Kalman Filter Estimation of Property Price Bubbles in Seoul

Xiao Qin

Division of Economics

School of Humanities and Social Sciences

Nanyang Technological University of Singapore

cqxiao@ntu.edu.sg

Tan Gee Kwang, Randolph

Division of Economics

School of Humanities and Social Sciences

Nanyang Technological University of Singapore

arandolph@ntu.edu.sg

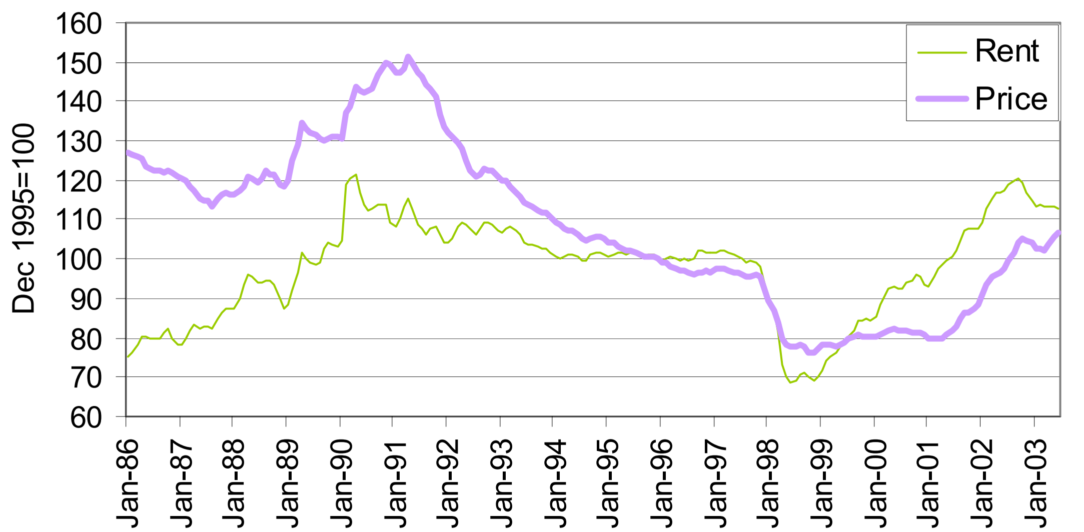
Abstract

The literature of asset pricing in general and speculative bubble in particular is massive. Nonetheless the theoretical development and empirical studies are more or less focused on the stock, the foreign exchange and the bond markets. The less populous studies on property market, when available, are usually conducted using data from OECD countries, such as US, UK and Japan. This paper is a study on the Seoul property market. We estimate, using Kalman filter technique and EM algorithm, a type of bubble in the property price which collapses periodically. The estimation is based on the state space form representation of the present value model. We found that bubble do appear to be important in driving the Seoul property prices. However the simple present value model without bubbles and the AR(1) model performs better in terms of in-sample estimation errors and out-of sample forecast accuracies.

1. Introduction

In the 1990s, many Asian property markets experienced rapid price increases. The property fever came to a sudden end in most of these markets with the Asian financial crises in late 1997. However, the Korean property prices move opposite of this general trend---rising in the second half of 1980s, falling in 1990s, and rising steadily again after 2001. The recent rise in the property prices is suspected, in the popular press, to be driven by bubbles. Figure 1.1 shows the CPI deflated housing price and rent indices in Seoul, Korea, between January 1986 and June 2003.

Figure 1.1. **Seoul Housing Price and Rent Indices**



Source: CEIC database

The purpose of this paper is to investigate the behavior of the property prices, in particular the possibility of existence of speculative bubbles, in Seoul, Korea in the past two decades.

Price bubble is a sharp, temporary price increase that cannot be plausibly explained by changes in fundamental value drivers. A bubble can be rational or irrational. Rational bubble occurs if investors have asymmetric and incomplete information; irrational bubble occurs if investors buy already overpriced assets on the belief that they can resell it at a even higher price.

The literature of asset pricing in general and speculative bubble in particular is massive. Nonetheless the theoretical development and empirical studies are more or less focused on the stock, the foreign exchange and the bond markets. The less populous studies on property market, when available, are usually conducted using data from OECD countries, such as US, UK and Japan.

Like bonds and equities, the value of commercial real estate depends on the expected cash flows it can generate in its lifetime. Unlike bonds and equities, the markets for real estate are thin and the information on market transactions is scarce, which makes the valuation of real estate much more difficult. Furthermore, the high transaction costs in the real estate market limit the ability of investors to trade on their opinion, hence market price may fail to fully reflect information available on the market; the lengthy lags in bringing substitute assets to the market can extend the period in which excess cash flows will be earned. Such excess cash flows seem to rationalize high prices but fail to reflect changes in the fundamentals of the economy. Lax bank lending is another factor conducive of property bubbles. Developments are often funded with lines of credit. This virtually guarantees eventual completion of a development once a loan is negotiated, even the underlying economics of the property market have soured in the process. New substitute assets being created in a property boom may continue to be produced well after the boom ends. Consequently, overshooting in property value on the downside may occur. All those factors make the real estate market more prone to rational bubbles than many other asset markets.

2. Theoretical Background and Empirical Evidences

In standard asset pricing model, individuals estimate asset values based on their expectations of future cash flows and required rates of return (discount rate). They buy (sell) when prices fall below (rise above) their value estimates. Knowledgeable investors recognize that their estimates contain errors which, combined with risk-aversion and capital constraints, make them unwilling to buy (sell) unless prices deviate significantly from what they perceive to be true. How much a risk-averse investor is willing to commit beyond a particular asset position depends on his wealth and ability to borrow, as well as the size of his current position. Investors' flow demand (supply) schedules will systematically create an increased demand (supply) for an asset whose price falls (rises). Consequently, prices should be fairly stable in the absence of new information.

Investors are heterogeneous in their estimates of the intrinsic asset values. When investors have different opinions, the market price will reflect their willingness and ability to trade. If most investors believe an asset is under-valued, their willingness to buy causes the price of the asset to rise. Investors who are generally correct in their estimates of asset values tend to make money and accumulate wealth, thus have more influence on price-setting. This will improve asset pricing overtime.

In all, asset prices will eventually return to its fundamental true value, but disturbances, which cause information noise, can drive the asset price way above or below its true value for lengthy periods of time.

2.1 Present Value Model and Speculative Bubble---the Theory

If economic agents are risk neutral, the price of one equity share, P_t, would be equal to the expected discounted present value of the dividend accruing to ownership of the equity share during the ownership period, D_t, plus the price at which the share can be sold at the end of the ownership period, P_{t+1}. Mathematically,

$$P_t = \frac{E_t[P_{t+1} + D_t]}{1 + R_t} \qquad (2.1.1)$$

where

P_t: the real price of the property asset at time t;

D_t: the real total rents received during the period t;

R_t: the time-varying real discount rate.

