpe's model and the S-L model.

First the relationship between the single index and multi-index security model is explained followed by the development of a model which combines Sharpe's and the S-L model. Next, the usefulness of the new model is demonstrated using data from several oil companies. Finally, the results of the paper are summarized.

II. Combining Single and Multi-Index Models: A Synthesis Approach

Cohen and Pogue [1967], as previously discussed have suggested that an industry factor can be included in equation (3) to increase the explanatory power of Sharpe's model. After including an industry factor (I_t) in equation (3), we have

$$R_{jt} = \alpha_j' + \beta_j' R_{mt} + \delta_j I_t + E_{jt}' \tag{6}$$

where α_j', β_j' and δ_j are the coefficients and E_{jt}' is the disturbance term. The explanatory power of equation (6) is generally larger than that of equation (3).[3] It also is known that the summation of β_j' and δ_j is approximately equal to β_j.

Following Theil's [1971] specification analysis technique, we can derive the relationship between the systematic risk obtained from Sharpe's model and the coefficients obtained from the S-L model. For simplicity, S-L's seven-equation simultaneous equation system is employed to demonstrate that the systematic risk associated with Sharpe's model can be written in terms of the regression parameters of S-L model as:

$$\beta_j = \gamma_{j1}D_1 + \gamma_{j2}D_2 + \ldots + \gamma_{j7}D_7 + b_{j1}D_8 + b_{j2}D_9 + b_{j3}D_{10} + \beta_{mj} \quad (j = 1, 2, \ldots, 7 \text{ and } \gamma_{jj} = 0) \tag{7}$$

where D_1, D_2, ... and D_{10} represent the auxiliary regression coefficients of regressing each explanatory variable of S-L model on R_m respectively. Actually, D_1, D_2, D_3, D_4, D_5, D_6 and D_7 are the systematic risks obtained from Sharpe's model.

Since firm related variables do not represent common factors in the security market, the relationship between the firm related variables and market rate of return is expected to be insignificant; therefore, equation (7) can approximately be rewritten as:

$$\beta_j \doteq \gamma_{j1}D_1 + \gamma_{j2}D_2 + \ldots + \gamma_{j7}D_7 + \beta_{jm} \quad (j = 1, 2, \ldots, 7 \text{ and } \gamma_{jj} = 0) \tag{8}$$

Expression (8) implies that the standard systematic risk (β_j) is a weighted average of the coefficients of all endogenous variables and the coefficient of R_m in S-L model. Statistically, the existence of equation (8) depends upon whether the auxiliary regression coefficients, D_8, D_9 and D_{10} are insignificant. Essentially that is an empirical question but if D_8, D_9 and D_{10} are insignificant, then the estimated systematic risk associated with the market index obtained from the new model defined in equation (9) will not be significantly different from that obtained from the right hand side of equation (8), and it will also not be significantly different from the estimated systematic risk obtained from Sharpe's model.

Thus, by using this result and imposing condition (2iii) on equation (1) the following obtains:

$$R_{jt} = \alpha_j'' + \beta_j'' R_{mt} + b_{j1}X_{j1t} + b_{j2}X_{j2t} + b_{j3}X_{j3t} + \varepsilon_{jt} \quad (j = 1, 2, \ldots, n) \tag{9}$$

where α_j'', β_j'', b_{j1}, b_{j2}, and b_{j3} are regression parameters and ε_{jt} is a disturbance.

This new model defined in equation (9) is obtained by assuming that rates of return on all other securities might be ommitted from equation (1). It also is a model formulated by integrating Sharpe's model with the S-L model. There are two advantages to this model. First, the explanatory power of this model is generally larger than that of Sharpe's model; secondly, this model allows SUR method to be applied to pool the cross-section and the time-series data to improve the efficiency of some estimators.[4] This new model synthesizes advantages of both Sharpe's model and S-L model. This "synthesis model" can simultaneously be used to obtain the systematic risk associated with Sharpe model and to incorporate the information of firm related variables into capital asset pricing. It also avoids some problems associated with specification and estimation of the S-L model, particularly, those involving multicollinearity.

