Economics of Planning **30:** 33–49, 1997.

Applying Kalman Filter on Solving Simultaneous Equations with Overidentifying Rank Restrictions: The Analysis of the Demand and Supply Model of Medium-size Scooter Market in Taiwan

CHYAN YANG and WEN DEN CHEN
National Chiao Tung University, Institute of Management Science, Taipei, Taiwan

Key words: simultaneous equations, demand-supply functions, state space model, exogenous variables, endogenous variables

Abstract. Once the structure form of demand and supply is translated into a reduced form, one can solve the reduced form with a state space model of the Kalman filter method. This paper discusses an innovation representation that links the structure form with the state space model. For the state space model, the recursive Expectation Maximization (EM) algorithm is used to estimate the parameters of a structure form. This research successfully applied the Kalman filter method to the estimation of the coefficients of simultaneous equations with overidentifying rank restrictions. The empirical monthly data set came from the medium-size scooter market in Taiwan during 1987 to 1992 period.

1. Introduction

This article deals with the simultaneous equations of multiple time series via state space model, where the output equations are the linkage between simultaneous equations and state space model. In particular, this research focuses on solving simultaneous equations with overidentifying rank restrictions problem by using a revised expectation maximization (EM) algorithm.

Yang and Chen (1995) did not apply Kalman filter in solving simultaneous equations due to the difficulty of overcoming the problem of overidentifying rank restrictions; a linear stuctural form system which is overidentified will give rise to a set of non-linear parameter restrictions in the reduced form. This then means that the Kalman filter will not be a correct precedure to apply subject to these restrictions. However in the case of this demand-supply system the restrictions can be expressed in a purely linear form and so the precedure proposed here gives full ML estimates of the reduced form subject to the overidentifying restrictions of the structure. This research uses Kalman filter in a Demand-Supply simultaneous model and obtains the coefficients of structure form via EM algorithm. In estimating the parameters of endogenous variables, this research assumes that the relationship of the endogenous variables are constant (in order to simplify the model). This research concentrates on how the exogenous variables influence the endogenous variables over time. In this paper, a recursive EM algorithm is examined to obtain the parameters by the maximum likelihood method. The EM algorithm is basically to use the forward and backward recursions in finding the expected maximum log likely function until it converges.

Variables used in most estimation models are non-stationary time series data. Most time series forecasting models, however, are based on the assumption that they are stationary or co-integration time series (although the real data used for a time series are not). In other words, the forecasting by using non-cointegration data in the stable model has the essential problem which guarantees significant errors. The state-space model, however, can adapt to the time-varying nature and is very compatible to the applications of unstable time series.

Boas (1989) discusses why the conventional model failed and how this forecasting problem was tackled by applying the filter technique, a model developed for predictions in non-stationary situations. Harvey (1985) and Aoki (1987) applied the state space model by using Kalman filter in economics when analyzing the influence of GNP of U.S.A. Yang and Chen (1995) presented the general results concerning simultaneous equations with demand-supply model by using the reduced form to forecast the multiple inputs and multiple outputs problems. The research revealed that the traditional approach in econometrics had been to specify a representation based on economic theory, convert to a reduced form if necessary, impose indentifiability constraints, and then estimate model parameters. However, the previous paper did not discuss how to impose identifiability constraints of overidentifying rank restrictions and how to estimate the coefficients of the structure form. Shumway and Stoffer (1982) proposed the EM algorithm for smoothing and forecasting; this research revised the algorithm in estimating the coefficients of the structure form.

This paper is organized as follows: Section 2 briefly outlines the recursive EM algorithm and its background materials. In section 3, an empirical study characterizes the supply-demand model of the medium-size scooter motorcycle market in Taiwan. Section 4 gives the concluding remarks.

2. Innovation Representation in State Space Model

Consider the time-varying Demand-Supply model of the structure form represented in equation (1).

$$\begin{aligned}
Q_{D,t} &= P_t\alpha + U_{D,t}A_t + U_{C,t}C_t + \epsilon_{1,t} \quad \text{(demand)}\\
Q_{S,t} &= P_t\beta + U_{S,t}B_t + U_{C,t}D_t + \epsilon_{2,t} \quad \text{(supply)}\\
Q_{D,t} &= Q_{S,t} = Q_t \quad \text{(at equilibrium).}
\end{aligned} \tag{1}$$

In equation (1), $Q_{D,t}$ and $Q_{S,t}$ represent the quantity of demand and supply at time t respectively; P_t represents the sales price at time t. $U_{D,t}$ and $U_{S,t}$ represent the vectors of demand and supply function's exogenous variables at time t, respectively. $U_{C,t}$ represents the joint-exogenous variables vector and appears in both demand and supply equations at time t. Additionally, a and b are the price coefficients and are assumed constant over time. The relaxation of this assumption is an ongoing research and will be reported soon. A_t and B_t are the coefficient vectors of exogenous variables in demand and supply functions respectively at time t. C_t and D_t are the coefficients vectors of joint-exogenous variables in demand and supply equations. Error terms of this structure form at time t are expressed as $e_{1,t}$ and $e_{2,t}$. Reduced form of equation (1) can be expressed as follows:

$$\begin{aligned} Q_t &= U_{D,t}\Pi_{1,t} + U_{S,t}\Pi_{2,t} + U_{C,t}\Pi_{3,t} + \epsilon^*_{1,t} \\ P_t &= U_{D,t}\Pi_{4,t} + U_{S,t}\Pi_{5,t} + U_{C,t}\Pi_{6,t} + \epsilon^*_{2,t} \end{aligned} \tag{2}$$

$$\text{with } \begin{bmatrix} \Pi_{1,t} & \Pi_{4,t} \\ \Pi_{2,t} & \Pi_{5,t} \\ \Pi_{3,t} & \Pi_{6,t} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -\alpha & -\beta \end{bmatrix} = \begin{bmatrix} A_t & 0 \\ 0 & B_t \\ C_t & D_t \end{bmatrix}. \tag{3}$$