Define

$$r_t \equiv \log(1 + R_t) \qquad (2.1.2)$$

Hence

$$r_t \equiv \log(E_t[P_{t+1} + D_t]) - \log(P_t) \qquad (2.1.3)$$

In a static world, dividends grows at constant rate, g, and the log of dividend-to-price ratio is also a constant. That is,

$$\log\left(\frac{D_t}{D_{t-1}}\right) = d_t - d_{t-1} = \Delta d_t = g \tag{2.1.4}$$

and

$$\log\left(\frac{D_{t-1}}{P_t}\right) = d_{t-1} - p_t = \delta \tag{2.1.5}$$

Equation (2.4) and (2.5) would imply that the asset price grows at the same rate as the dividend, and the ratio of asset price to the sum of asset price and dividend is also a constant. That is

$$\log\left(\frac{P_{t+1}}{P_t}\right) = \log\left(\frac{D_t/\delta}{D_{t-1}/\delta}\right) = \log\left(\frac{D_t}{D_{t-1}}\right) = g \tag{2.1.6}$$

and

$$\frac{P_t}{P_t + D_{t-1}} = \frac{1}{1+\frac{D_{t-1}}{P_t}} = \frac{1}{1+\exp(\delta)} \equiv \rho \tag{2.1.7}$$

In such a world the log of the gross discount rate, r_t, would also be a constant. To see this, define

$$\xi_t \equiv \kappa + \rho p_{t+1} + (1-\rho)d_t - p_t \tag{2.1.8}$$

Given the characteristics of the static world, $\xi_t \equiv \kappa + g + (1-\rho)\delta \equiv \xi$, which is a constant. If we set

$$\kappa = -\log(\rho) - (1-\rho)\delta \tag{2.1.9}$$

Then

$$\xi = -\log(\rho) + g = \log\left(\frac{E_t\left[P_{t+1} + D_t\right]}{P}\right) = r \tag{2.1.10}$$

Thus in a static world

$$r_t = \xi_t = \kappa + \rho p_{t+1} + (1-\rho)d_t - p_t = \xi \tag{2.1.11}$$

When the world is evolving over time, equation (2.11) would hold only approximately.

Solve (2.11) for p_t by forward iteration

$$\begin{aligned} p_t &= \kappa - \xi + \rho E_t\left[p_{t+1}\right] + (1-\rho)d_t \\ &= \kappa - \xi + \rho\left\{\kappa - \xi + \rho E_t\left[p_{t+2}\right] + (1-\rho)E_t\left[d_{t+1}\right]\right\} + (1-\rho)d_t \\ &= \ldots \\ &= \frac{\kappa - \xi}{1-\rho} + \rho^i E_t\left[p_{t+i}\right] + (1-\rho)\sum_{j=0}^{i-1}\rho^j E_t\left[d_{t+j}\right] \end{aligned} \tag{2.1.12}$$

If the boundary condition, $\lim_{i\to\infty}\rho^i E_t\left[p_{t+i}\right] = 0$, is satisfied, we would have the fundamental solution for property price

$$p_t = p_t^f = \frac{\kappa - \xi}{1-\rho} + (1-\rho)\sum_{j=0}^{\infty}\rho^j E_t\left[d_{t+j}\right] \tag{2.1.13}$$

This is the present value model for asset pricing.

However, the transversality condition may fail to hold. In such case, we would expect the bubbly solution for the asset price

$$p_t = p_t^f + b_t \tag{2.1.14}$$

where the bubble component

$$E_t\left[b_{t+i}\right] = \frac{1}{\rho^i} b_t \tag{2.1.15}$$

which implies

$$b_{t+1} = \frac{1}{\rho} b_t \tag{2.1.15`}$$

If the log property prices and the log property rents are I(1) processes, i.e. having one unit root, the following model may be estimated instead

$$\Delta p_t^{\ f} = p_t^{\ f} - p_{t-1}^{\ f} = (1-\rho)\sum_{j=0}^{\infty} \rho^j \left\{E_t\left[d_{t+j}\right] - E_t\left[d_{t+j-1}\right]\right\} \tag{2.1.16}$$

Suppose the growth of the dividends follows an AR(1) process (justification will be given for this assumption in section 3.)

$$\Delta d_t = \phi \Delta d_{t-1} + \delta_t; \qquad E\left[\delta_t\right] = 0, \qquad Var\left[\delta_t\right] = \sigma_\delta^2 \tag{2.1.17}$$

then

$$\Delta p_t^{\ f} = \frac{1}{1-\phi\rho} \Delta d_t - \frac{\phi\rho}{1-\phi\rho} \Delta d_{t-1} \equiv \psi \Delta d_t + (1-\psi) \Delta d_{t-1} \tag{2.1.18}$$

If bubble is present

$$\Delta p_t = \Delta p_t^{\ f} + \Delta b_t \tag{2.1.19}$$

where

$$\Delta b_{t+1} = \frac{1}{\rho} \Delta b_t \tag{2.1.20}$$

Suppose the bubble process is itself stochastic, we can write the bubble equation as,

$$\Delta b_{t+1} = \frac{1}{\rho} \Delta b_t + \zeta_t \qquad E\left[\zeta_t\right] = 0, \qquad Var\left[\zeta_t\right] = \sigma_\zeta^2 \tag{2.1.20`}$$

<u>2.2 Speculative Bubbles---Empirics</u>

The methods employed in the empirical studies can be broadly divided into two classes: one based on a completely specified fundamental model, the other tests the implications of efficient market hypothesis on the behavior of the asset prices without parametric specification. Most authors testing the existence of speculative bubbles do not distinguish between rational and irrational bubbles, given the obvious difficulty of the task.

Campbell and Shiller (1987) based their tests on the implication of the present value model of the form

$$p_t = \theta(1-\delta)\sum_{i=0}^{\infty}\delta^i E_t d_{t+i} \tag{2.2.1}$$

If the variables in equation (2.2.1) are stationary in their first differences, one can define

$$S_t \equiv p_t - \theta d_t \tag{2.2.2}$$

Equation (2.2.1) would imply

$$S_t = \theta\sum_{i=1}^{\infty}\delta^i E_t \Delta d_{t+i} \equiv E_t S_t^* \tag{2.2.3}$$

They estimated the following VAR model

$$\begin{bmatrix} \Delta d_t \\ S_t \end{bmatrix} = \begin{bmatrix} \alpha(L) & \beta(L) \\ \gamma(L) & \lambda(L) \end{bmatrix}\begin{bmatrix} \Delta d_{t-1} \\ S_{t-1} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \tag{2.2.4}$$

Restriction of equation (2.2.3) implies

$$\gamma_i = -\theta\alpha_i \text{, i=1…p;} \tag{2.2.5a}$$

$$\lambda_1 = \frac{1}{\delta} - \theta\beta_1 \tag{2.2.5b}$$

and $$\lambda_i = -\theta\beta_i \text{, i=2…p.} \tag{2.2.5c}$$

Their test shows that the present value model hypothesis is rejected statistically at the conventional significance levels. However, the strength of the evidence depends sensitively on the discount rate assumed in the test. And when applied to bubble tests, the procedure is biased towards accepting the hypothesis of no-bubble when the rational bubble collapses periodically.