III. Some Empirical Results with the Synthesis Model

To demonstrate the usefulness of the synthesis model relative to Sharpe's and the S-L model as well as to illustrate the problems with the S-L formulation, annual data of stock price and firm related variables from the period 1945-1973 for seven oil companies are used to calculate the related rates of return, the profitability index, the leverage index and the dividend policy index.[5] The appropriate rates of return for each company are adjusted for dividends and stock splits. The annual Standard and Poor index (S & P) with dividends is used to calculate the annual rate of return on the market. Using such a long time period raises the question of stationarity as parameter estimation assumes a stationary distribution. Tests using the Box-Pierce Q-statistic, however, show that the hypothesis that the time series is white noise cannot be rejected.[6]

First, to investigate the validity of the S-L model, a traditional 2SLS method is used to estimate the simultaneous relationship of security returns for the seven oil companies. These results are listed in Table I. Then, Sharpe's model specified in equation (3) is used to calculate systematic risks for seven

TABLE I

2SLS Structure Estimates of Oil Industry; R_m Included
(t-values appear in parentheses beneath coefficient)

Dependent Variables (R_i)**	R_1	R_2	R_3	R_4	R_5	R_6	R_7
Endogenous Variables							
R_1		4.3251	-.1172	.6199*	-.3461	-.0323	-.2089
	1.0	(1.3504)	(.4711)	(3.5029)	(.8766)	(.1506)	(1.0269)
R_2	.1034	-.1228	.1640	.7604*	-.0921	-.1893	
	(.4875)	1.0	(.4595)	(.7610)	(2.9124)	(.5674)	(1.0468)
R_3	-.0898	.2105		-.0750	-.2341	.2934	.0404
	(.4203)	(.7578)	1.0	(.4101)	(.6692)	(1.6715)	(.2134)
R_4	1.0012*	-.0504	.1120		.2702	.3970	.6824*
	(3.4242)	(.1113)	(.3054)	1.0	(.4267)	(1.3637)	(2.5069)
R_5	.0529	.4704*	.2552	-.1723	.1805	-.0156	
	(.2930)	(2.8126)	(.1531)	(1.0571)	1.0	(1.4404)	(.1105)
R_6	-.1792	.5125	.2818	1.0234*	.1824		
	(.5281)	0	(1.6127)	(1.0447)	(1.8031)	1.0	(.6254)
R_7	-.1618	-.3657		.1519	.1945	-.1330	
	(.5419)	(1.0446)	0	(.6284)	(.3923)	(.4571)	1.0
Exogenous Variables †							
Constant	-.1168	-.2300	-.2781	.3241	.1825	-.0529	-.0025
X_{j1}	1.4595	3.6608	2.8658*	-5.4350*	-3.2714	-.1330	-.4648
	(.4939)	(1.0009)	(2.0256)	(1.8298)	(.9801)	(.0990)	(.2658)
X_{j2}	1.6809	-1.0986	2.3367*	.8283	2.4936*	.8941	1.1500*
	(1.2933)	(1.3711)	(2.6637)	(.8969)	(2.1135)	(1.2919)	(2.3707)
X_{j3}	0	-17.5875	-6.7611	30.5213*	1.4031	59.5394*	2.5121
		(1.0725)	(.7694)	(1.8208)	(1.0286)	(1.8322)	(.4229)
R_m	.2486	.6470*	-.2720	.2774	.1446	.6498*	
	(.6626)	0	(2.2035)	(.7610)	(.4428)	(.4362)	(2.3336)
$\bar{R}^2$ ††	.4576	.3934	.5982	.6341	.5500	.559	.6206

NOTES: † subscript j, j = 1, ..., 7 denotes variables in second stage of relevant equation.
†† second stage estimates for the single equation.
* denotes significant at .10 level of significance or better for two-tailed test.
** see Table II for firm identification of the various R_i's.