The vectors $\Pi_{1,t}$, $\Pi_{2,t}$, $\Pi_{3,t}$, $\Pi_{4,t}$, $\Pi_{5,t}$ and $\Pi_{6,t}$ represent the coefficients of reduced form (equation (2)) at time t. These vectors express how the exogenous variables ($U_{D,t}$, $U_{S,t}$, $U_{C,t}$) influence the endogenous variables (P_t, Q_t). and $\epsilon^*_{2,t}$ are error terms of the reduced form. From equation (3), two entries of zeros in right hand side matrix lead to the following relationship:

$$\begin{aligned} \Pi_{2,t} &- \alpha\Pi_{5,t} = 0 \\ \Pi_{1,t} &- \beta\Pi_{4,t} = 0 \end{aligned} \tag{4}$$

For equation (4) to be solvable, the vector pair of either ($\Pi_{2,t}$, $\Pi_{5,t}$) or ($\Pi_{1,t}$, $\Pi_{4,t}$) must have a scalar multiplier relationship. In other words, the scalar values of α and β can be solved by given ($\Pi_{2,t}$, $\Pi_{5,t}$) and ($\Pi_{1,t}$, $\Pi_{4,t}$), respectively. Given either vector pair of ($\Pi_{2,t}$, $\Pi_{5,t}$) or ($\Pi_{1,t}$, $\Pi_{4,t}$) to solve for scalar α or β, however, is a problem with overidentifying rank restrictions in general (unless equation are just identified).

Other relationships can be derived from equation (3) are listed in equation (5):

$$\begin{aligned} \Pi_{1,t} &- \alpha\Pi_{4,t} = A_t \\ \Pi_{2,t} &- \beta\Pi_{5,t} = B_t \\ \Pi_{3,t} &- \alpha\Pi_{6,t} = C_t \\ \Pi_{3,t} &- \beta\Pi_{6,t} = D_t \end{aligned} \tag{5}$$

Note that A_t, B_t, C_t, D_t, are the coefficients of structure form at time t. In other words, equation (5) is a linkage between the coefficients of reduced form (equation (2)) and that of the structure form (equation (1)). Furthermore, from equation (4) and (5), one can see that if the values of α, β, $\Pi_{2,t}$, $\Pi_{3,t}$, $\Pi_{4,t}$, $\Pi_{6,t}$, are known then one can obtain $\Pi_{1,t}$, $\Pi_{5,t}$, A_t, B_t, C_t and D_t easily. Therefore, for the parsimonious principle, one can define the state variables at time t as X_t in equation (6a). Since $\Pi_{3,t}$ and $\Pi_{6,t}$ relate the joint exogenous variables matrix $U_{C,t}$ they must be in the state variables. Additionally, if one includes either ($\Pi_{2,t}$, $\Pi_{4,t}$) or ($\Pi_{1,t}$, $\Pi_{5,t}$) in the state variables then one can construct a state vector. For example, X_t can be

$$X_t = [\Pi_{2,t}^{\tau}\, \Pi_{4,t}^{\tau}\, \Pi_{3,t}^{\tau}\, \Pi_{6,t}^{\tau}]^{\tau} \tag{6a}$$

Let matrix Θ be defined with two parameters, θ_1 and θ_2, as follows:

$$\Theta = \begin{bmatrix} 0 & I_D \times \theta_1 & 0 & 0 \\ I_s & 0 & 0 & 0 \\ 0 & 0 & I_c & 0 \\ 0 & I_D & 0 & 0 \\ I_s \times \theta_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & I_c \end{bmatrix}. \tag{6b}$$

X_t is a $\{$ rank(U_D)+rank(U_S)+2rank$(U_C)\}\times 1$ matrix and Θ is a $\{2[$ rank$(U_D)+$ rank(U_S) + rank$(U_C)]\} \times \{$ rank(U_D) + rank(U_S) + 2rank$(U_C)\}$ matrix. Additionally, I_D is a rank$(U_D) \times$ rank(U_D) identity matrix; I_S is a rank$(U_S) \times$ rank(U_S) identity matrix. I_C is a rank$(U_C) \times$ rank(U_C) identity matrix. Θ is the translation matrix and is derived from the constraints of equations (4).

Note that

$$\Theta \times X_t = [\theta_1 \Pi_{4,t}^T\, \Pi_{2,t}^T\, \Pi_{3,t}^T\, \Pi_{4,t}^T\, \theta_2 \Pi_{2,t}^T\, \Pi_{6,t}^T]^T. \tag{7}$$

When $\theta_1 = \beta$, $\theta_2 = 1/\alpha$, equation (7) can be rewritten as equation (8).

$$\Theta \times X_t = [\Pi_{1,t}^T\, \Pi_{2,t}^T\, \Pi_{3,t}^T\, \Pi_{4,t}^T\, \Pi_{5,t}^T\, \Pi_{6,t}^T]^T. \tag{8}$$

Since $\Pi_{1,t,\ldots}$, $\Pi_{6,t}$ are the coefficients of the reduced form in equation (2), equation (8) can be thought of as the coefficients of the reduced form. Therefore, by using Θ and X_t one can generate all $\Pi_{i,t}$, and subsequently A_t, B_t, C_t, D_t in equation (5) without solving equation (4). To construct the innovation equations, one can imitate the reduced form pattern and define the output matrix Ψ_t as follows:

$$\Psi_t = \begin{bmatrix} U_{D,t} & U_{S,t} & U_{C,t} & 0 & 0 & 0 \\ 0 & 0 & 0 & U_{D,t} & U_{S,t} & U_{C,t} \end{bmatrix}, t = 1, 2, \ldots, n, \tag{9}$$

where Ψ_t is a known $2\times\{2[$rank(U_D) + rank(U_S) + rank$(U_C)]\}$ design matrix that translates the unobserved stochastic vector, ΘX_t, into the two observed series Q_t and P_t, i.e. Y_t. Again, $U_{D,t}$, $U_{S,t}$ and $U_{C,t}$ in equation (9) are the exogenous variables shown in equation (1). The parameter n in equation (9) is the number of observations. Combining equations (6a), (6b), (8) and (9), the output equation can be expressed as follows:

$$Y_t = \Psi_t \Theta X_t + W_t, \quad t = 1, 2, \ldots, n, \tag{10}$$

where $W_t = \begin{bmatrix} W_{1,t} \\ W_{2,t} \end{bmatrix}$, and $Y_t = \begin{bmatrix} Q_t \\ P_t \end{bmatrix}$

In equation (10), Q_t (Quantity) and P_t (Price) are the endogenous variables. X_t, is the state variable at time t and contains the elements of $\Pi_{2,t}$, $\Pi_{4,t}$, $\Pi_{3,t}$, $\Pi_{6,t}$ (see equation (6a)).

