Price Discovery in Time and Space: The Course of Condominium Prices in Singapore

by

Min Hwang
University of California
Berkeley, CA
min@econ.berkeley.edu

John M. Quigley
University of California
Berkeley, CA
quigley@econ.berkeley.edu

This draft: March 22, 2002

ABSTRACT: There is increasing evidence that aggregate housing price are predictable. Despite this, a random walk in time and independence in space are two maintained hypotheses in the empirical models for housing price measurement used by government agencies and by commercial companies as well. This paper examines the price discovery process in individual dwellings over time and space by relaxing both assumptions, using a unique body of data from the Singapore private condominium market. We develop a model that tests directly the hypotheses that the prices of individual dwellings follow a random walk over time and that the price of an individual dwelling is independent of the price of a neighboring dwelling. The model is general enough to include other widely used models of housing price determination, such as Bailey, Muth, and Nourse (1963), Case and Shiller (1987) and Redfearn and Quigley (2000), as special cases. The empirical results clearly support mean reversion in housing prices and also diffusion of innovations over space. Our estimates of the level of housing prices, derived from a generalized repeat sales model, suggest that serial and spatial correlation matters in the computation of price indices and the estimation of price levels. The finding of mean reversion may suggest that housing prices are forecastable and that excess returns are possible for investors. We use the monthly price series derived from condominium sales to investigate this issue. We compute gross unleveraged real returns monthly. When returns are computed from models which assume a random walk without spatial autocorrelation, we find that they are strongly autocorrelated. When returns are calculated from more general models that permit mean reversion, the estimated autocorrelation in investment returns is reduced. Finally, when they are calculated from models permitting mean reversion and spatial autocorrelation, predictability in aggregate investment returns is completely absent.

JEL codes: E31, C23, R32

A previous version of this paper was presented at the ASSA Annual Meetings, Atlanta, GA, January 2002.

I. Introduction

The durability, fixity and heterogeneity of dwellings imply that transaction costs are significant in the housing market. Certainly in comparison to financial markets, and in comparison to the markets for most consumer goods, housing purchases require costly search to uncover the prices and attributes of commodities. Given the many frictions associated with the purchase of housing, it is hardly surprising that price behavior deviates from that predicted by simple models of economic markets.

Case and Shiller (1990) report that both real and excess returns in the housing market were forecastable; subsequently several other researchers (for example, Guntermann and Norrbin, 1991; Gatzlaff, 1994; and Malpezzi, 1999) have documented predictable returns in housing markets by demonstrating that aggregate price series exhibit inertia in percentage changes. Less is known about the dynamics of house prices at the individual level. Englund, Gordon, and Quigley (1999) and Quigley and Redfearn (2000), using very different techniques, rejected a random walk in individual housing prices by examining repeat sales of single family dwellings. This suggests that the inertia reported in the aggregate may also characterize micro behavior.

But in this geographical market, price signals exist in space as well as time. Many of the features which can lead to autocorrelation in the time domain could have analogous effects over space. Price information diffuses over space as well as time, and information costs alone can cause prices to deviate from random fluctuations.

This paper examines price discovery over time in a spatial market using a body of data almost uniquely suited to the problem. We examine the prices of condominium dwellings in Singapore using all sales reported in the country during an eleven-year period. Multiple sales of the same condominium unit are observed, and all dwellings with market transactions are geocoded. We develop a model of housing prices that more faithfully represents the temporal and spatial features unique to housing markets, and we incorporate a more general and more appropriate structure of prices at the level of the individual dwelling.

The model and the data support a direct test of the hypotheses that the prices of individual dwellings follow a random walk over time and that the price of an individual dwelling is independent of the price of a neighboring dwelling. We link these results to movements in aggregate measures of housing prices and their spatial and temporal properties.

The model is more general than other widely used methods of measuring aggregate housing prices. Indeed, the method used by government agencies (e.g., OFHEO) and commercial firms (e.g., MRAC, Inc.) to estimate the course of house prices is a special case of the model developed below. The framework presented supports tests of the assumptions implicit in more conventional models.

There are a few studies that use spatial econometric methods in analyzing housing prices, but none of them are based on a theory of price diffusion. For example, Can and Megbolugbe (1997) estimated hedonic house price models incorporating lagged values of neighborhood house prices to reflect spatial dependencies in prices. Goetzmann and Spiegel (1997) developed a "distance-weighted-repeat-sales procedure," where distance is defined in terms of geographical and socio-economic factors (such as neighborhood income, education attainment and racial composition) and where "distance weights" are estimated using an ad hoc procedure. Dubin (1998) postulated a specific form for a correlogram relating the correlation between housing prices as a function of distance. The estimates of an empirical correlogram were used to in hedonic models of housing prices to reflect spatial dependencies. A more sophiscated hedonic model relying upon empirical semivariograms was estimated by Basu and Thibodeau (1998).

Pace et al (1998) developed an empirical model for house prices which evolves thorough time and space. Their model specified an autoregressive structure of house prices and a spatial dependency among prices. Given an irregular panel of house prices (in which there are few transactions in any period), ad hoc procedures were used to filter house price sales by time and location. (Indeed, different results are obtained depending upon the ordering of filtering process.)