oil companies which are shown in Table II. From Tables I and II, it is found that the adjusted coefficients of determination ($\bar{R}^2$) of S-L model are consistently higher than those of Sharpe's model. This implies that the explanatory power of S-L model is higher than that of Sharpe's model. However, it is found that the market rate of return is always the most important exogenous variable in estimating the endogenous variables. In other words, the estimated endogenous variables used in the second stage are highly correlated with the market rate of return which appears in every second stage regression. These results imply that indeed the methodology developed by S-L can be subject to the multicollinearity associated with the 2SLS estimation method. It should be noted that in Table I, R_m is significant at the 10% level in two regressions. However, if the modified 2SLS developed by Klein and Nakamura [1962] is used to estimate the simultaneous relationship, there exist five regression coefficients associated with R_m which are significant at the 10% level. Thus, R_m still plays a relatively important role in the S-L model if the multicollinearity is reduced to a manageable level.[7,8]

Before testing the validity of the synthesis model, it is necessary to determine whether the assumption holds that the estimated β''_j using the right hand side of equation (8) is not significantly different from the $\hat{\beta}_j$ estimated by either the Sharpe model or the right hand side of equation (7). S-L's empirical results of both oil and drug industries are used to test if this proposition holds more generally. Table III shows the estimated $\hat{\beta}_j$ and $\hat{\beta}''_j$ for each of the firms. Using a t-test it can easily be shown that $\hat{\beta}''_j$ for each firm estimated from equation (8) is not significantly different from $\hat{\beta}_j$ determined by equation (7) at the 95% level; therefore, the hypothesis that $\hat{\beta}_j$ and $\hat{\beta}''_j$ are not significantly different cannot be rejected and equation (9) can be used with confidence.[9]

To test the validity of the synthesis model specified in equation (9), first OLS is used to estimate the necessary parameters of seven oil companies. Residuals from these equations are used to test the degree of interrelationship

TABLE II

OLS Parameter Estimates of Oil Industry - Sharpe Model†

		α	β	$\bar{R}^2$
R_1	Imperial Oil	.04421	.7421 (2.6699)	.1851
R_2	Phillips Petroleum	-.0141	.6578 (1.6758)	.0626
R_3	Shell Oil	-.0043	1.2353 (4.2429)	.3864
R_4	S. O. of IN	.0366	.6869 (2.8619)	.2082
R_5	S. O. of OH	-.0161	.8720 (1.8160)	.0782
R_6	Sun Oil	.0049	.6240 (2.7130)	.1906
R_7	Union Oil of CA	.0058	1.0228 (4.6704)	.4353

† - t-values appear in parentheses beneath the corresponding coefficients

TABLE III

SYSTEMATIC RISKS FOR DRUG AND OIL INDUSTRIES

Industries	(1) $\hat{\beta}_j$*	(2) $\hat{\beta}_j''$**
Drug Industry		
American Home Product	1.0105	1.0033
	(.1579)	(.1350)
Bristol Myers	1.0081	.9856
	(1.9968)	(.1812)
Gillete	.8639	.8909
	(.1628	(.2013)
Johnson & Johnson	.8281	.9881
	(.1843)	(.2021)
Merck	.6228	.7057
	(.1591)	(.1602)
G. D. Searle	1.2375	1.2617
	(.2570)	(.2891)
Smith, Kline & French	1.2248	1.0715
	(.2088)	(.1856)
Oil Industry		
Imperial Oil	.7905	.7454
	(.1460)	(.1522)
Phillips Petroleum	.8011	.7893
	(.1300)	(.1588)
Shell Oil	.7384	.6982
	(.1249)	(.1112)
S. O. of IN	1.0650	.9891
	(.2315)	(.2481)
S. O. of OH	.7880	.8478
	(.1851)	(.2055)
Sun Oil	.5421	.5587
	(.1745)	(.1735)
Union Oil of CA	1.0521	1.1134
	(.1643)	(.1533)