The innovation representation can be defined follows:

$$\begin{aligned} &X_{t+1} = \Phi X_t + V_t \\ &Y_t = \Psi_t \Theta X_t + W_t, \quad t = 1, 2, \ldots, n. \end{aligned} \tag{11}$$

Φ is the transition matrix that translates state variable X_t to X_{t+1}. In equation (11), Vt and Wt are the error or noise terms and are assumed to be zero-mean uncorrelated normally distributed noise vectors with covariance matrix S and R. Having formulated the innovation representation, we now explain the procedure for estimating the coefficients of structure form.

Procedure for estimating the coefficients of structure form:

Note that, first, the joint log likelihood of the complete data X_0, $X_1, \ldots$, X_n, $Y_1, \ldots$, Y_n can be written in the form of equation (12) (see Appendix A).

$$\begin{aligned} \log L \;=\; & -\frac{1}{2}\ln|\Sigma| - \frac{1}{2}tr[\Sigma^{-1}(X_0-\mu)(X_0-\mu)^T]. \\ & -\frac{n}{2}\ln|S| - \frac{1}{2} - \frac{1}{2}tr[S^{-1}\Sigma_{t=1}^{n}(X_t-\Phi X_{t-1})(X_t-\Phi X_{t-1})^T] \\ & -\frac{n}{2}\ln|R| - \frac{1}{2}tr[R^{-1}\Sigma_{t=1}^{n}(Y_t-\Psi\Theta X_t)(Y_t-\Psi\Theta X_t)^T]. \end{aligned} \tag{12}$$

Note that *trA*, or trace of a matrix A, is defined as the scalar sum of diagonal entries of matrix A. The initial value X_0 is assumed to be a normally random vector with mean vector μ and the covariance matrix Σ. According to Shumway and Stoffer (1982), when log L is maximized means that the products of joint probabilities of V_t, W_t in equation (11) are maximized. Since the log likelihood given above depends on the unobserved state variables, X_t, t = 0, 1,. . .,n, the EM algorithm is considered to be applied conditionally with the observed series Y_1, $Y_2, \ldots$, Y_n. The parameters of (r+1)st iteration are defined as the values of μ, Σ, Φ, S, R, and Θ that maximize

$$G(\mu, \Sigma, \Phi, S, R, \Theta) = E_r(\log L|Y_1, \ldots, Y_n). \tag{13}$$

where E_r denotes the conditional expectation relative to a density containing the r-th iteration values of $\mu(r)$, $\Sigma(r)$, $\Phi(r)$, $S(r)$, $R(r)$ and $\Theta(r)$. The formula is similar to that in Shumway and Stoffer(1982) except adding a parameter Θ in G function.

By using the recursive EM algorithm, one can estimate the log likelihood with parameter μ, Σ, Φ, S, R, Θ, at each iteration. But there is one problem left to be solved: we must know how to estimate the θ_1, θ_2 to construct Θ. From equation (10), we can estimate θ_1, θ_2 as follows: (see Appendix B):

$$\hat{\theta}_1 \;=\; n^{-1}\Sigma_{t=1}^{n}\left[\frac{Q_t - U_{S,t}\Pi_{2,t} - U_{C,t}\Pi_{3,t}}{U_{D,t}\Pi_{4,t}}\right],$$

$$\hat{\theta}_2 = n^{-1}\Sigma_{t=1}^{n}\left[\frac{P_t - U_{D,t}\Pi_{4,t} - U_{C,t}\Pi_{6,t}}{U_{S,t}\Pi_{2,t}}\right] \tag{14}$$

Since Π_1 and Π_5 in equation (5) can be expressed in Π_2 and Π_4 respectively through equation (4), the coefficients of structure form can be computed by solving equation (5).

$$\begin{aligned} A_t &= (\beta - \alpha)\Pi_{4,t} \\ B_t &= (1 - \beta/\alpha)\Pi_{2,t} \\ C_t &= \Pi_{3,t} - \alpha\Pi_{6,t} \\ D_t &= \Pi_{3,t} - \beta\Pi_{6,t}. \end{aligned} \tag{15}$$

Since $\theta_1 = \beta$ and $\theta_2 = 1/\alpha$, one can compute α, β from equation (14). $\Pi_{2,t}$, $\Pi_{3,t}$, $\Pi_{4,t}$, $\Pi_{6,t}$ are the elements of state variables at time t and can be generated from the EM algorithm (see Appendix C).

3. Empirical Study

The motorcycle industry in Taiwan started in 1952. Motorcycles were first imported by some trading companies, and later the imported parts were assembled locally. In 1961 the government of Taiwan wanted to protect the industry and restricted the import of motorcycles except the parts. Since San Yang Co., Ltd., established the first mortorcycle manufactoring factory, several others joined the production. The domestic sales of motorcycles in 1966 reached 144,000 units, worth U.S. $ 50 million dollars, with more than forty companies. The market became competitive after Taiwan started a 20% advorum tax in 1968. Due to the growth of the market of Taiwan, the annual sales of motorcycles reached 746,000 units in 1979. However, the sales were decreased for the next three consecutive years due to the oil crisis. The market has been revived since 1986.

Taiwan is now the 14th largest international trading country and is the 3rd largest producer of motorcycles (preceded by Japan and Italy). The major firms that produce motorcycles are Sang-Yang Ltd., Kuang-Yang Ltd., Taiwan Sun-Yeh Ltd., Tai-Ling Ltd. and Taiwan vespa Ltd. Due to the low price, mobility, and easy parking, motorcycles are becoming the hottest commodity in Taiwan and many firms are extending their markets to mainland China. The motorcycle industry in Taiwan can categorize its market into five sub-markets: 50 c.c and below, 50–125 c.c. (medium-size), above 125 c.c., racing motorcycles, and the others. The medium-size scooter motorcycles market is the largest of these five markets and reaches a monthly sale of 50,000 units (see Figure 1). Therefore, this research focuses on the market of medium-size scooter.