Reliance upon ad hoc procedures to analyze the spatial and temporal pattern of housing prices is understandable, given the infrequency of transactions on dwellings. This means that a panel of houses typically contains a relatively small and irregular number of observations on the sales prices of these houses. The temporal correlation in prices depends upon the time interval between sales, and with irregular intervals, inference in a model which also accounts for spatial dependence may be quite difficult. (See, for example, Pace et al, 1998: 18-22.)

The model developed in this paper employs an explicit model of the spatial and temporal dependence of housing prices, and estimates the importance of spatial and temporal factors in the estimation of the course of housing prices. We do this using a repeat sales model of price determination, not a hedonic model. Not surprisingly, the introduction of an explicit micro model presents certain computational difficulties in estimation.

We devote considerable attention to the implications of our statistical findings for investment in the housing markets. In particular we demonstrate the importance of these findings for investor returns using a variety of investment rules. We also investigate the link between spatial and temporal dependence in prices and aggregate indices of house prices.

Section II develops a general model of housing prices that supports explicit tests for the spatial and temporal pattern of price movements. This section links our model to the widely employed method for measuring housing prices proposed almost forty years ago by Bailey, Muth, and Nourse (1963), as well as its subsequent extensions (e.g., Case and Shiller, 1987). The data are described tersely in Section III. Our empirical results are presented in Sections IV and V. We test for random walks in space and time, against the alternative of mean reversion, and we examine the link between pricing deviations at the individual level and aggregate price movements. We also investigate investor behavior in some detail. Section VI is a brief conclusion.

II. A Micro Model of House Prices

The objects of exchange in the housing market are imperfect substitutes for one another. Indeed, the fixity of housing implies that dwellings with identical physical attributes may differ in price simply because the price incorporates a complex set of site-specific amenities and access costs. But few dwellings have identical physical characteristics; thus comparison shopping is more difficult and more expensive than in most other markets.

Moreover, housing transactions are made only infrequently, so households must consciously invest in information to participate in this market. As a result, the market is characterized by a costly matching process. Market agents, buyers and sellers, are heterogeneous and differ in information and motivation; commodities are themselves heterogeneous. Consequently an observed transaction price for a specific unit may deviate from the price ordained in the fully informed perfect market of the intermediate micro textbook.

Buyers, sellers, appraisers, and real estate agents estimate the “market price” of a dwelling by utilizing the information embodied in the set of previously sold dwellings. The usefulness of these sales as a reference depends upon their similarity across several dimensions: physical, spatial, and temporal. Inferences about the “market price” of the dwelling can be drawn only imperfectly from the set of past sales, because dwellings differ structurally, enjoy different locational attributes, and are valued under different market conditions by different actors over time. Because housing trades infrequently, the arrival of new information about market values is slow. From an informational standpoint, the closest comparable sale across these various dimensions may be the last sale of the same dwelling. Alternatively, the closest comparable sale may be the contemporaneous selling price of another dwelling in close physical proximity.

An attempt to uncover the market value of a dwelling is further complicated by the fact that an observed sales price is not only a function of observable physical characteristics, but also of unobserved buyer and seller characteristics such as their urgency to conclude a transaction (Quan and Quigley, 1991). For any given sale, all that is known is that an offer was made by a specific buyer that was higher than a specific seller’s reservation price.

We develop a model with spatially and temporally correlated errors in a repeat sales framework. Innovation processes over time are assumed to be continuous, but sales are obtained at irregular intervals. At any point in time, the prices of houses are dependent over space. In the determination of the price of a house, the weights attributable to neighboring houses are fixed and depend upon their distances from that house. Again, the prices of neighboring houses are observed infrequently.

Let the log sale price of dwelling i at time t be

(1) $V_{it} = P_t + Q_{it} + e_{it} = P_t + X_{it}\beta + e_{it}$,

where V_{it} is the log of the observed sales price of dwelling i at t, and P_t is the log of aggregate housing prices. Q_{it} is the log of housing quality, and can be parameterized by X_{it}, the set of housing attributes and by a set of coefficients, β, which price those attributes. If a sale is observed at two points in time, t and τ, and if the quality of the dwelling remains constant during the interval, then

(2) $$V_{it} - V_{i\tau} = P_t - P_\tau + (X_{it} - X_{i\tau})\beta + e_{it} - e_{i\tau}$$

$$= P_{it} - P_{i\tau} + e_{it} - e_{i\tau}.$$

With constant quality, (2) identifies price change in the market.

Let the error term, e_{it}, consist of two components that are realized for each individual dwelling at the time of sale: η_{it}, an idiosyncratic innovation without persistence, and ε_{it}, an idiosyncratic innovation with persistence, $\varepsilon_{it} = \lambda\varepsilon_{i,t-1} + \mu_{it}$. In addition, assume that the value of any particular dwelling depends also on innovations that occur to other dwellings contemporaneously. We assume this spatial correlation depends on the distance between units.