Wests (1987) introduced a test based on present value model and Hausman's (1978) specification test. In his test, the null hypothesis of no bubble is rejected by the data. However, the equation used for the basis of test could be mis-specified.

The direct test approach, which is based on the formulation and estimation of a complete parametric specification, is susceptible to the criticism that it is unable to detect bubbles other than those belonging to the specific parametric class under consideration. So failure to reject the no-bubble hypothesis does not necessarily imply the absence of other unspecified types of bubbles.

Many empirical works testing the existence of bubbles do not specify an equilibrium model to which the market adjusts in the long run. One such test is based on the second moments of the asset prices.

According to Shiller (1981), the perfect foresight rational asset price is

$$p_t^* = \sum_{i=1}^{\infty} \left[\frac{1}{1+r} \right]^i d_{t+i} \qquad (2.2.6)$$

the actual asset price

$$p_t = E\left[p_t^*\right] \tag{2.2.7}$$

since

$$p_t^* = E\left[p_t^*\right] + u_t = p_t + u_t \tag{2.2.8}$$

therefore

$$Var\left(p_t\right) \leq Var\left(p_t^*\right) \tag{2.2.9}$$

Notice that equation (2.2.1) is not measurable. Grossman and Shiller (1981) hence estimated the following equation

$$\hat{p}_t = \sum_{i=1}^{T-t}\left[\frac{1}{1+r}\right]^i d_{t+i} + \left[\frac{1}{1+r}\right]^{T-t} p_T \tag{2.2.10}$$

where T is the last observation in the sample. They then compare the variance of $\hat{p}_t$ with the variance of the actual asset price. However, this test is not valid, since if bubble exists, p_T in equation (2.2.5) would contain the bubble, i.e. $p_T = p_T^f + b_T$. Shiller (1981) uses $p_T^* = \frac{\sum_{t=1}^{T} p_t}{T}$ to replace p_T. However $p_T^* \neq \sum_{i=T+1}^{\infty}\left[\frac{1}{1+r}\right]^{i-T} d_i$, and the results can be misleading if dividends are smoothed by the management.

Other non-parametric tests in the literature are variance decomposition tests introduced by Campbell (1991) and Cochrane (1992), sign tests by Evans (1986), run tests and tail tests by Blanchard and Watson (1982), etc. Those tests show mixed results.

Diba and Grossman (1984, 1988) and Hamilton and Whiteman (1985) put forward a class of indirect tests based on checking the order of integration of a given pair of variables.

The idea is that if asset prices are not more explosive than the relevant driving fundamental variable, then it can be concluded that rational bubbles are not present. However, this indirect procedure is not always reliable. In the majority of applications, part of the difficulty is that the stationarity properties of the relevant time series are analyzed by testing the null hypothesis of a unit root in the levels and differences of the series against one-sided stationary alternatives rather than the more relevant explosive ones. Although in theory such tests should be capable of revealing the existence of a rational bubble, in small samples, series with explosive bubble components could look very much like stationary processes when differenced a sufficient number of times.

In addition to these difficulties, Evans (1991) has demonstrated that the use of standard unit root and cointegration tests for prices and underlying fundamentals can erroneously lead to acceptance of the no-bubble hypothesis for an important class of rational bubbles that collapse periodically. In such cases, even a direct test for explosive behavior in the levels of the relevant series may fail to detect the bubble since such a test tends to have low power. Evans' basic argument is that integration and cointegration tests are likely to have some power to detect a rational bubble only when the latter lasts for most of the period under investigation. In theory, a deterministic bubble will continue to infinity, and integration tests should clearly provide a good way of detecting such an event. Nevertheless, the real world bubbles, if exist, are stochastic, or periodic, so that periods of expansion will eventually be followed by a collapse or contraction. Evans argued, despite having explosive conditional means, such collapsing bubbles will appear to standard unit root tests as stationary processes, and tests with either stationary or explosive alternatives will have little or no power to detect the bubbles. The basic problem is that the periodically collapsing bubbles only exhibit characteristic bubble behavior during their expansion phase. As a result, a test is more likely to find evidence of systematic divergence between asset prices and fundamentals if it is based only on data points that are associated with the expansion phase of the bubble.

There are researchers who, instead of testing for the existence of bubble, attempt to estimate the bubble in equation (2.1.14). Yangru Wu (1997) estimated the type of rational bubbles in the stock market which can collapse and restart continuously. In his paper, the bubble is treated as an unobserved state variable and can be estimated using the Kalman filter. His study shows that much of the deviations of stock prices from the present-value model are captured by the bubble.

Our paper applies the Kalman filter approach to estimate bubbles in the property prices of Seoul, Korea.

3. The Data

The time series selected for estimation in this paper are the CPI deflated housing price index and housing rent index for Seoul. The data are taken from CEIC database. Both series run from January 1986 to June 2003---hence there are 210 observations for each. After differencing and taking lags, the actual sample size for estimation is 208.

The data identification process shows that both series have slowly decaying sample ACF and the Phillips-Perron tests indicate that each has one unit root. Performing log-transformations on these series do not improve the situation. Hence the first difference is applied to the log-transformed series.

The sample partial ACF of the first difference of the log-price exhibit AR(1) pattern. Thus a preliminary AR(1) model, $\Delta p_t = \mu_p + \varphi \Delta p_{t-1}$, is fit to the data by conditional least square estimation. The result shows that the intercept is not distinguishable from zero, but the coefficient on the first lag is highly significant. However the residuals are not white noise, indicating that the model is not adequate. The same results are obtained for the first difference of the log-rent series. Although based on AIC statistics, an AR(15) model is selected for the latter series, all coefficients, except for the first and the fifth lag are statistically insignificant. Besides, adding 14 extra parameters do no change the fact that the residuals are not white noise. The summary statistics are presented below in table 3.1 and table 3.2. Given these result, we assume in our paper that rents follows a AR(1) process with zero intercept.