The monthly data between January 1987 to December 1992, from the Bureau of Statistics, Ministry Administration and San Yang Lct., Co. of R. O. C., are available for this research. A model capable of expressing the market behavior

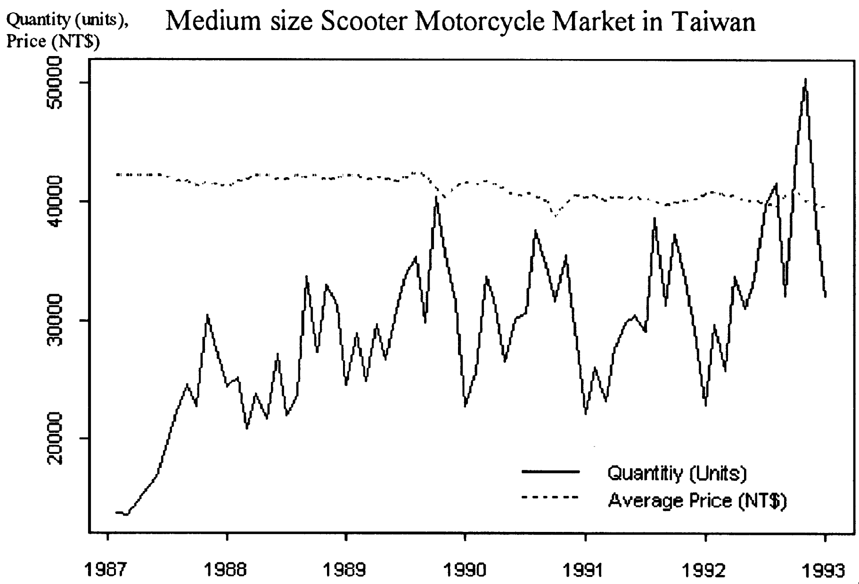

Figure 1. Number of Transaction Units and Average Price of medium-size scooter in Taiwan. Monthly data set of 72 months; January 1987–December 1992.

should therefore consist of at least two equations: demand and supply. These considerations lead to the following structure form:

$$Q_t = \alpha \times P_t + (ADV_t NOH_t) \times A_t + \text{ Constant } \times C_t + \epsilon_{1t} \text{ (Demand)}$$
$$Q_t = \beta \times P_t + (\text{ MWAGE}_t \text{ TREND}_t\ ) \times B_t + \text{Constant} \times D_t + \epsilon_{2t} \text{ (Supply)}$$

Q_t	the transaction quantity observed at time t, an endogenous variable in Demand-Supply model (unit: 100 units).
P_t	the transaction price observed at time t, an average price of various brands medium size scooter, an endogenous variables in Demand-Supply model (unit: NT$ 100).
ADV:	total expenditure of the advertisement in medium-size scooter (unit: NT$ 100,000).
NOH:	the number of houses in Taiwan (unit: 100,000 units).
MWAGEt	Average (labor costs) wage in manufacture industry (unit: NT$1,000).
TREND	time trend (unit: 1, 2, ..., n).
Constant	Constant term.
ϵ_{1t} and ϵ_{2t}	are the error or noise terms of demand and supply equations, respectively.

Table I. The $-2\log$ likelihood and BIC at each iteration by the EM algorithm

Iteration	BIC (Bayes information critical)	$-2\log(L)$
1	1523.91	1626.55
2	1320.69	1423.33
3	1208.12	1310.76
4	1135.24	1237.88
5	1082.98	1185.62
6	1044.25	1146.89
:	:	:
:	:	:
:	:	:
24	954.90	1057.49
25	954.82	1057.46

Refer to equation (1), α and β relate endogenous variables. α expresses how much the price influences the quantity in the demand equation or dQ_D/dP. If the commodity is a normal good then a should be negative. β expresses how much the price influences the quantity in supply function or dQ_S/dP. If the commodity is a normal good then β should be positive. A is the structure coefficients of the exogenous variable (ADV and NOH) at time t. (The ADV is the total expenditure of the advertisement in medium-size scooter. The NOH is the number of houses in Taiwan.) If the ADV and NOH increases, the quantity of medium-size scooter Q_t demand should increase and the value of A_t should be positive. The (MWAGE TREND) is a column vector and B_t is a 2×1 vector at time t where MWAGE is the industry average wage in manufactures. If MWAGE, a production cost, increases then the quantity of supply should decrease and the value of $B_{1,t}$ (the first element of B_t) should be negative.

In order to apply the Kalman filter state space (KFSS) EM procedure (see Appendix C), initial values are required for the parameters. Note that, to simplify the calculation, we set the initial values of S, R, μ, Σ by identity matrices. After setting the initial points, we can compute each iteration's log likelihood by using equation (12). If the estimation is defined as a stationary point of the likelihood function, then it is a non-decreasing likelihood, i.e. $-2\log(L)$ is a non-increase series (Dempster et al., 1977). From the equation (12), we can obtain the values of $-2\log(L)$ and BIC (Bayes information critical) for each iteration (Shumway, 1982).

Table 1 shows that the values of $-2\log(L)$ and BIC are decreasing. This meets Dempster assumption (Dempster et al., 1977) that $-2\log(L)$ is a non-increase series. At the convergency we can obtain the estimation of parameters θ_1, θ_2 and Θ. At the final convergent point, the estimate of Θ is,

$$\hat{\Theta} = \begin{pmatrix} 0 & 0 & 0 & 8.5547 & 0 & 0 \\ 0 & 0 & 0 & 0 & 8.5547 & 0 \\ 1.0000 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.0000 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1.000 & 0 \\ 0 & 0 & 1.0000 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1.0000 & 0 & 0 \\ -2.1969 & 0 & 0 & 0 & 0 & 0 \\ 0 & -2.1969 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1.0000 \end{pmatrix}$$

where, $\hat{\theta}_1$ = 8.5547 and $\hat{\theta}_2 = -2.1969$. Since $\alpha = 1/\theta_2$, $\beta = \theta_1$, the estimates of α and β can be evaluated as follows:

$$\begin{aligned} \hat{\alpha} &= -0.4552 \\ \hat{\beta} &= 8.5548. \end{aligned}$$