(3) $$e_{it} = \rho\sum_{j=1}^{N} w_{ij}e_{jt} + \xi_{it} = \rho\sum_{j=1}^{N} w_{ij}e_{jt} + \varepsilon_{it} + \eta_{it} = \rho\sum_{j=1}^{N} w_{ij}e_{jt} + \lambda\varepsilon_{i,t-1} + \eta_{it} + \mu_{it}$$

where w_{ij} is some function of the distance between unit i and j and N is the number of dwellings in the economy. Let $E(\eta_{it}\eta_{jt}) = 0$ and $E(\varepsilon_{it}\varepsilon_{jt}) = 0$, $E(\eta_{it}^2) = \sigma_\eta^2$, $E(\mu_{it}^2) = \sigma_\mu^2$.

The value of a particular dwelling depends, not only on its own past and contemporaneous innovations, but also on innovations of other dwellings, past and contemporaneous. Note that the model of housing prices in (2) and (3) specializes to that of

Bailey, Muth and Nourse (1963) when $\lambda = \rho = 0$, and to that of Case and Schiller (1987) when $\lambda = 1$, $\rho = 0$, and to that of Quigley and Redfearn (2000) when $\rho = 0$.

In vector notation, expression (3) is

(4) $\mathrm{e_t} = \rho \mathrm{We_t} + \xi_\mathrm{t}$,

where $\mathbf{e}_t$ is a vector of e_{it} for all the dwellings, $\mathbf{W}$ is a weight matrix, some measure of the distance between dwellings, and ξ_t a vector of $\xi_{it} = \lambda \varepsilon_{i,t-1} + \eta_{it} + \mu_{it}$, for all the dwellings. By solving for $\mathbf{e}_t$ and taking the difference between two sales at times *t* and *s*, we have

(5) $\mathrm{e_t} - \mathrm{e_s} = (\mathrm{I} - \rho \mathrm{W})^{-1}(\xi_\mathrm{t} - \xi_\mathrm{s})$.

The variance-covariance matrix of (5) is

(6) $$E\left[(\mathbf{e}_t - \mathbf{e}_s)(\mathbf{e}_t - \mathbf{e}_s)'\right] = (\mathbf{I} - \rho\mathbf{W})^{-1} E\left[(\xi_t - \xi_s)(\xi_t - \xi_s)'\right](\mathbf{I} - \rho\mathbf{W})^{-1}$$

Transactions on dwellings occur only irregualrly. Consider the covariance in errors between a dwelling *i* sold at *t* and *s* and another dwelling *k* sold at τ and ς, $E\left[(e_{it} - e_{is})(e_{k\tau} - e_{k\varsigma})\right]$. Let $\mathbf{\Psi} = E\left[(\xi_t - \xi_s)(\xi_\tau - \xi_\varsigma)'\right]$ and $\mathbf{\Pi} = (\mathbf{I} - \rho\mathbf{W})^{-1}$. Thus,

(7) $$E\left[(\mathbf{e}_t - \mathbf{e}_s)(\mathbf{e}_\tau - \mathbf{e}_\varsigma)'\right] = \mathbf{\Pi\Psi\Pi} = \begin{bmatrix} \boldsymbol{\pi}_1' \\ \boldsymbol{\pi}_2' \\ \vdots \\ \boldsymbol{\pi}_N' \end{bmatrix} [\boldsymbol{\psi}_1 \ \boldsymbol{\psi}_2 \cdots \ \boldsymbol{\psi}_N][\boldsymbol{\pi}_1 \ \boldsymbol{\pi}_2 \cdots \ \boldsymbol{\pi}_N].$$

The elements of this expression are,

(8) $$E\left[(e_{it} - e_{is})(e_{k\tau} - e_{k\varsigma})\right] = \boldsymbol{\pi}_i' \mathbf{\Psi} \boldsymbol{\pi}_k.$$

Now consider an element of the covariance matrix, $\mathbf{\Psi}$. Note that

(9) $$E(\xi_{it}\xi_{j\tau}) = \lambda^{|t-\tau|}\left(\frac{\sigma_\mu^2}{1-\lambda^2}\right) + I(t = \tau)\sigma_\eta^2, \text{ if } i = j$$

$= 0$, otherwise.

where $I(\bullet)$ is an indicator function. For sales of a given dwelling at time t, s, τ and ς,

$$(10)\qquad E\left[(\xi_{it}-\xi_{is})(\xi_{i\tau}-\xi_{i\varsigma})\right]=\left(\lambda^{|t-\tau|}-\lambda^{|t-\varsigma|}-\lambda^{|s-\tau|}+\lambda^{|s-\varsigma|}\right)\left(\frac{\sigma_\mu^2}{1-\lambda^2}\right)$$

$$+\left(I(t=\tau)-I(t=\varsigma)-I(s=\tau)+I(s=\varsigma)\right)\sigma_\eta^2 .$$

Therefore, the variance-covariance matrix is

$$(11)\qquad \boldsymbol{\Psi}=E\left[(\xi_t-\xi_s)(\xi_\tau-\xi_\varsigma)'\right]=E\left[(\xi_{it}-\xi_{is})(\xi_{i\tau}-\xi_{i\varsigma})\right]\times\mathbf{I} .$$