Remarks: standard errors appear in parentheses beneath the corresponding coefficients.
*estimated by Sharpe's model
**estimated using the right hand side of equation (8)

among these oil companies. The residual correlation coefficient matrix for these seven companies (shown in table IV) indicates that there exist 10 residual correlation coefficients significantly different from zero at .05 significance level. This implies that SUR estimation method can be used to improve the efficiency of some estimators. The SUR estimators are obtained using equation (4). Results of both OLS and SUR estimates of these seven oil companies are listed in Table V. Five of the seven $\bar{R}^2$ associated with OLS estimates of the synthesis model shown in Table V are higher than those of Sharpe model shown in Table II. As the SUR estimation method is applied to the synthesis model, the efficiency of some estimators for Shell Oil, Standard Oil of Indiana, Standard Oil of Ohio, Sun Oil and Union Oil of California are improved.[10]

Results of the SUR estimation method have a great deal of intuitive appeal. For example, Imperial Oil which is a Canadian firm shows the lowest correlation with other firms as might be expected. It is also found in Table IV that the residuals of Phillips Petroleum are highly correlated with those of Standard Oil of Ohio and Sun Oil. However, the SUR estimation method does not improve the efficiency of estimators for Phillips Petroleum. One possible reason is that the financial management policies of this company may be highly correlated with those of other companies in the oil industry. As the explanatory variables of a regression become more similar to those of other regressions in the same industry, the gain of SUR estimation method will be getting smaller. It is found that systematic risk estimates from the synthesis model are not significantly different from those obtianed from Sharpe's model.

While these results are interesting, they can be viewed with confidence only if the assumptions of the regression model are fulfilled. Besides the stationarity assumption previously discussed, the homoscedasticity assumption of the regression residuals is an additional condition required to ascertain the stability of the estimated systematic risk as discussed by Blume [1971]. It is even more important

TABLE IV

Residual Correlation Coefficient Matrix

	R_1	R_2	R_3	R_4	R_5	R_6	R_7
R_1	1.0000	.1725	.1687	.4422*	.0571	.1129	.1450
R_2		1.0000	.2062	.2312	.7487*	.4420*	-.0770
R_3			1.0000	.1634	.3542*	.5748*	.2183
R_4				1.0000	.1789	.3645*	.3329*
R_5					1.0000	.6697*	.6234*
R_6						1.0000	.3154*
R_7							1.0000

*Denotes significantly different from zero at .05 level of significance.

TABLE V

OLS AND SUR Estimates of Oil Industry - Synthesis Model†

		α_j	β_j	b_{j1}	b_{j2}	b_{j3}	$\bar{R}^2$
R_1	OLS	-.4479 (1.6480)	.7602 (2.4770)**	6.043 (1.7270)	.4582 (.2932)	-2.0840 (-.1749)	.2202
	SUR	-.2861 (1.1890)	.7663 (2.5560)**	4.2020 (1.3600)	.8655 (.6268)	-6.9720 (-.6648)	
R_2	OLS	-.8203 (-.2680)	.5413 (1.3140)	1.0680 (.2359)	-1.2170 (-1.2180)	5.2960 (.2725)	.0248
	SUR	-.1600 (-.8001)	.5822 (1.4340)	2.4550 (.8517)	-.8193 (-1.2710)	-8.509 (.6768)	
R_3	OLS	-.2251 (-1.7230)	1.1200 (3.9800)**	2.0550 (1.4220)	2.5110 (2.1510)**	-.5165 (-.0677)	.4540
	SUR	-.3003 (2.633)**	1.0880 (3.8870)**	2.9140 (2.393)**	2.665 (2.812)**	-2.532 (-.4126)	
R_4	OLS	-.1550 (-.8648)	.7283 (3.0540)**	2.1070 (.7852)	-.2395 (-.2707)	25.2900 (1.8170)*	.3040
	SUR	-.0538 (-.3519)	.6458 (2.7530)**	-.9832 (-.4340)	.4262 (.5884)	26.3500 (2.337)**	
R_5	OLS	-.0029 (-.0111)	1.1280 (2.0820)**	-.7520 (-.2295)	1.8370 (1.3410)	15.600 (.8942)	.0558
	SUR	-.0047 (-.0271)	1.1640 (2.3090)**	-.8607 (-.4143)	1.8860 (2.4970)**	.9103 (.9555)	
R_6	OLS	-.1472 (1.4350)	.7098 (3.0960)**	.6888 (.5501)	.9155 (1.3050)	62.2000 (1.7140)	.2755
	SUR	-.1509 (1.9960)*	.7070 (3.1751)**	.8508 (1.0101)	.9016 (2.0230)**	54.1100 (2.4020)**	
R_7	OLS	-.1140 (-.9173)	1.0141 (4.7590)**	1.347 (.7899)	.6427 (1.2400)	4.5300 (.7579)	.4693
	SUR	-.0661 (-.5823)	1.0190 (4.6800)**	.6500 (.4206)	.9632 (2.0890)**	3.9660 (.7431)	