In other words, when the price increases one unit then the quantity of demand will decrease 0.4552 units and the quantity of supply will increase 8.5548 units. These results are consistent with consumers' behavior. Similarly, from equation (12), we can obtain transition matrix, the covariance matrices of V_t and W_t as follows:

transition matrix (see Appendix B)

$$\hat{\Phi} = \begin{pmatrix} 0.8607 & -0.1767 & -2.2222 & -0.7813 & 0.0043 & 0.0555 \\ 0.0713 & 0.9724 & 0.3969 & 0.0496 & 0.0002 & -0.0064 \\ 0.0037 & 0.0160 & 0.5178 & 0.0422 & -0.0003 & -0.0015 \\ -0.0271 & -0.1844 & 0.0319 & 0.3254 & 0.0064 & 0.0446 \\ 0.0533 & 0.0566 & -0.0741 & 0.2444 & 0.9954 & -0.0152 \\ 0.1547 & 0.2750 & 0.2113 & 1.3330 & -0.0189 & 0.8988 \end{pmatrix}$$

covariance matrix of the noise term V_t

$$\hat{S} = \begin{pmatrix} 0.4711 & -0.1712 & 0.0118 & -0.0031 & 0.0101 & 0.0102 \\ -0.1712 & 0.0683 & -0.0035 & 0.0047 & -0.0044 & 0.0097 \\ 0.0118 & -0.0035 & 0.0033 & -0.0060 & 0.0012 & 0.0031 \\ -0.0031 & 0.0047 & -0.0060 & 0.0217 & -0.0037 & -0.0031 \\ 0.0144 & -0.0069 & 0.0012 & -0.0037 & 0.9822 & -0.0755 \\ 0.0092 & 0.0102 & 0.0031 & -0.0010 & -0.0656 & 0.6668 \end{pmatrix}$$

and covariance matrix of the noise term W_t

$$\hat{R} = \begin{pmatrix} 0.9993 & -0.0000 \\ -0.0000 & 0.9636 \end{pmatrix}.$$

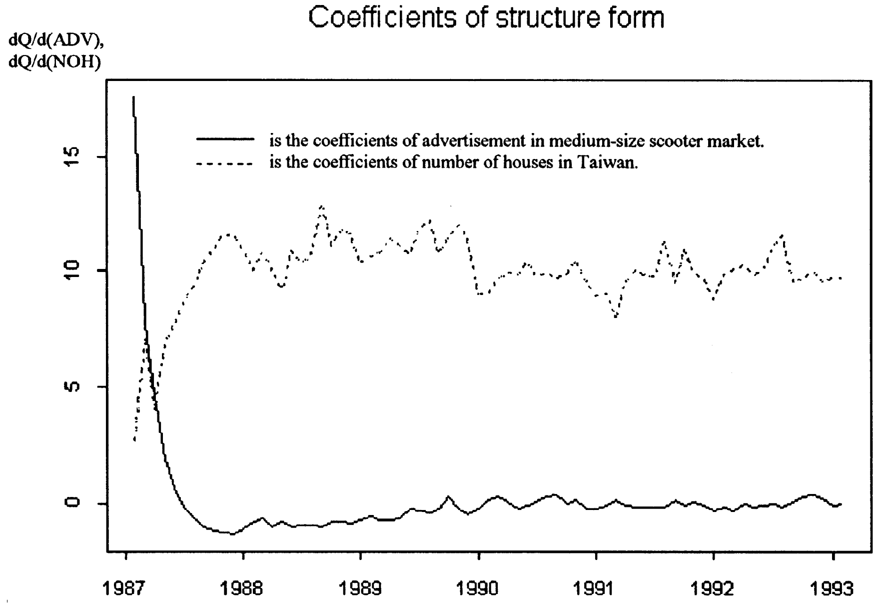

Figure 2. The coefficients A_t of ADV (total expenditure of the advertisement in medium-size scooter) and NOH (the number of houses in Taiwan) in Demand function at last iteration. January 1987–December 1992.

The estimates of $\hat{\Phi}$, $\hat{S}$, $\hat{R}$ are all fixed points of convergency over time at the last iteration of EM algorithm. The next step, the dynamic coefficients of exogenous variables are calculated. Figures 2–4 exhibit how the exogenous variables coefficients change over time.

The demand side exogenous variables, coefficients $\hat{A}_t$ of ADV (total expenditure of the advertisement in medium-size scooter) and NOH (the number of houses in Taiwan), derived from equation (15), at last iteration by EM algorithm) are shown in Figure 2.

The coefficients of ADV and NOH express how advertisement and the growth of housing supply influence the qunatity of demand of medium- size scooter.

Second, about the supply side exogenous variables, the coefficients of MWAGE and TREND ($\hat{\beta}_t$ t = 1,2,. . .,144, calculated by equation (15) at last iteration), are shown in Figure 3.

Figure 3 shows that the coefficients of the MWAGE increase and then decrease whereas the coefficients of TREND increase. The coefficients of constant term $\hat{C}_t$ in demand equation and $\hat{D}_t$ in supply equation, t = 1,2,. . .,144, are shown in Figure 4.

Having calculated the estimates of transition equation residuals (V_t) and output equation residuals (W_t), one should test that whether these residuals are normally distributed. $W_{1,t}$ and W_{2t} are tested to see whether they are normally distributed. Figure 5 shows the residuals, $W_{1,t}$, $W_{2,t}$ of output equation (10). Note that $W_{1,t}$ and $W_{2,t}$ are within the range of -0.6 and 0.6.

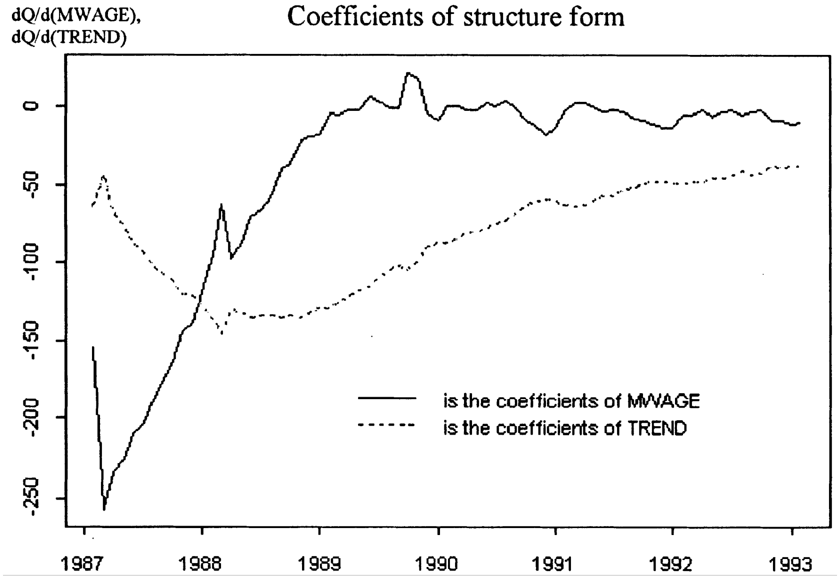

Figure 3. The coefficients of MWAGE and REND, B_t, in supply function at last iteration. January 1987–December 1992.