Finally, the variance-covariance matrix of innovations between a dwelling *i* sold at *t* and *s* and another dwelling *k* sold at τ and ς is,

$$(12)\qquad E\left[(e_{it}-e_{is})(e_{k\tau}-e_{k\varsigma})\right]=\boldsymbol{\pi}_i'\,\boldsymbol{\Psi}\,\boldsymbol{\pi}_k=\boldsymbol{\pi}_i'\left\{E\left[(\xi_{it}-\xi_{is})(\xi_{i\tau}-\xi_{i\varsigma})\right]\times\mathbf{I}\right\}\boldsymbol{\pi}_k$$

$$=\left\{\left(\lambda^{|t-\tau|}-\lambda^{|t-\varsigma|}-\lambda^{|s-\tau|}+\lambda^{|s-\varsigma|}\right)\left(\frac{\sigma_\mu^2}{1-\lambda^2}\right)+\right.$$

$$\left.\left[I(t=\tau)-I(t=\varsigma)-I(s=\tau)+I(s=\varsigma)\right]\sigma_\eta^2\right\}\boldsymbol{\pi}_i'\boldsymbol{\pi}_k .$$

Equation (12) indicates how the variance-covariance matrix of residuals from the regression specified in (2) can be used to identify the temporal and spatial components of house price persistence, λ and ρ, respectively. Identification requires observing at least two transactions for each dwelling and observing the distance of each dwelling from all others in the market.

III. Data

The data uitilized in this paper have been compiled by the Singapore Institute of Surveyors and Valuers (SISV) and consist of all transactions involving dwellings in multifamily housing during the period from Jan 1, 1990 to Dec 31, 2000. SISV gathers transactions data from a variety of sources including legal registration records and developers' sales records. The dataset is complete – each condominium sale in the entire country is recorded. In addition, an extensive set of physical characteristics of dwellings is recorded. The date of the sale is recorded as well as the date of occupancy. In addition, the address, including the postal code, is reported. The postal code identifies the physical location, often the specific building. A matrix of distances among Singapore's fifteen hundred postal codes permits each dwelling to be located spatially. The data set includes transactions among dwellings in the standing stock, sales of newly constructed dwellings, and presales of dwellings under construction (where the contract date may be several months before the date construction is completed).

The panel nature of the data permits us to distinguish dwellings sold more than once, and the multiple sales feature of the data identifies the models specified in section II. By confining the sample to dwellings in multifamily properties, we eliminate types of dwellings for which additions and major renovations are feasible. The sample of multifamily dwellings is thus less likely to include those for which the assumption of constant quality between sales is seriously violated.

Singapore data offer another advantage in estimating the model of housing prices, namely a spatial homogeneity of local public services (e.g., police protection, neighborhood schools), especially when compared to cities of comparable size in North America. During the decade of the 1990s, there was no discernible trend in the quality of neighborhood attributes of the bundle of housing services.[1]

Table 1 presents a summary of the repeat sales data used in the empirical analysis reported below. There are several points worth noting in the table. First, confirming the

[1] One possible exception to this may be accessibility, where improvement in the transport system and its pricing may have altered the workplace access of various neighborhoods.

infrequency of housing transactions, the number of dwellings sold more than once is less than twenty percent of the population of the dwellings sold during the eleven year period. Only three percent of the 52,337 dwellings were sold more than twice in the eleven year period.

Second, the average selling prices tend to be higher for dwellings sold more frequently. The rate of appreciation is also higher. On average, dwellings sold five times appreciate almost twice as fast as dwellings sold only twice. For the dwellings sold more frequently, price appreciation tends to be more volatile. Transactions involving high-turnover dwellings are apparently riskier, but this risk is compensated by higher returns.

Third, the intervals between sales are longer for dwellings sold infrequently. In part, this is an artifact of the fixed sampling framework. For presold dwellings, the average length of time between sale and completion of construction is highest for those sold least frequently, which is not consistent with popular belief that presales are associated with speculation in the housing market.

Fourth, there are some differences in the characteristics of dwellings sold more frequently. They tend to be larger in area, containing more rooms, and they are more centrally located to the CBD, but their transit access is similar to the dwellings sold less frequently.

The data on condominium sales supports a regression model of the form

(13) $$V_{it} - V_{is} = P_t D_{it} - P_s D_{is} + \gamma\kappa_{it} - \gamma\kappa_{is} + e_{it} - e_{is},$$

where D_{it} is a variable with a value of 1 for the month *t* in which condominium *i* is sold and zero in other months and P_t is the estimated coefficient for this variable. There are 132 of these time variables, one for each month between 1990 and 2000. If dwelling *i* has been presold, κ_{it} is the time interval between the transaction and the completion of construction. For dwellings sold after completion of construction, κ_{it} is set to zero. Thus, γ, the estimated coefficient for

κ_{it}, measures the monthly discount rate for presold dwellings, i.e., the discount for unrealized service flows from presold dwellings. The purchase of a dwelling before completion, or even before construction, is not uncommon in Singapore. One aspect of this institution may, however, be uncommon – namely that the entire purchase price is paid at the time the contract is signed, not at the time the dwelling is first occupied.

Of the 11,883 pairs of transactions noted in Table 1, 305 consist of presale pairs. For another 5,024 pairs, the first sale was made before the property was completed.