† - t-values appear in parentheses beneath the corresponding coefficients.

*denotes significant at .10 level of significance or better for two-tailed test.

**denotes significant at .05 level of significance or better for two-tailed test.

to investigate this assumption in light of the recent work by Rogalski and Vinso 1975 who found that the OLS estimation of the CAPM for over 25% of all securities show heterosecdasticity. To test for homoscedasticity of the regression residuals for each equation associated with the synthesis model, the Goldfield-Quandt [1965] test is used. To test whether the variance-covariance matrix obtained for the SUR equation system is stable over time, Anderson's 1958, Chapter 10] approximate χ^2 statistic is used. The results show that the assumption of homoscedasticity cannot be rejected at the .05 level of significance for any firm except Imperial Oil. Likewise, the assumption of a constant covariance matrix cannot be rejected either at the .05 significance level. Another advantage of the synthesis model concerns the problem of sample size. While the S-L model has explicitly specified the full structural simultaneous relationship of capital asset pricing for a particular industry, the multicollinearity problem explored here generally makes the statistical results of S-L model become less meaningful. As the number of equations in the system becomes larger, the number of regressors in the first stage of 2SLS generally is getting too large to be handled.[11] Since the regressor of the synthesis model is not affected by the number of equations, the problem of undersized sample does not exist in the synthesis model.

Now that the validity of the synthesis model has been shown, it would be of interest to investigate the importance of the three firm related variables used by Simkowitz and Logue [1973] in capital asset pricing. They have shown that the roles played by three firm related variables are to identify the simultaneous equation system of security market and to improve the explanatory power of the diagonal security market model. These same firm related variables also are explicitly included in the synthesis model indicated in equation (9). Using the SUR estimates of the synthesis model, the importance of these firm related variables in the return generation process can be analyzed. The profitability index is significant in explaining the rates of return of Shell Oil; the dividend policy index is significant in explaining the rates of return of Standard Oil of Indiana, Standard Oil of Ohio and Sun Oil; and the leverage index is significant in explaining the rates of return of Shell Oil, Standard Oil of Ohio, Sun Oil and Union Oil of California. These results imply that both leverage index and dividend policy index can be additional important factors in capital asset pricing. From a financial management viewpoint both leverage and dividend policies are unique factors of an industry so the market index itself can hardly be used to take care of the change of these two policies associated with a particular industry.

Thus the synthesis model is formulated by introducing the accounting information - the profitability index, the leverage index and the dividend policy index into Sharpe's model. It has explicitly taken into account the arguments on the possible impacts of accounting information on the behavior of security price. This multi-index model differs with other multi-index models from several aspects. First, the additional indices employed in the synthesis model are the accounting information of an individual firm rather than general economic activity indicators. Secondly, the indices of accounting information are relatively orthogonal to the market rate or return and the multicollinearity problem is much less essential relative to that of other multi-index models. Finally, the SUR estimation method can be used to take care of the interdependent relationship among securities of a particular industry. As quarterly data instead of annual data is employed to estimate the synthesis model for a particular industry, then the gain associated with the SUR estimation method will become much more important.[12]

The relationship among the systematic risks associated with Sharpe's model, S-L model, and the synthesis model can also be used to explore the possible applications of this model in security and portfolio analyses. Since the systematic risk obtained from Sharpe model is approximately identical to the systematic risk associated with the synthesis model, this kind of systematic risk can be regarded as an aggregated systematic risk. The statistical relationship defined in equation (8) implies that the aggregated systematic risk can be decomposed into components associated with the securities within an industry and the component associated with all other securities outside the industry. The disaggregated information has shed more light on the systematic risks which are generally used in security and portfolio analyses.