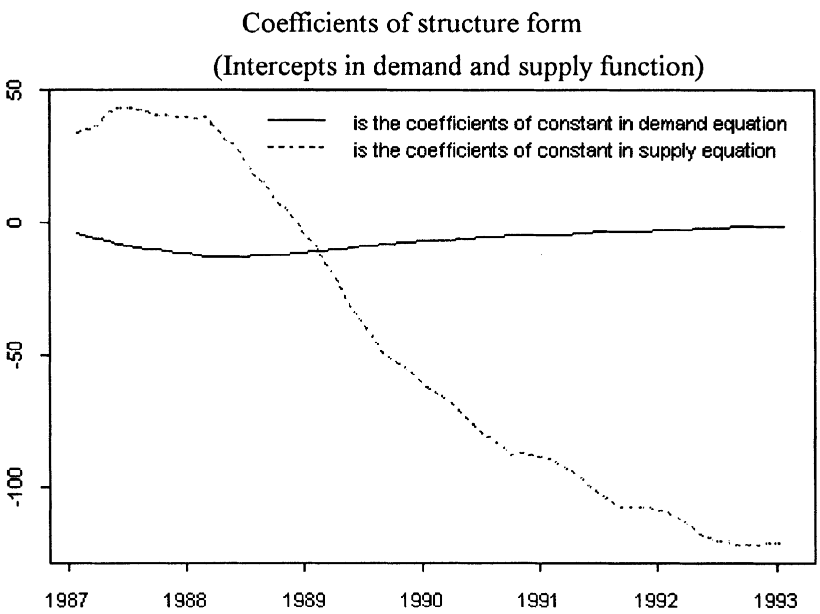

Figure 4. The coefficients C_t, D_t of constant terms in Demand and Supply functions at last iteration. Jan. 1987-Dec. 1992

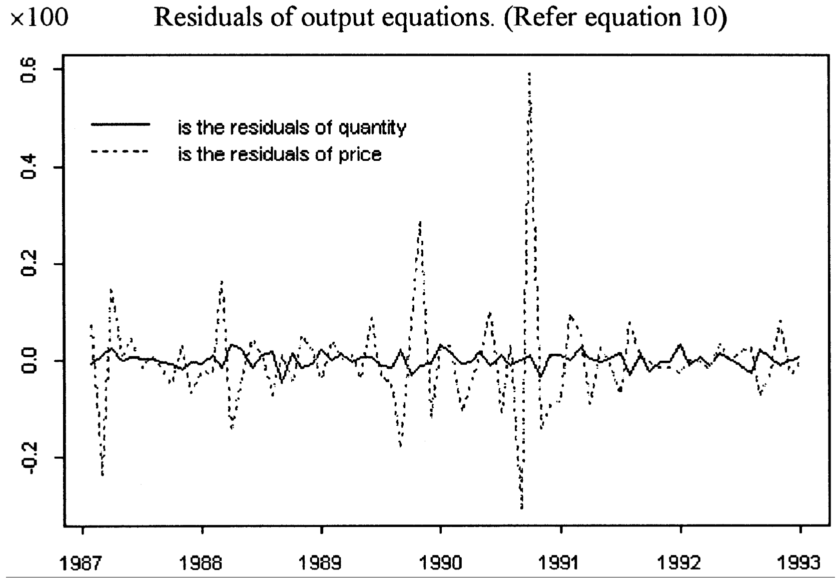

Figure 5. The residuals in the output equations at last iteration. Jan. 1987-Dec. 1992

Through the quantile-quantile plot and Bowman and Shenton's (1975) test for normality, one can evaluate the residuals of the Q_t and P_t to see whether they are normally distributed. Figure 6 shows the characteristics of residual $W_{1,t}$.

The picture at upper left of Figure 6 is the histogram. The density function by using bandwidth is shown at lower left. The box plot is shown at upper right. At the lower right of Figures 6 is the quantile- quantile plot with normal quantile.

Bowman and Shenton (1975) proposed an alternative test for normality. Let $\sqrt{\beta_1} = \mu_3/\mu_2^{3/2}$ and $\beta_2 = \mu_4/\mu_2^2$ μ_r (r=2,3,4) where is the rth moment about the mean. For the normal distribution β_1=0 (since the distribution is symmetric) and β_2=3. The estimates of β_1 and β_2, say, β_1 and β_2, are obtained by replacing the μ's with their sample estimates and are defined as

$$\hat{\mu}_r = \frac{1}{n}\Sigma_{t=1}^{n} W_{1,t}^{r} \quad (r = 2, 3, 4. \quad \text{n is the number of observations}).$$

Therefore, test for normality is then a test of the null hypothesis

$$H_0 : \beta_1 = 0 \text{ and } \beta_2 = 3.$$

We may use the following test statistic which, under the null hypothesis, has a chi-square distribution with two degrees of freedom (Bowman, 1975).

$$n[b_1/6 + (b_2 - 3)^2/24] \sim \chi^2(2) \quad \text{n is the number of observations.}$$

The test statistics for $W_{1,t}$ and $W_{2,t}$ are the values of 0.3253499 and 0.9062722 respectively, the p values are 0.8498674 and 0.6356316 > 0.05. If we were willing