IV. The Diffusion of House Price Innovations

The model can be estimated by maximum likelihood methods. In particular, if we assume the error terms in equation (3), η_{it} and μ_{it}, are normally distributed, the log likelihood function for the observed sample of condominium sales is

(14) $$\log L\left(P,\gamma,\lambda,\rho,\sigma_\eta^2,\sigma_\mu^2\right) = -\log\left(\left|\boldsymbol{\Sigma}\right|\right) - \left(\delta'\boldsymbol{\Sigma}^{-1}\delta\right),$$

where $\boldsymbol{\Sigma} = \left[\boldsymbol{\pi}_i'\,\boldsymbol{\Psi}\,\boldsymbol{\pi}_k\right]$ and $\delta = V_{it} - V_{is} - P_t D_{it} + P_s D_{is} - \gamma\kappa_{it} + \gamma\kappa_{is}$.

Note that the parameters in the $\boldsymbol{\Sigma}$ matrix are λ, ρ, σ_μ^2 and σ_η^2. Conditional on values for λ and ρ, the consistent estimates of the error variances, σ_μ^2 and σ_η^2, can be obtained from the regression

(15) $$\left(\hat{e}_{it} - \hat{e}_{is}\right)^2 = 2\left(1-\lambda^{t-s}\right)\left(\frac{\sigma_\mu^2}{1-\lambda^2}\right) + 2\sigma_\eta^2,$$

where the vector $\left(\hat{e}_{it} - \hat{e}_{is}\right)$ is the set of residuals from a first-stage regression. Then, the remaining parameters of the repeat sales model, equation (13), can be estimated by generalized least squares. The vector of residuals, δ, and the matrix $\boldsymbol{\Sigma}$, computed from λ and ρ, are sufficient to compute the values of the log likelihood function. The function can be maximized by a grid search over λ and ρ.

Consider the matrix $\mathbf{\Sigma}$. It is the product of three submatrices, $\boldsymbol{\pi}_i'$, $\mathbf{\Psi}$ and $\boldsymbol{\pi}_k$. The matrix $\mathbf{\Psi}$ is large, with rows and columns equal to the number of dwellings in the sample, but it is diagonal. The elements of the matrix are computed from the time intervals between sales, given λ, according to Equation (10). Absent spatial correlation, i.e., when $\rho = 0$, the matrix $\boldsymbol{\pi}_i$ is the *i*-th vector of the identity matrix. So is $\boldsymbol{\pi}_k$. Thus the matrix $\mathbf{\Sigma}$ is block diagonal. The size of each block is determined by the number of paired sales for a given house. Using a variety of techniques for large sparse matrices, the inverse, $\mathbf{\Sigma}^{-1}$, can be computed.

As noted in Section I, there is ample reason to expect mean reversion in house prices. We begin by assuming no spatial dependence and analyze antocorrelation. Figure 1A presents the maximized value of log likelihood function, equation (14), assuming $\rho = 0$, and hence the matrix $\mathbf{\Sigma}$ is block diagonal and sparse. The estimation is based on 11,883 observations on repeat sales on 10,288 dwellings sold two or more times. The likelihood function is well behaved with a maximum at $\lambda = 0.72$. Likelihood ratio tests reject a random walk in house prices ($\lambda = 1$) and serially uncorrelated house prices ($\lambda = 0$) by a wide margin, $\chi^2 = 2{,}489.68$, and $\chi^2 = 32{,}129$ respectively. The estimated value of λ suggests that the half life of a one unit shock to housing prices is about 33 days.[2]

We now estimate the parameters of spatial and temporal autocorrelation simultaneously. As noted above, when $\rho \neq 0$, the $\mathbf{\Sigma}$ matrix is no longer block diagonal. The appendix illustrates the nature of the numerical problems encountered. One way of addressing these problems is to

[2] This half life is considerably shorter than the values reported by Quigley and Redfearn (2000) in analogous quarterly models of the price movements of single family housing in eight Swedish housing markets.

note that $\boldsymbol{\pi}_i$ is the *i*-th row of $(\mathbf{I}-\rho\mathbf{W})^{-1}$, and when $\mathbf{W}$ is sparse, most of elements of $\boldsymbol{\pi}_i$ will be zeros. This, in turn, will make $\boldsymbol{\Sigma}$ matrix sparse, since $\boldsymbol{\Sigma} = [\boldsymbol{\pi}_i' \boldsymbol{\Psi} \boldsymbol{\pi}_k]$. One inconsequential way of making $\mathbf{W}$ sparse is to set small values of weights to zero, implying that when two dwellings are sufficiently far apart, then there is no spatial correlation between them. In the following, we assume that the elements of the weight matrix are the reciprocals of the distance between dwellings and that dwellings further than 250 meters apart are not spatially correlated.

Figure 1B presents the ML estimates of the likelihood values for different values of $\boldsymbol{\lambda}$ and $\boldsymbol{\rho}$. The values of $\boldsymbol{\lambda}$ and $\boldsymbol{\rho}$ that maximize the log-likelihood values are 0.78 and 0.55, respectively. The ML estimate of the serial correlation coefficient, $\boldsymbol{\lambda}$, is rather similar to that reported in Figure 1A, but the half life of a unit shock is now estimated to be 53 days, more than 60 per cent longer. The value of 0.55 for the spatial correlation coefficient, $\boldsymbol{\rho}$, is quite modest. Figure 2 shows the contemporaneous impact of a unit shock over a grid where the distance between houses is 30 meters, when the spatial correlation coefficient is 0.55. This illustrates how a unit shock at point 0 diffuses over the grid. The impulse quickly dissipates over the space; most of the impulse completely dissipates within 100 meters.