IV. Summary

In this paper, a synthesis model has been derived to integrate the Sharpe model with the Simkowitz-Logue model to avoid the problem of multicollinearity associated with 2SLS used by S-L. It is shown that the seemingly unrelated regression estimation method developed by Zellner can be applied to the synthesis model to take care of the interdependent relationship of security returns in the same industry. From the relationships among the Sharpe model, S-L model, and the synthesis model, it is shown that the aggregated systematic risk can be decomposed into components in accordance with possible sources of impacts. Empirical results of seven oil firms are used to demonstrate that some accounting information - leverage and dividend policy indices-might be used to increase the explanatory power of the diagonal security market model in capital asset pricing.

FOOTNOTES

*Associate Professor of Finance, University of Illinois and Assistant Professor of Finance, The Wharton School, University of Pennsylvania respectively. We would like to thank Irwin Friend for his helpful comments on an earlier draft.

1. It should be noted that S-L choose indices of profitability, leverage, and dividend policy as firm specific accounting information to include in the model as it is assumed that these factors will have the greatest impact on determining security returns. While others could be added or substituted, these same variables will be used in the present study to maintain continuity with the S-L study.

2. Klein and Nakamura [1962] and Fox [1968, 472-475] have discussed the possible problem of multicollinearity associated with 2SLS in detail.

3. According to King [1966] and Blume [1971], the market explains only 30 percent of the total variation in stock prices. King also found that approximately 12 percent of total variance could be explained by the industry influence.

4. Kamenta [1971] has explained why the SUR estimation method can be used to pool the time series information into the cross-section information.

5. Following Simkowitz and Logue [1973], the profitability index is defined as annual retained profit (retained earning plus interest and preferred dividend) divided by total assets; the leverage index is defined as annual change of long term debt plus annual change of outstanding preferred stock divided by total assets; and the dividend policy index is defined as annual change of total dividends divided by the book value of equity. Since annual instead of quarterly data are used in this study, the annualizing procedure used by S-L is not applicable here. It should be reemphasized that other firm-related variables can be added or substituted for those chosen by S-L.

6. A white noise series is a random sequence which in independent, normally distributed with zero mean and constant variance. As no discussion of this test is better than a necessarily brief one, the reader is referred to the original work by Box and Jenkins [1970] for a complete discussion of the Q-statistic.

7. In the same vein, the 2SLS has been used by Miller and Modigliain [1966] (m & M) to test the effectiveness of dividend policy in the electric utility industry. M & M regarded their 2SLS results as striking. However, it can easily be shown that the M & M results face similar problems with those of S-L's as discussed here. As the explanatory power of M & M's first-stage regression is essentially due to the dividend variable (see M & M, p. 361), the correlation coefficient between the estimated earning and dividends is approximately equal to unity. Hence, the dividend coefficients of M & M's second stage regression are subject to strong multicollinearity problem associated with 2SLS. Moreover, since the $\bar{R}^2$ of M & M's first stage regressions are only .50, .49, and .40 for 1957, 1956 and 1954 respectively, M & M's 2SLS estimators also are inefficient. From M & M and S-L's empirical studies of applying 2SLS in financial research, it is clear that the multicollinearity associated with 2SLS should be carefully examined to avoid misleading conclusions.

8. The 3SLS is also applied to the S-L model to obtain the simultaneous relationship of security returns for these same firms. It is found that the efficiency of 12 estimators has been significantly improved. These results are available on request.

9. Since S-L's empirical results do not give us enough information to estimate the right hand side of equation (7), and therefore the $\hat{\beta}_j$ obtained from Sharpe's model is used to do the empirical tests. These empirical results are employed to demonstrate that the firm related variables are essentially uncorrelated to the market rates of return. The same conclusion can also be drawn by comparing the estimated systematic risks listed in Table II with those listed in Table V.