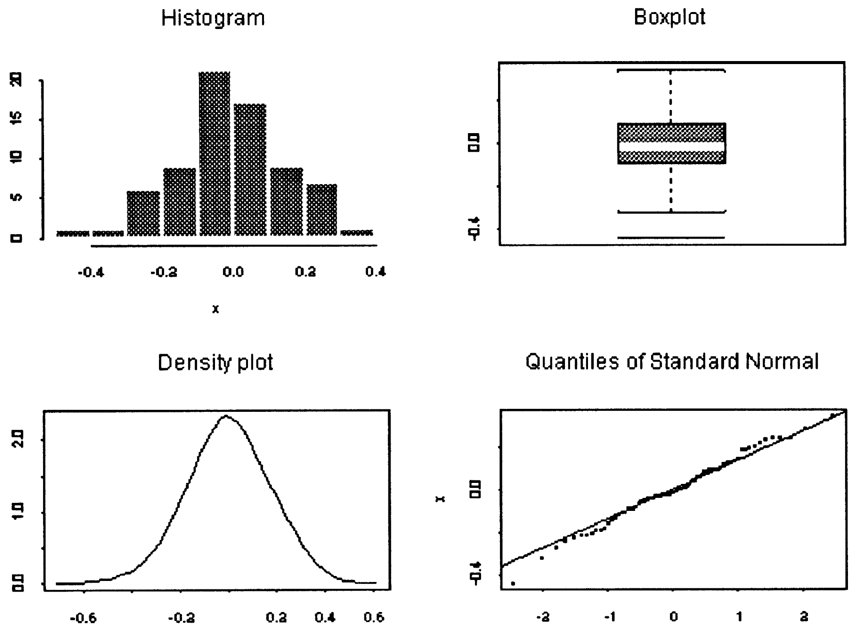

Figure 6. Histogram, box plot, density plot and quantile quantile plot of the residual $W_{1,t}$. The picture at upper left is the histogram that expresses the frequency of the residuals. The picture at lower left is the density function by using bandwidth that expresses the distribution of the residuals. The picture at upper right is the box plot that expresses mode and mean of the residuals. It shows a little skewed toward right. The one at lower right is the well-known Q-Q plot with normal quantile that expresses the residuals meet the normal distribution.

to use the asymptotic test procedure, we would not reject the hypothesis of normality at the 5% level.

4. Conclusions

We successfully applied the Kalman filter method to estimate the coefficients of simultaneous equations with overidentifying rank restrictions. The recursive EM algorithm shows that the estimation of likelihood is non-decreasing and model parameters (Φ, Θ, R, S) are convergent. The method dynamically adjusts the exogenous variables coefficients of structure form to estimate the supply-demand equations. In other words, this method indicates how the exogenous variables affect the endogenous variables (price and quantity) in the market over time.

The coefficients of the endogenous variables were assumed as constant in this research. To relax this assumption, the Kalman covariance transition equation can be revised. The relaxation of this assumption deserves further researches.

5. Acknowledgement

The authors are indebted to Dr. Paul D. Gilbert for his instructive comments that motivate our investigation on this subject. The authors are also thankful to the anonymous referees for their helpful sugguestions that make the paper more valuable and more readable. This research is supported in part by NSC84–3015–p009–001.

Appendix A: Finding the Estimator of Parameters that Used in Kalman Filter Procedure

By adding the Θ into the Shumway and Stoffer's (1982) smoothing theory, the log likely function can be expressed as follows:

$$\begin{aligned} G(\mu, \Sigma, \Phi, S, R, \Theta) \;=\; & -\frac{1}{2}\ln|\Sigma| - \frac{1}{2}tr\left\{\Sigma^{-1}(P_{0|n}(X_0-\mu)(X_0-\mu)^T)\right\} \\ & -\frac{n}{2}\ln|S| - \frac{1}{2}tr\left\{S^{-1}(E - H\Phi^T - \Phi H^T + \Phi M\Phi^T)\right\} \\ & -\frac{n}{2}\ln|R| - \frac{1}{2}tr\left\{R^{-1}\Sigma_{t=1}^n\left[(Y_t - \Psi\Theta X_t)^T(Y_t - \Psi\Theta X_t)\right]\right\}, \end{aligned}$$

where the conditional mean

$$X_{t|n} = E(X_t|Y_1, Y_2, \ldots, Y_n)$$

and covariance functions

$$P_{t|n} = \mathrm{cov}(X_t|Y_1, Y_2, \ldots, Y_n)$$

and

$$P_{t,t-1|n} = \mathrm{cov}(X_t, X_{t-1}|Y_1, Y_2, \ldots, Y_n).$$

From Shumway and Stoffer's (1982) smoothing theory, M, H, E, Φ, Θ, R are defined as follows:

$$M = \Sigma_{t=1}^n(P_{t-1|n} + X_{t-1|n}(X_{t-1|n})^T)$$

$$H = \Sigma_{t=1}^n(P_{t,t-1|n} + X_{t|n}(X_{t-1|n})^T)$$

$$E = \Sigma_{t=1}^n(P_{t|n} + X_{t|n}(X_{t|n})^T)$$

$$\Phi(r+1) = HM^{-1}$$

$$S(r+1) = n^{-1}(E - HM^{-1}H^T)$$

$$R(r+1) = n^{-1}\Sigma_{t=1}^n\left[(T_t - \Theta_t\Psi_t X_t)(Y_t - \Theta\Psi_t X_t)^T + \Theta_t\Psi_t P_{t|n}\Psi_t^T\Theta_t^T\right].$$

This research defines a parameter Q(r+1) as follows:

$$\Theta(r+1) = \begin{bmatrix} 0 & I_D \times \theta_1(r+1) & 0 & 0 \\ I_s & 0 & 0 & 0 \\ 0 & 0 & I_c & 0 \\ 0 & I_D & 0 & 0 \\ I_s \times \theta_2(r+1) & 0 & 0 & 0 \\ 0 & 0 & 0 & I_c \end{bmatrix},$$

where

$$\theta_1(r+1) = n^{-1}\Sigma_{t=1}^{n}\left[\frac{Q_t - U_{S,t}\Pi_{2,t} - U_{C,t}\Pi_{3,t}}{U_{D,t}\Pi_{4,t}}\right].$$

$$\theta_2(r+1) = n^{-1}\Sigma_{t=1}^{n}\left[\frac{P_t - U_{D,t}\Pi_{4,t} - U_{C,t}\Pi_{6,t}}{U_{S,t}\Pi_{2,t}}\right].$$