Appendix Table 1 presents estimates of the price index, Equation (13), under different assumptions on the error structure. Price index estimates for early years tend to be insignificantly different from each other while those for later years are significant. In part, this arises from the sampling design: there are more observations for later years, which allows more precise estimation of coefficients for later years. Among three indices, the two that allow stationary processes for error terms tend to move more closely. Figure 3A, B and C report the estimated price indices with January of 1990 as a basis year. The three indices generate a similar course of aggregate prices for private condominiums in Singapore during the period.

The estimated coefficients in Appendix Table 1 for the period between sale and dwelling completion (for presold units) are around 11 basis points; this is between a 1.2 percent and 1.5 percent discount for an incomplete dwelling unit sold today for occupancy a year hence. The magnitude of the discount is not trivial: aggregate housing prices rise, on average, by 0.3 percent monthly and the presales discount reduces monthly increases by one third.

V. The course of Condominium Prices and Investment Returns

Figure 3A through 3C present apparently similar patterns for the course of housing prices for Singapore dwellings during the period 1990-2000. However, a closer examination of returns implied by these housing indices reveals substantial differences among them. In other words, while the different assumptions about the error structure do not yield substantial differences in estimated prices, the real return series implied by the estimates are quite different.

The economic returns from investment in housing depend upon the course of real prices and rents. In particular, ignoring transaction costs and leverage, the real return in any period, R_t, is the change in the value plus the dividend (i.e. the rental stream, r_t, enjoyed during the period)

(16) $$R_t = \left(\frac{P_t + r_t}{P_{t-1}}\right)\left(\frac{I_{t-1}}{I_t}\right),$$

where I_t is an index of the cost of living, less housing.

Figure 4 uses the estimates presented in Figure 3 and the monthly CPI in Singapore to chart the course of investment returns during the eleven-year period. The estimated returns are strikingly different. Table 2 reports the forecastability of investment returns. There is no apparent trend in the data. Tables 2A, B and C report more explicit information on trends in real gross returns. The table reports the forecastability of monthly returns based upon lags of returns of one, two, three and four months.

As reported in the table, there is considerable disparity in the forecastibility of returns estimated by the three procedures. With a random walk and no spatial correlation (Table 2C), there is a considerable evidence of overshooting in monthly returns, so a contrarian investment policy would maximize investment returns: sell on price increases, buy on price decreases. There is no evidence that a more complicated lag structure improves the forecastibility of investment returns. With mean reversion but no spatial autocorrelation, there is again evidence of overshooting, and also weaker evidence that a more complicated lag structure improves forecastability. Using the maximum likelihood estimates, (λ=0.78; ρ=0.55) there is no evidence of forecastability in aggregate house prices at all. There is no predictability in aggregate returns.

V. Investment performance

These results may have significant implications for investment in the housing market. Consider an investment decision in housing based on housing price determination models such as (1). In this context, a better specification of error structure can lead to superior investment decisions in two ways. First, improvement comes through better estimates of aggregate housing price trends. In the regression models graphed in Figure 3, different assumptions about error structure have relatively small effects on the large sample properties of slope coefficients, but they do have large effects on efficiency of those parameters. Therefore, investment decisions based on the correct error structure are more important when investment horizons are relatively short. Second, additional improvement comes from basing the investment decision on more complete information. In other words, when errors are spatially correlated, knowledge of past and present innovations in neighboring dwellings may provide valuable information, useful for predicting the future course of prices for other dwellings. If one assumes there is no spatial correlation and does not use information from housing transactions in neighborhood, the investor may lose important information in making price forecasts.

This section highlights the consequences of different assumptions on error structures on measured investment performance in housing market. We use investment rules which depend upon forecasts of future housing returns. These forecasts depend on investor's assumptions about the underlying housing price generating processes.

The investment rule applied in this section is quite simple. Given assumptions on error structures and the consequent parameter values for underlying house price processes, we make forecasts for housing returns using all the available current information. The investor is instructed to "Buy" if the expected return is greater than some preset threshold. The threshold may be interpreted to as the known transaction costs in the housing market.

The sizes of actual transaction costs vary with housing market characteristics, financial market characteristics and tax systems and it is difficult to present a particular number as a universal estimate for the cost. We use 0 percent , 5 percent and 10 percent thresholds, comparable with a range of plausible transaction costs[3].

The investment holding period is set arbitrarily at 24 months.

When spatial correlations exist among dwellings, error distributions of individual dwelling prices are heterosckedastic since dwellings have different neighborhoods. Further, the variance-covariance matrix of error terms depends on distances to neighbors, so this varies across dwellings. In this exercise, we consider a housing development, where each dwelling is located at a point on a 51 by 51 grid, and each house is 30 meters away from its nearest neighbors. It is possible to consider the investment performance all 2601 dwellings, but for convenience, we chose the dwelling at the center of the town.