Table 3.1. **Autocorrelation and Unit Root Tests**

	Differenced Log Price			Differenced Log Rent		
H0: no auto-correlation	Lag	Box Ljung statistics (χ_p)	Probability	Lag	Box Ljung statistics (χ_p)	Probability
	6	*143.11*	*0.000*	6	*100.93*	*0.000*

H0: exist one unit root	Lag=0	Phillips –Perron statistics	Probability	Lag=0	Phillips –Perron statistics	Probability
	Zero mean	*-7.3736*	*0.0001*	Zero mean	*-7.8034*	*0.0001*
	Single mean	*-7.3693*	*0.0001*	Single mean	*-7.8219*	*0.0001*
	Trend	*-7.4162*	*0.0001*	Trend	*-7.8095*	*0.0001*

Table 3.2. **Conditional Least Square Estimates of ARMA Models**

		Differenced Log Price	Differenced Log Rent
AR(1)	Intercept (t ratio)	*-0.0009* *(-0.49)*	*0.0020* *(0.78)*
	AR 1 (t ratio)	*0.5817* *(10.27)*	*0.5425* *(9.29)*
	Standard error estimate	*0.0101*	*0.017*
	H0: residuals are white noise, lag=6 (probability)	*16.05* *(0.007)*	*49.07* *(0.000)*
ARMA(1,1)	Intercept (t ratio)	*-0.0008* *(-0.45)*	*0.0020* *(0.83)*
	MA1 (t ratio)	*0.1243* *(1.05)*	*-0.2192* *(-1.83)*
	AR1 (t ratio)	*0.6649*	*0.3983*

		(7.45)	*(3.53)*
	Standard error estimate	*0.0106*	*0.0171*
	H0: residuals are white noise, lag=6 (probability)	*15.52* *(0.004)*	*30.65* *(0.000)*
AR(15)	Intercept (t ratio)	*-.0009* *(-0.38)*	*0.0019* *(0.94)*
	AR1 (t ratio)	*0.5385* *(7.48)*	*0.5329* *(7.47)*
	AR2 (t ratio)	*0.1886* *(2.30)*	*0.0606* *(0.75)*
	AR3 (t ratio)	*-0.2001* *(-2.41)*	*-0.1328* *(-1.64)*
	AR4 (t ratio)	*-0.0074* *(-0.09)*	*-0.0150* *(-0.19)*
	AR5 (t ratio)	*0.0988* *(1.18)*	*0.2002* *(2.48)*
	AR6 (t ratio)	*0.1127* *(1.36)*	*0.0221* *(0.27)*
	AR7 (t ratio)	*0.0311* *(0.37)*	*-0.0637* *(-0.78)*
	AR8 (t ratio)	*-0.1107* *(-1.33)*	*-0.0727* *(-0.89)*
	AR9	*-0.0223*	*0.0277*

	(t ratio)	*(-0.27)*	*(0.34)*
	AR10 (t ratio)	*0.1837* *(2.20)*	*-0.0346* *(-0.42)*
	AR11 (t ratio)	*-0.1296* *(-1.54)*	*0.0581* *(0.72)*
	AR12 (t ratio)	*-0.1527* *(-1.79)*	*0.1385* *(1.71)*
	AR13 (t ratio)	*-0.0496* *(-0.58)*	*-0.0025* *(-0.03)*
	AR14 (t ratio)	*-0.0549* *(-0.66)*	*-0.1050* *(-1.30)*
	AR15 (t ratio)	*-0.0225* *(-0.30)*	*-0.1354* *(-1.89)*
	Standard error estimate	*0.0102*	*0.0155*
	H0: residuals are white noise, lag=6 (probability)	*0.00* *(0.000)*	*0.00* *(0.000)*

4. State Space Form, Kalman Filter and Estimation Strategies

4.1 State Space Form

To make use of the Kalman filter, we first express the present value model, presented in section 2.1, in a time-invariant state space form.

Let z_t be a $nz-$vector of state variables, x_t a $l-$vector of inputs, and y_t a $ny-$vector of outputs. The state space model consists of two equations:

The measurement equation:

$$y_t = Hz_t + Bx_t + \varepsilon_t\,;\; E(\varepsilon_t) = 0\,,\; \mathrm{var}(\varepsilon_t) = R \tag{4.1}$$

and the transition equation:

$$z_t = Fz_{t-1} + Ax_t + \eta_t\,;\; E(\eta_t) = 0\,,\; \mathrm{var}(\eta_t) = V \tag{4.2}$$

where the system matrices H, B, F, A, the measurement variance R and the transition variance V are all time-invariant. We also assume that ε_t and η_t and uncorrelated.

In our case,

$$y_t = \begin{pmatrix} \Delta p_t \\ \Delta d_t \end{pmatrix}, \quad z_t = \Delta b_t\,,\; x_t = \begin{pmatrix} \Delta d_t \\ \Delta d_{t-1} \end{pmatrix}, \qquad \varepsilon_t = \begin{pmatrix} 0 \\ \delta_t \end{pmatrix}, \qquad \eta_t = \zeta_t$$

$$H = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad F = \frac{1}{\rho}\,, \qquad B = \begin{pmatrix} \psi & 1-\psi \\ 0 & \phi \end{pmatrix}, \qquad A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

Hence

$$R = \begin{pmatrix} 0 & 0 \\ 0 & \sigma_\delta^2 \end{pmatrix}, \qquad V = \sigma_\zeta^2 \tag{4.3}$$

4.2 Kalman Filter and the Bubble Estimation

The bubble component, Δb_t, is not observable but can be estimated as a state variable using the Kalman filter, assuming the system parameters are known. The Kalmal filter consists of a set of recursive equations. Suppose we estimate the initial value of the state variable to be z_0, with estimation error P_0. The predicted value of the state variable and the prediction error at time t, given information set available at time $t-1$, $\Xi_{t-1} = \{y_1,...y_{t-1}, x_1,..., x_{t-1}\}$, can be calculated using the prediction equations recursively forward:

$$z_{t|t-1} = Fz_{t-1|t-1} + Ax_{t-1} \tag{4.4a}$$

$$P_{t|t-1} = FP_{t-1|t-1}F` + V \tag{4.4b}$$

When time t information becomes available, we can update our estimation of the bubbles and their estimation errors using the filtering equations recursively forward:

$$z_{t|t} = z_{t|t-1} + \kappa_t \varepsilon_{t|t-1} \tag{4.5a}$$

$$P_{t|t} = P_{t|t-1} - \kappa_t HP_{t|t-1} \tag{4.5b}$$

where

$$\kappa_t = P_{t|t-1}H`D_{t|t-1}^{-1} \tag{4.5c}$$

$$D_{t|t-1} = (HP_{t|t-1}H` + R) \tag{4.5d}$$

$$\varepsilon_{t|t-1} = y_t - Hz_{t|t-1} - Bx_t \tag{4.5e}$$

Once we obtained the sequences $\{z_{t|t-1}\}_{t=1}^{T}$, $\{P_{t|t-1}\}_{t=1}^{T}$, $\{z_{t|t}\}_{t=1}^{T}$ and $\{P_{t|t}\}_{t=1}^{T}$, we can have a more efficient estimation of the state variable and its estimation errors, using the full set of information, $\Xi_T = \{y_1,...y_T, x_1,..., x_T\}$, and the following smoothing equations by backward recursion:

$$z_{t|T} = z_{t|t} + J_t(z_{t+1|T} - z_{t+1|t}) \tag{4.6a}$$

$$P_{t|T} = P_{t|t} + J_t(P_{t+1|T} - P_{t+1|t})J_t \tag{4.6b}$$

where

$$J_t = P_{t|t} F` P_{t+1|t}^{-1} \tag{4.6c}$$

The starting values for smoothing are $z_{T|T}$ and $P_{T|T}$ obtained from the filtering process.