10. The gain associated with the SUR estimation method is measured using the t-statistic of the regression coefficient as the coefficient of determination for the SUR estimation method is not provided by the SUR computer program. It is obvious, however, that the efficiency of SUR is greater than with OLS.

11. If the number of predetermined variables exceeds the number of observation on each variable, we cannot apply the standard two stage and three stage least squares procedures to estimate the parameters. There is a so-called problem of undersized sample. See Swamy and Holmes [1971] for detail.

12. In this circumstance, the sample size increase sharply and the gain associated with the SUR estimation method is substantiated. See Zellner [1962] for details.

REFERENCES

1. Anderson, T. W. (1958). An Introduction to Multivariate Statistical Analysis. New York: John Wiley and Sons, Inc.

2. Aber, J. W. (1973. Beta Coefficients and Models of Security Return. Lexington: C. C. Heath and Company.

3. Beaver, W. H. (1972), "The Behavior of Security Price and Its Implications for Accounting Research (Methods)," supplement to The Accounting Review, pp. 407-437.

4. Blume, M. (1971), "On the Assessment of Risk," Journal of Finance, 29, pp. 1-10.

5. Box, B. E. P., and Jenkins, G. (1970). Time Series Analysis, Forecasting and Control. San Francisco: Holden Day.

6. Cochran, W. G., "Some Effects of Measurement on Multiple Correlation Coefficients," Journal of American Statistical Association, Vol. 65 (1972), pp. 22-34.

7. Cohen, K., and J. Pogue (1967), "An Empirical Evaluation of Alternative Protfolio Selection Models," Journal of Business, 40, pp. 166-193.

8. Dowens, D., and T. Dyckman (1973), "Efficient Market Research and Accounting Information," Accounting Review, XLVIII, pp. 300-317.

9. Fama, E. F. (1968), "Risk, Return and Equilibrium: Some Clarifying Comments," Journal of Finance, 23, 1, pp. 29-40.

10. Fox, K. A. (1968). Intermediate Economic Statistics. New York: John Wiley and Sons, Inc.

11. Goldfeld, S. M., and R. E. Quandt (1965), "Some Tests for Homoscedasticity," Journal of the American Statistical Association, 60, pp. 539-57.

12. King, B. F. (1966), "Market and Industry Factors in Stock Price Behavior," Journal of Business, 39, pp. 139-190.

13. Klein, L. R. and Mitsugu Nakamura (1962), "Singularity in the Equation Systems of Econometrics: Some Aspects of the Problem of Multicollinearity," International Economic Review, 3, pp. 274-299.

14. Kmenta, J. (1971). Element of Econometrics. New York: The MacMillan Company.

15. Lintner, J. (1965), "Security Prices, Risk and Maximal Gains from Diversification," Journal of Finance, 20, pp. 587-616.

16. Miller, M. H. and F. Modigliani (1966), "Some Estimates of the Cost of Capital to the Electric Utility Industry, 1954-1957," American Economic Review, 56, pp. 333-391.

17. Mossin, J. (1966), "Equilibrium in a Capital Asset Market," Econometrica, 34, pp. 768-783.

18. Rogalski, Richard and Joseph D. Vinso, "Heteroscedasticity and the Market Model," Columbia University Working Paper No. 104, 1975.

19. Sharpe, W. F. (1964), "Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risks," Journal of Finance, 19, pp. 425-442.

20. Swamy, P. A. V. B., and J. Holmes (1971), "The Use of Under-sized Samples in the Estimation of Simultaneous Equation Systems," Econometrica, 39, pp. 455-459.

21. Simkowitz, M. A. and D. E. Logue (1973), "The Interdependent Structure of Security Return," Journal of Financial and Quantitative Analysis, 8, pp. 259-272.

22. Theil, H. (1971). Principles of Econometrics. New York: John Wiley and Sons, Inc.

23. Zellner, A. (1962), "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias," Journal of The American Statistical Association, 57, pp. 348-368.