Appendix B: The Relationship Between the Exogenous Variables, Coefficients and Θ

$Y_t = \Psi_t \Theta X_t + W_t$, is the short form of

$$\begin{bmatrix} Q_t \\ P_t \end{bmatrix} = \begin{bmatrix} U_{D,t} & U_{S,t} & U_{C,t} & 0 & 0 & 0 \\ 0 & 0 & 0 & U_{D,t} & U_{S,t} & U_{C,t} \end{bmatrix} \begin{bmatrix} 0 & I_D \times \theta_1 & 0 & 0 \\ I_s & 0 & 0 & 0 \\ 0 & 0 & I_c & 0 \\ 0 & I_D & 0 & 0 \\ I_s \times \theta_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & I_C \end{bmatrix} \begin{bmatrix} \Pi_{2,t} \\ \Pi_{4,t} \\ \Pi_{3,t} \\ \Pi_{6,t} \end{bmatrix} + \begin{bmatrix} W_{1,t} \\ W_{2,t} \end{bmatrix}.$$

To evaluate the θ_1 and θ_2, the matrix form can be rewritten as:

$$Q_t = \theta_1 \Pi_{4,t} U_{D,t} + \Pi_{2,t} U_{S,t} + \Pi_{3,t} U_{C,t} + W_{1,t}$$
$$P_t = \theta_2 \Pi_{2,t} U_{S,t} + \Pi_{4,t} U_{D,t} + \Pi_{6,t} U_{C,t} + W_{2,t}.$$

as E(W$_{1,t}$) = 0 and E(W$_{2,t}$) = 0, the estimates of θ_1 and θ_2 are

$$\hat{\theta}_1 = n^{-1}\Sigma_{t=1}^{n}\left[\frac{Q_t - U_{S,t}\Pi_{2,t} - U_{C,t}\Pi_{3,t}}{U_{D,t}\Pi_{4,t}}\right].$$

$$\hat{\theta}_2 = n^{-1}\Sigma_{t=1}^{n}\left[\frac{P_t - U_{D,t}\Pi_{4,t} - U_{C,t}\Pi_{6,t}}{U_{S,t}\Pi_{2,t}}\right].$$

Appendix C: A Revised Shumway EM Algorithm by Adding Θ

X denotes state vector

Y denotes observed output
μ_0 Θ Φ are defined in equation (12)
Given: $\Phi(0)$, S(0), R(0), Θ(0), μ0, Y_i, i = 1,. . . n.
Find: max E_r(log $L|Y_i$), X_i(r)| i = 1,. . .,n.
Note: $P_{t,t-1}$ is the covariance matrix of X_t, X_{t-1}.
Q_t is the covariance of V_t.
R_t is the covariance of W_t.
Step 1: Forward recursions. Do the following for t = 1,2,. . .n

$$\{$$

$$\begin{aligned}
1.\ P_{t,t-1} &= \Phi_{t-1}P_{t-1,t-1}\Phi^T_{t-1} + S_{t-1}\\
2.\ X_{t,t} &= X_{t,t-1} + G_t(Y_t - \Theta_t\Psi_t X_{t,t-1}\\
3.\ G_t &= P_{t,t-1}\Psi^T_t\Theta^T_t(\Theta_t\Psi_t P_{t,t-1}\Psi^T_t\Theta^T_t + R_t)^{-1}\\
4.\ P_{t,t} &= (I - G_t\Theta a_t\Psi_t)P_{t,t-1}
\end{aligned}$$

$$\}$$

Step 2: Backward recursions.
Do the following for t = n, n–1,. . .,1

$$\{$$

$$\begin{aligned}
1.\ J_{t-1} &= P_{t-1|t-1}\Theta^T_t(P_{t|t-1})^{-1}\\
2.\ X_{t-1|n} &= X_{t-1|t-1} + J_{t-1}(X_{t|n} - \Phi_t X_{t-1|t-1})\\
3.\ P_{t-1|n} &= P_{t-1|t-1} + J_{t-1}(P_{t|n} - P_{t|t-1})J^T_{t-1}\\
4.\ P_{n,n-1|n} &= (I - \Theta_t\Psi_t)\Phi_n P_{n-1|n-1}
\end{aligned}$$

$$\}$$

Step 3: Compute $E_r(\log L|Y_i)$.

Step 4: Compute parameters $\Phi(r+1), S(r+1), R(r+1), \Theta(r+1)$ (for next time).

Step 5: If E_r converges then exit else goto step 1 (forward recursion).

References

Aoki, Masanao (1987), State Space Modeling of Time Series, Springer-Verlag, Berlin Heidelberg.

Boas, J. (1989), 'Forecasting under unstable conditions: A case study of the cocoa market', European Journal of Operational Research, 41, 15–22.

Bowman, K.O. and Shenton, L.R. (1975), 'Omnibus Contours for Departures from Normality Based on $b_1^{1/2}$ and b_2', Biometrika, 62, 243–250.

Burmeister, E., Wall, K.D. and Hamilton, J.D. (1986), 'Estimation of unobserved expected monthly inflation using Kalman filtering', Journal of Business & Economic Statistics, 4(2), 147–160.

Dempster, A.P., Laird, N.M. and Rubin, D.B. (1977) 'Maximum likelihood from incomplete data via the EM algorithm'. J.R. Statist. Soc. B 39, 1–38.

Gilbert, P.D. (1995), 'Combining VAR estimation and state space model reduction for simple good predictions.' Journal of Forecasting, 14, 229–250.

Harvey, A.C. (1989). Forecasting, Structural Time Series Models and Kalman Filter, Cambridge University Press.

Mehra, R.K. (1970), 'On the identification of variances and adaptive Kalman filtering,' IEEE Transactions on Automatic Control, AC–19(2), 175–184.
Myeres, K.A. and Tapley, B.D. (1976), 'Adaptive sequential estimation with unknown noise statistics', IEEE Transactions on Automatic Control, August 520–523.
Shumway, R.H., and Stoffer, D.S. (1982), 'An approach to time series smoothing and forecasting using the EM algorithm', Journal of Time Series Analysis, 3(4).
Yang, C. and Chen, W.D. (1995), 'An application of the state space model by using Kalman filter: the analysis of the demand and supply model of medium-size scooter market in Taiwan', Management Review, 14(2), 41–58.