4.3 The Model Parameter Estimation Strategy

There are only four unknown parameters in the model given in section 2.1, which are ρ, ϕ, σ_ζ^2 and σ_δ^2. These parameters are estimated by maximizing the log likelihood function of y_t, $t = 1,2,...T$, which is

$$LogL(y; x, \theta) = -\frac{nyT}{2}\log 2\pi - \frac{1}{2}\sum_{t=1}^{T}\log\left|D_{t|t-1}\right| - \frac{1}{2}\sum_{t=1}^{T}\varepsilon_{t|t-1}` D_{t|t-1}^{-1}\varepsilon_{t|t-1} \tag{4.7}$$

Where

$$\begin{aligned} \theta &= \begin{pmatrix}\rho & \phi & \sigma_\zeta^2 & \sigma_\delta^2\end{pmatrix}` \\ y &= (y_1, y_2,..., y_T)` \\ x &= (x_1, x_2,..., x_T)` \end{aligned} \tag{4.8}$$

We obtained the estimates of θ using EM algorithm.

In applying the EM algorithm, we first obtain $\frac{\partial Log(y;x,\theta)}{\partial \theta_i}$, $i = 1,2,3,4$, then take the expectation of $\frac{\partial Log(y;x,\theta)}{\partial \theta_i}$ with respect to information set $\Xi_T = \{y_1,...y_T, x_1,..., x_T\}$ and set it to zero. By solving the set of four equations hence obtained, we get the ML estimators:

$$
\begin{aligned}
\widetilde{R} &= \frac{1}{T}\sum_{t=1}^{T}\left[HP_{t|T}H` + \left(y_t - Hz_{t|T} - Bx_t\right)\left(y_t - Hz_{t|T} - Bx_t\right)`\right] \\
\widetilde{V} &= \frac{1}{T}\left[S_t(0) - S_t(1)\left(S_{t-1}(0)\right)^{-1} S_t(1)`\right] \\
\widetilde{F} &= S_t(1)\left(S_{t-1}(0)\right)^{-1} \\
\widetilde{D} &= \left(\sum_{t=1}^{T} y_t x_t` - H\sum_{t=1}^{T} z_{t|T} x_t`\right)\left(\sum_{t=1}^{T} x_t x_t`\right)^{-1}
\end{aligned}
\tag{4.9}
$$

where

$$
S_t(1) = \sum_{t=1}^{T} P_{t,t-1|T} + z_{t|T} z_{t-1|T}`
$$

$$
S_t(0) = \sum_{t=1}^{T} P_{t|T} + z_{t|T} z_{t|T}` \tag{4.10}
$$

$P_{t,t-1|T}$: estimated covariance between z_t and z_{t-1}.

In computing equation (4.9), we need estimated values of the state variable and their estimation errors. Thus we have to provide a guess starting value for the system parameters. The estimation process is as following:

Step 1: initiate guessed system parameter values;

Step 2: run through equations (4.4), (4.5) and (4.6), to obtain the sequences $\{z_{t|t-1}\}_{t=1}^{T}$, $\{P_{t|t-1}\}_{t=1}^{T}$, $\{z_{t|t}\}_{t=1}^{T}$, $\{P_{t|t}\}_{t=1}^{T}$ $\{z_{t|T}\}_{t=1}^{T}$ and $\{P_{t|T}\}_{t=1}^{T}$;

Step3: compute ML estimates of system parameters using equation (4.9).

Step4: repeat step 2 and 3 until convergence occurs.

The convergence criterion is set to be $\left\{ \max\left[abs\left(\frac{x^{i+1} - x^i}{x^i + 10^{-6}} \right) \right] < 10^{-2} \right\}$[1] for the estimates of the parameters, and $\left\{ \max\left[abs\left(\frac{x^{i+1} - x^i}{x^i + 10^{-6}} \right) \right] < 10^{-3} \right\}$ for the log likelihood.

The initial system parameter values in our experiment are based on the preliminary OLS estimate of the simple present value model without bubble component. The initial values of the state variable for the $(i+1)^{th}$ iteration is updated using estimates from the i^{th} iteration by the set of equations,

$$\begin{aligned} z_0^{\,i+1} &= F^i z_{0|T}^{\,i} \\ P_0^{\,i+1} &= F^i P_0 F^{i\,`} + V^i \end{aligned} \tag{4.11}$$

4.4 Asymptotic Properties of the ML Estimators

Suppose $\tilde{\theta}$ is the ML estimator of θ obtained by maximizing (4.7).

Subject to certain regularity conditions (Caines, 1988, Ch7),

$$\sqrt{T}\varphi_{2D,T}^{\frac{1}{2}}(\tilde{\theta} - \theta_0) \xrightarrow{d} N(0, I) \tag{4.12}$$

i.e.

$$\tilde{\theta} \xrightarrow{d} N(\theta_0, T^{-1}\varphi_{2D,T}^{-1}) \qquad (4.12`)$$

where $\varphi_{2D,T}$ is the information matrix from the sample of size T

---

[1] x^i is the estimate of x obtained in the i^{th} iteration.

$$\varphi_{2D,T} = -\frac{1}{T} E\left(\sum_{t=1}^{T} \frac{\partial^2 LogL_t}{\partial\theta\partial\theta`} \middle| \theta = \theta_0 \right) \tag{4.13}$$

with

$$LogL_t = -\frac{ny}{2}\log 2\pi - \frac{1}{2}\log\left|D_{t|t-1}\right| - \frac{1}{2}\varepsilon_{t|t-1}`D_{t|t-1}^{-1}\varepsilon_{t|t-1},\ t = 1,2,\ldots,T \tag{4.14}$$

and

$$\lim_{T\to\infty} \varphi_{2D,T} \xrightarrow{p} \hat{\varphi} = -\frac{1}{T}\left(\sum_{t=1}^{T} \frac{\partial^2 LogL_t}{\partial\theta\partial\theta`} \middle| \theta = \tilde{\theta} \right) \tag{4.15}$$