In the simulation, the time series prices of each individual dwelling are generated twice for 24 months by using the parameter values obtained in the maximum likelihood estimation along with an appropriate weight matrix. The first set of prices is assumed to be observed by the investor, who uses this information together with his estimates of parameter values to make a precise forecast for next 24 months. If the forecasted return is greater than the threshold, then he will buy the house. The second set of prices is then used to evaluate the performance of investment.

We consider three investors with differing amounts of information. Investor A is armed with the ML estimates that $\rho = 0.55$ and $\lambda = 0.78$ reported in Figure 1B. She uses this information together with 24 months of history on housing prices to generate a price forecast for the given house at the end of the next 24 months. Figure 5A through D report the probability distribution of the investor's returns. Part A reports the returns from the naïve rule: always invest. Part B reports the distribution of returns for the same investor using the rule: invest if the expected return exceeds zero. Part C reports the distribution under a more stringent five percent rule, and part D imposes a ten percent rule.

Investor B is armed with less complete information. Based upon the results reported in Figure 1A, she assumes $\rho = 0$ and $\lambda = 0.72$. Figure 6A through C report the distribution of returns to this investor using the 0, 5 and 10 percent rules respectively.

---

[3] For more thorough examination of such costs, see Söderberg (1995).

Finally, Figures 7A and B report the distribution of returns for an investor who assumes no spatial correlation and a random walk in house prices, i.e, $\rho = 0$ and $\lambda = 1$. The performance of this investor is almost as bad as that of the person who always invests.

It is quite clear that the best econometrician is the richest investor.

VI. Conclusions

Because of the special features of the housing market, we may anticipate that price discovery and the diffusion of price information is more complicated than in many other markets. In this paper, we test the departures from instantaneous diffusion of price information over time and space. Using information on all condominium sales in Singapore during an eleven year period, we test for random walks, mean reversion and serial correlation in house prices. We rely upon multiple sales of more than ten thousand dwellings over the period to analyze the structure of pricing errors.

Our empirical results quite clearly support mean reversion in house prices. Our statistical tests reject the hypothesis of a random walk and they also reject the hypothesis of no serial correlation against the alternative hypothesis of mean reversion. We also find significant spatial dependence in prices.

The maximum likelihood estimate of serial correlation, 0.78 per month, suggests rapid dissipation of any innovation in housing prices. After two months, about 39 percent of any mispricing error is dissipated (i.e., $1-.78^2$). After six months, 77 percent is dissipated, and after a year 98 percent is dissipated.

Our estimates of the level of housing prices, derived from the repeat sales model, do suggest that there are only small differences in the house price levels estimated when serial and spatial correlation is recognized. However, there are substantial differences in the estimated returns to housing investment.

The finding of mean reversion may suggest that housing prices are forecastable and that excess returns are possible for investors in this market. We use the monthly price series derived from condominium sales to investigate this issue. We compute gross unleveraged real returns monthly. In misspecified models, we do find evidence of a one period lag in real returns, i.e., real returns today are negatively related to real returns last month. When aggregate house prices are calculated from micro models that permit mean reversion and spatial autocorrelation, predictability in investment returns is completely absent.

Finally, we investigate the economic value of information about the spatial and temporal autocorrelation in house prices in affecting investment returns in the housing market. Our analysis suggests that recognition of spatial and temporal factors can substantially increase the returns to investment in the housing market.

References

Bailey, M. J., R. F. Muth and H. O. Nourse (1963), "A Regression Method for Real Estate Price Index Construction," *Journal of the American Statistical Association*, 58, 933-42.

Basu, S. and T. Thibodeau (1998), "Analysis of Spatial Autocorrelation in House Prices," *Journal of Real Estate Finance and Economics* 17, 61-85.

Can, A. and I. Megbolugbe (1997), "Spatial Dependence and House Price Index Construction," *Journal of Real Estate Finance and Economics*, 14, 203-22.

Case, K. E., and R. J. Shiller (1987), "Price of Single-Family Homes since 1970: New Indexes for Four Cities," *New England Economic Review*, 45-56.

________ (1990), "Forecasting Prices and Excess Return in the Housing Markets," *American Real Estate and Urban Economics Association Journal*, 18(3), 253-73.

Dubin, R. (1998), "Predicting House Prices Using Multiple Listings Data," *Journal of Real Estate Finance and Economics* 17, 35-59.

Englund, P., T. M.Gordon, and J. M. Quigley (1999), "The Valuation of Real Capital: A Random Walk Down Kungsgaten," *Journal of Housing Economics*, 8(3), 205-216.

Gatzlaff, D. H. (1994), "Excess Returns, Inflation, and the Efficiency of the Housing Markets," *Journal of the American Real Estate and Urban Economics Association*, 22(4), 553-81.

Goetzmann, W. and M. Spiegel (1997), "A Spatial Model of Housing Returns and Neighborhood Substitutability," *Journal of Real Estate Finance and Economics*, 14, 11-31.

Guntermann, K. L. and S. C. Norrbin (1991), "Empirical Tests of Real Estate Market Efficiency," *Journal of Real Estate Finance and Economics*, 4(3), 297-313.

Malpezzi, S. (1999), "A Simple Error Correction Model of Housing Prices," *Journal of Housing Economics*, 8(1), 27-62.