The reported standard errors for $\tilde{\theta}$ are the square roots of the diagonal elements of

$$(T\hat{\varphi})^{-1} = -\left(\sum_{t=1}^{T} \frac{\partial^2 LogL_t}{\partial\theta\partial\theta`} \middle| \theta = \tilde{\theta} \right)^{-1}$$

The Hessian is calculated numerically in our paper. The method is described below. First we collect the estimated parameters in a $r \times 1$ vector $\tilde{\theta}$. We perturb one parameter at a time by $\Delta = +0.01$. Running through the Kalman filter again and recalculate the log-likelihood of the data. We then perturb one parameter at a time by $\Delta = -0.01$. Running through the Kalman filter again and recalculate the log-likelihood of the data. The Hessian is calculated using the formula[2]

$$\frac{\partial^2 LogL(y;\theta)}{\partial\theta_i^2} \approx \frac{LogL(y;\tilde{\theta} + \Delta_i) - 2 \times LogL(\theta) + LogL(\tilde{\theta} - \Delta_i)}{\Delta_i^2} \tag{4.16}$$

where

Δ_i: a $r \times 1$ vector with all elements, except the i-th which is 0.01, to be zero;

and

[2] Refer to Gerald Wheatley "applied numerical analysis", seventh edition, page 267.

θ_i : the i^{th} element of $\tilde{\theta}$.

The standard error of θ_i is approximated using the equation,

$$SE(\theta_i) = \sqrt{\left(-\frac{\partial^2 LogL(y;\tilde{\theta})}{\partial \theta_i^2}\right)^{-1}} \tag{4.17}$$

5. Estimation Results

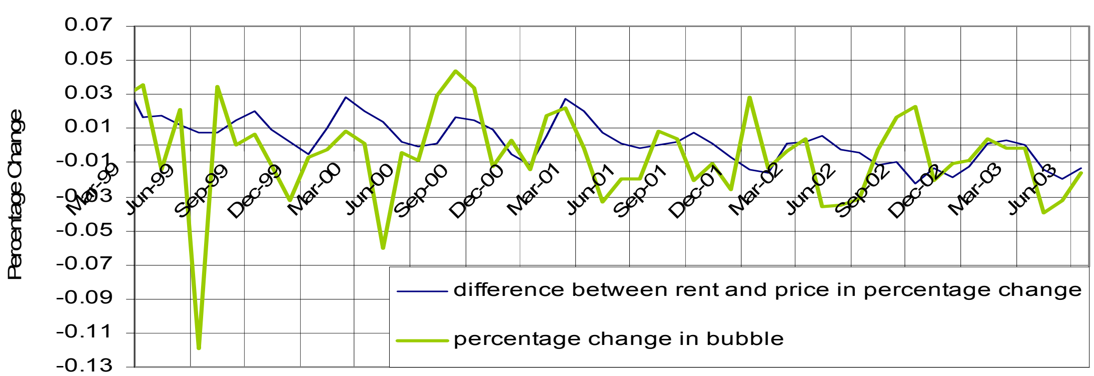
In the estimation process, we encountered negative variances. In such cases, following the usual practices, we set the variance to zero. The estimation of the model achieved convergence after 61 iterations. The results are shown in Table 5.1, which indicates that only ϕ is significant at the conventional levels.

Table 5.1. **ML Parameter Estimates and Statistical Properties**

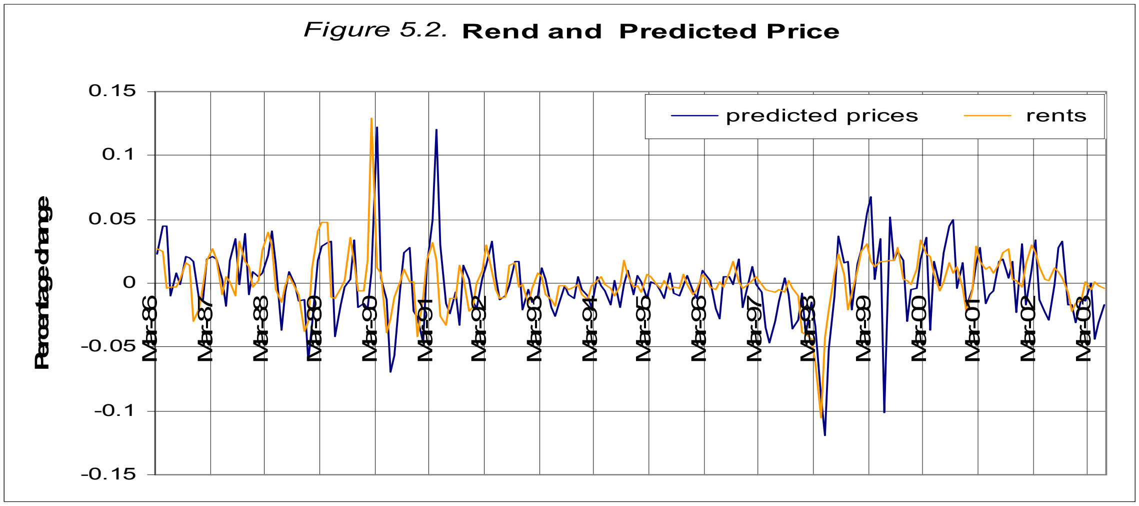
	$\beta = \frac{1}{\rho}$	ϕ	σ_ζ^2	σ_δ^2
ML estimates	0.0067	0.5462	0.0004	0.00030
Standard Error	0.0205	0.0568	0.0005	0.0004
t ratio	0.3268	9.6162	0.8	0.75

The fact the $\frac{1}{\rho}$ is indistinguishable from zero implies that $\psi = 0$ and $1-\psi = 1$ i.e. $\Delta p_t^{\ f} = \frac{1}{1-\phi\rho}\Delta d_t - \frac{\phi\rho}{1-\phi\rho}\Delta d_{t-1} \equiv \psi\Delta d_t + (1-\psi)\Delta d_{t-1} = \Delta d_{t-1}$. That is the fundamental price depends only on rent in the previous period. In another word, future cash flows can be approximated by rent received in the previous period. If the change in rent in the past period is zero, there should be no change in the fundamental price. Hence if change in the actual price occurs, it is attributable to bubble. Figure 5.1 shows that when the difference between the percentage change in rent and percentage change in price increases (decreases), the estimated percentage change in bubble also increases (decreases). The magnitude of change in bubble in general exceeds that in the rent-price differential. This could be caused by market information noise.

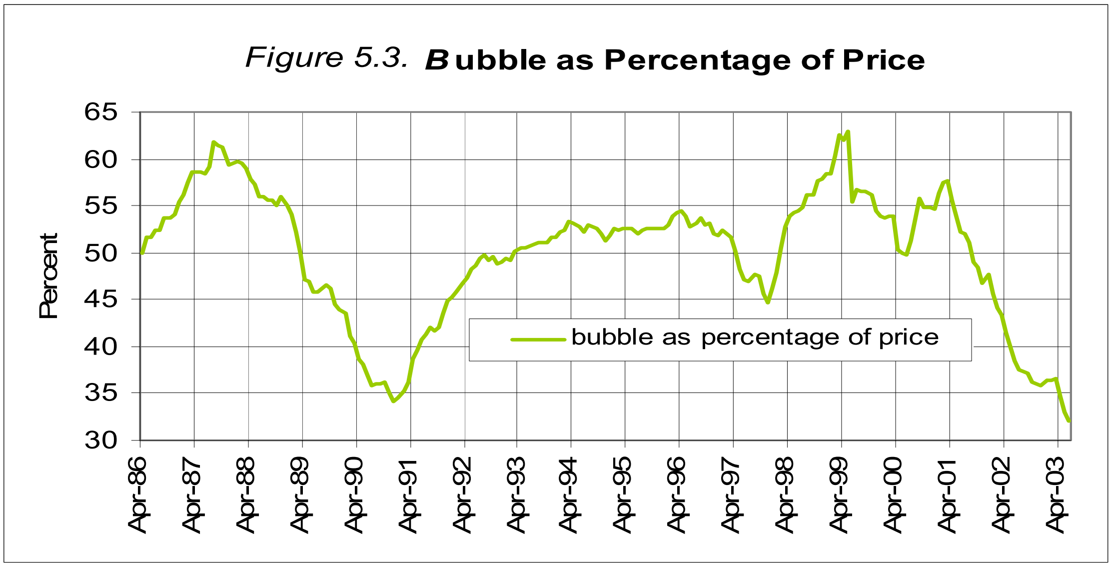
Figure 5.1. **Bubble and Differences between Prices and Rents**



The model implies that the predicted percentage change in price should mimic that in rent, except for the time when the change in bubble is large. This trend can be seen by comparing Figure 5.1 and Figure 5.2.



The Kalman filter computed bubble values are change in log values, i.e. percentage changes. To convert them back to level values, we need to make certain assumptions. According to Diba and Grossman (1988), the principle of free disposal rules out negative bubble, and if a positive rational bubble exists it can start only on the first date of trading of a stock. Assuming the starting of the sample to be the first trading date, and assuming the price at that data contains 50% of bubble, we can obtain the level values for bubble from its percentage changes. The results are shown in Figure 5.3. The figure shows that the portion of price which is non-fundamental fluctuates between 30% and 65%, given the assumption of initial condition. The fluctuation in the bubbly portion of price could be caused by noise in the market information. In times of better information, the bubbly portion declines (collapses). But when market information is plagued by noised, this portion swells up (expands). Figure 5.3 shows that, contradictory to the popular press, the rise of property prices after 2001 is accompanies by a squeeze out of bubble. That is the recent rise in property price in Seoul is due to improvement in economic fundamental rather than speculative bubbles.



The estimated bubble does appear to be an important driving force for the prices. Refer to figure 5.4. In general, percentage changes in bubble move closely together with percentage changes in prices. In times (April 1991, June 1999 and August 2000) the change in bubble far exceeds the change in price, reflecting, possibly, increased noise in market information in those time periods.

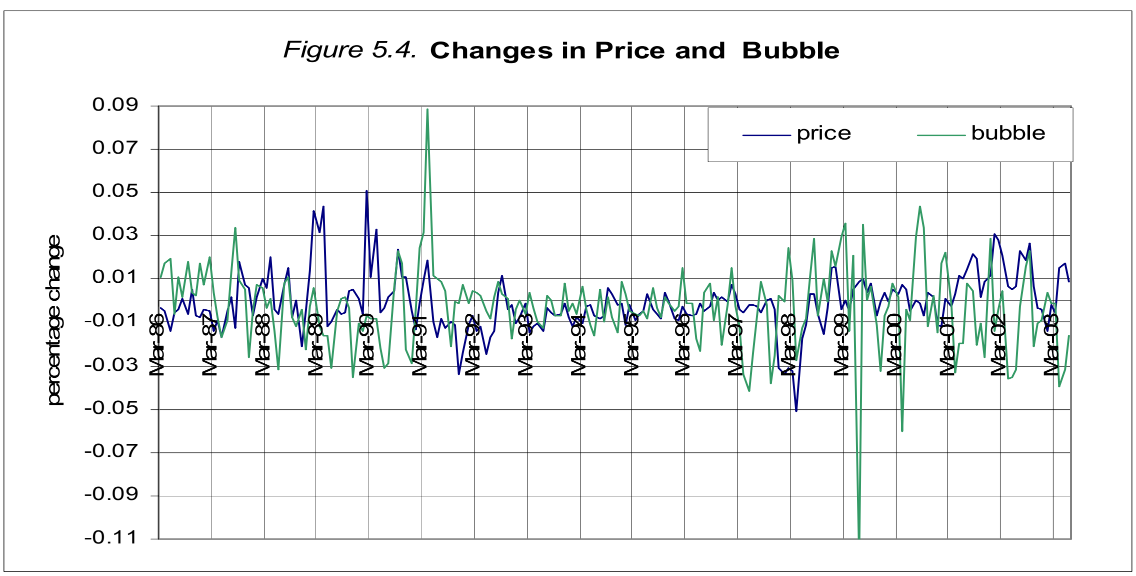


Table 5.2 compares the results of the state space model with the results from the simple present value model without bubble and from the AR(1) model. The simple present value model has the best in-sample performance in terms of sum of squared errors. But the AR(1) model is the best, speaking of forecasting accuracy. Overall the state space model under-perform both simple PV model and AR(1) model.

Table 5.2. **Model Comparison**

	State space model	Simple present value model	AR(1) model
In sample sum of squared error	*0.0787*	*0.0173*	*0.0235*
8-step forecast sum of squared errors	*0.0013*	*0.0010*	*0.0007*

6. Conclusion

This paper tries to find out the importance of bubble, estimated as an unobservable state variable using Kalman filter, in driving the property prices of Seoul, Korea. Based on the model introduced in section 2.1, we found that speculative bubble appears to be an important factor driving the Seoul property prices. But performance of the model is not entirely satisfactory, compared with the simple present value model with no bubble component. When forecasting performance is concerned, the AR(1) model does even a better job. This is indeed not surprising given the characteristics of the price and the rent series. Both series move closely together showing strong positive correlations, hence the simple present value model should do a reasonable job. These series also exhibit positive serial correlations in that rising prices and rents tend to cause further increases in prices and rents and vise versa, indicating AR(p) model might be useful in forecasting future prices.

We have a few comments before ending the paper. First, in our paper, the unobserved state variable is interpreted as speculative bubble, which may in fact represent some economic fundamentals unobserved by the researcher. Secondly, as argued in section one, due to the lengthy lags in bringing substitute assets into the market, excess rents might be earned for sustained period of time, which superficially justifies high property prices, but fail to reflect changes in economic fundamentals. For example, the twin rises in rent and price in late 1980s and in the period after 1999 could indeed be caused by speculative bubbles, but such bubble will not be captured by our model.

The unsatisfactory performance in our model, in terms of in-sample and out-of-sample sum of squared error, might have been caused by the existence of structural break. Future investigation can incorporate such fact and allow the model parameters to take on different values in different regimes. Furthermore, it is unclear, if there is any information loss in differencing the time series before estimation. Future studies might base the model on the level series instead, using appropriate techniques available in the literature (Gomez and Maravall, 1994).

List of Figures

List of Tables

Bibliography

1. Abraham, J. M. and Hendershott, P. H. “bubbles in metropolitan housing markets” *NBER working paper no. 4774,(June, 1994).*

2. Ansley, Craig F. and Kohn, Robert, “a geometrical derivation of the fixed interval smoothing algorithm”, Biometrika, vol, 69, No. 2 (Aug., 1982), 486-487.

3. Basile, A. and Joyce, J. P. “asset bubbles, monetary policy and bank lending in Japan: an empirical investigation” Applied Economics, 2001, 33, 1737-1744

4. Blanchard O. J. and Watson, M. W. “bubbles, rational expectations and financial markets” *NBER working paper no. 945 (Jul., 1982)*

5. Burmeister, Edwin and Kent D. Wall, “Kalman filtering estimation of unobserved rational expectations with an application to the German hyperinflation”, *Journal of Econometrics, January 1982, 20: 255-284.*

6. Campbell, J. Y. and Shiller R. J. “cointegration and tests of present value models” *The Journal of Political economy, Vol. 95, Iss. 5 (Oct., 1987), 1062-1088)*

7. Campbell, J. Y. and Shiller R. J., “stock prices, earnings, and expected dividend”, *The Journal of Finance*, vol. 43, No. 3, Papers and Proceedings of the Forty-seventy Annual Meeting of the American Finance Association, Chicago, Illinois, December 28-30, 1987 (Jul., 1988), 661-676.

8. Campbell, J. Y. and Shiller R. J., “the dividend-price ratio and expectations of future dividends and discount factors”, *The Review of Financial Studies, vol. 1, No. 3 (Autumn, 1988), 195-228.*

9. Diba, B. T. and Grossman H. I., “the theory of rational bubbles in stock prices”, *The Economic Journal,* Vol. 98, No. 392 (Sep., 1988), 746-754.

10. Diba, B. T. and Grossman, H. I. "explosive rational bubbles in stock prices?" *The American Economic Review*, Vol. 78, Iss. 3, 520-530 (Jun., 1988).

11. Engsted, T. and Tanggaard, C. "a new test for speculative bubbles based on return variance decompositions" *Aarhus School of Business, Department of Finance, working paper 01-9*

12. Evans, G. W. "a test for speculative bubbles in the sterling-dollar exchange rate: 1981-84", *The American Economic Review, Vol. 76, Iss. 4 (Sep., 1986), 621-636*

13. Flood, R. P. and Garber, P. M. "market fundamentals versus price-level bubbles: the first tests" *The Journal of Political Economy, Vol. 88, Iss. 4 (Aug., 1980), 745-770*

14. Ghahramani, Zoubin and Roweis, Sam T., "learning nonlinear dynamical systems using an EM algorithm", *Gatsby computational neuroscience unit, University College London*

15. Gomez, V and Maravall, A. "estimation, prediction, and interpolation for nonstationary series with the Kalman filter", *Journal of the American Statistical Association, Vol. 89, No. 426 (Jun., 1994), 611-624*

16. Hall, R. "struggling to understand the stock market---the 2001 ely lecture" *NBER (2001)*

17. Hall, S. G., Psaradakis, Z. and Sola, M. "detecting periodically collapsing bubbles: a markov-switching unit root test" *Journal of Applied Econometrics, 14, 143-154 (1999)*

18. Hamilton, James D., "state space model", *Handbook of Econometrics, vol.4*

19. Harvey, Andrew C., "forecasting, structural time series models and the Kalman filter", *Cambridge University Press, 1990.*

20. Hendershott, P. H. "property Asset Bubbles: Evidence from the Sydney office market" *Journal of Real Estate Finance and Economics, 20:1, 67-81 (2000)*

21. Hendershott, P. H, Hendershott, R.J. and Ward, C. R. W. "corporate equity and commercial property market 'bubbles'" *Urban studies, Vol. 40, Nos 5-6, 993-1009, 2003*

22. Herrera, S. and Perry, G. "tropical bubbles: asset prices in Latin America, 1980-2001", *World Bank working papers 2724*

23. Ito, T. and Iwaisako, T. "explaining asset bubbles in Japan" *NBER working paper 5358, (1995)*

24. Jong, Piet De, "the diffuse Kalman filter", *The Annals of Statistics, vol. 19, No. 2 (Jun., 1991), 1073-1083.*

25. Meinhold, Richard J. and Singpurwalla, Nozer D., "understanding the Kalman filter", *The American Statistician, vol. 37, No.2, (May, 1983), 123-127.*

26. Meinhold, Richard J. and Singpurwalla, Nozer D., "robustification of Kalman filter models", *Journal of the American Statistical Association, vol. 84, No. 406 (Jun., 1989), 479-486.*

27. Shiller R. J. "do stock prices move too much to be justified by subsequent changes in dividends?" *The American Economic Review, Vol. 71, Iss. 3 (Jun., 1981), 421-436*

28. Shiller, R. J. "market volatility and investor behavior" *The American Economic Review, Vol. 80, Issue 2, (May, 1990), 58-62*

29. Stoffer, David S. and Wall, Kend D. ‘bootstrapping state-space models: Gaussian maximum likelihood estimation and the Kalman filter” *Journal of the American Statistical Association, Vol 86, No.416 (Dec., 1991)*

30. West, K. D. “a specification test for speculative bubbles”, *The Quarterly Journal of Economics, Vol. 102, Iss. 3 (Aug., 1987), 553-580.*

31. Wheatley, G. “applied numerical analysis”, Pearson Education, Inc. 2004, 7th edit.

32. Wu, Yangru, “are there rational bubbles in foreign exchange markets? evidence from an alternative test”, *Journal of International Money and Finance, Vol. 14, No.1, 27-46, 1995.*

33. Wu, Yangru, “rational bubbles in the stock market: accounting for the U. S. stock-price volatility”, *Economic Inquiry, Iss. 0095-2583, vol. xxxv, April 1997, 309-319*