Pace, K., R. Barry, J. Clapp and M. Rodriquez (1998), “Spatiotemporal Autoregressive Model of Neighborhood Effects,” *Journal of Real Estate Finance and Economics* 17, 15-33.

Quan, D. C. and J. M. Quigley (1991), “Price Formation and the Appraisal Function in Real Estate Markets,” *Journal of Real Estate Finance and Economics*, 4(2), 127-46.

Quigley, J. M. and C. L. Redfearn (2000), “Investment Returns and Price Discovery in the Housing Market,” Paper presented at the Fifth Annual Conference of the Asian Real Estate Society, Beijing, July 2000.

Söderberg, B. (1995), “Transaction Costs in the Market for Residential Real Estate,” Department of Real Estate Economics Working Paper no.20, Royal Institute of Technology, Sweden, 20, 1-12.

Table 1. Summary of Sales Data on Singapore Condominiums

Number of Sales	Number of Units	Average Selling Price*	Interval between Sales**	Average Price Appreciation+	Std of Price Appreciation++	Average Presale Interval♦	Average Size◊	Number of Rooms	Distance to the nearest MRT stationsℵ	Distance to CBDℵ
1	42,169	861				31.98	129.93	2.77	1.421	9.239
2	8,791	913	47.77	0.52%	0.71%	8.48	137.44	2.74	1.437	8.577
3	1,195	1,030	28.33	0.68%	1.13%	2.51	154.42	2.76	1.500	7.464
4	190	1,087	20.88	0.73%	1.29%	1.53	159.51	2.63	1.383	6.838
5	28	1,418	15.23	0.92%	1.53%	2.06	208.90	2.90	1.427	4.633
6	4	1,129	15.85	0.86%	1.60%	0.00	187.40	2.80	1.362	6.296

* Thousands of Singapore Dollars

** Number of months

\+ Average price appreciation between sales divided by average interval between sales in months.

++ Standard deviation of price appreciation between sales divided by average interval between sales in months.

♦ Average number of months from sales to completion of construction of dwellings.

◊ Average size of dwellings in square meters.

ℵ Average distance in kilometers.

Table 2A. Forecastability of Investment Returns, Singapore Condominiums, 1990-2000 (λ = 0.78 and ρ = 0.55)

$$R_t = \alpha_0 + \sum_{i}^{n} \alpha_i R_{t-i} + \varepsilon_R$$

Constant	0.00183 (0.5892)	0.00162 (0.5151)	0.00168 (0.5426)	0.00111 (0.3579)
R_{t-1}	-0.12501 (1.5106)	-0.09863 (1.0991)	-0.11468 (1.2959)	-0.12307 (1.3603)
R_{t-2}		0.08302 (0.9866)	0.16707 (1.8771)	0.15579 (1.7469)
R_{t-3}			0.10402 (1.2529)	0.11251 (1.2511)
R_{t-4}				0.13853 (1.6644)
σ_R	0.001245	0.001253	0.001209	0.001199
$\bar{R}^2$	0.00984	0.00427	0.0272	0.036564
DW statistics	1.925854	1.978786	1.994373	1.972382
F test	2.299821	1.29469	2.23648	2.267443

Table 2B. Forecastability of Investment Returns, Singapore Condominiums, 1990-2000 (λ = 0.72 and ρ = 0)

$$R_t = \alpha_0 + \sum_i^n \alpha_i R_{t-i} + \varepsilon_R$$

Constant	0.00205 (0.5616)	0.00224 (0.6090)	0.00238 (0.6590)	0.00183 (0.5026)
R_{t-1}	-0.19596 (2.4396)	-0.20156 (2.2651)	-0.19813 (2.2785)	-0.21006 (2.3438)
R_{t-2}		-0.04686 (0.5654)	0.05924 (0.6666)	0.06439 (0.7249)
R_{t-3}			0.12224 (1.5063)	0.14232 (1.5981)
R_{t-4}				0.12797 (1.5651)
σ_R	0.001718	0.001741	0.001658	0.00165
$\bar{R}^2$	0.036965	0.023941	0.041281	0.04117
DW statistics	2.010751	1.953312	2.022047	1.984458
F test	5.998046	2.610564	2.891094	2.429682

Table 2C. Forecastability of Investment Returns,
Singapore Condominiums, 1990-2000
(λ = 1 and ρ = 0)

$$R_t = \alpha_0 + \sum_i^n \alpha_i R_{t-i} + \varepsilon_R$$

Constant	0.00257 (0.5350)	0.00297 (0.6204)	0.00352 (0.7289)	0.00239 (0.5022)
R_{t-1}	-0.33200 (4.1575)	-0.39925 (4.5480)	-0.39656 (4.4443)	-0.38875 (4.4217)
R_{t-2}		-0.16017 (1.8958)	-0.13323 (1.3938)	-0.10233 (1.0854)
R_{t-3}			-0.01123 (0.1241)	0.04448 (0.4633)
R_{t-4}				0.24902 (2.6255)
σ_R	0.002989	0.002947	0.002952	0.002826
$\bar{R}^2$	0.112091	0.128102	0.119128	0.15169
DW statistics	2.132955	1.979492	1.974094	1.983217
F test	17.42017	10.56818	6.887789	6.850136

Figure 1A. Values of Log Likelihood Function at various values of λ assuming ρ=0.

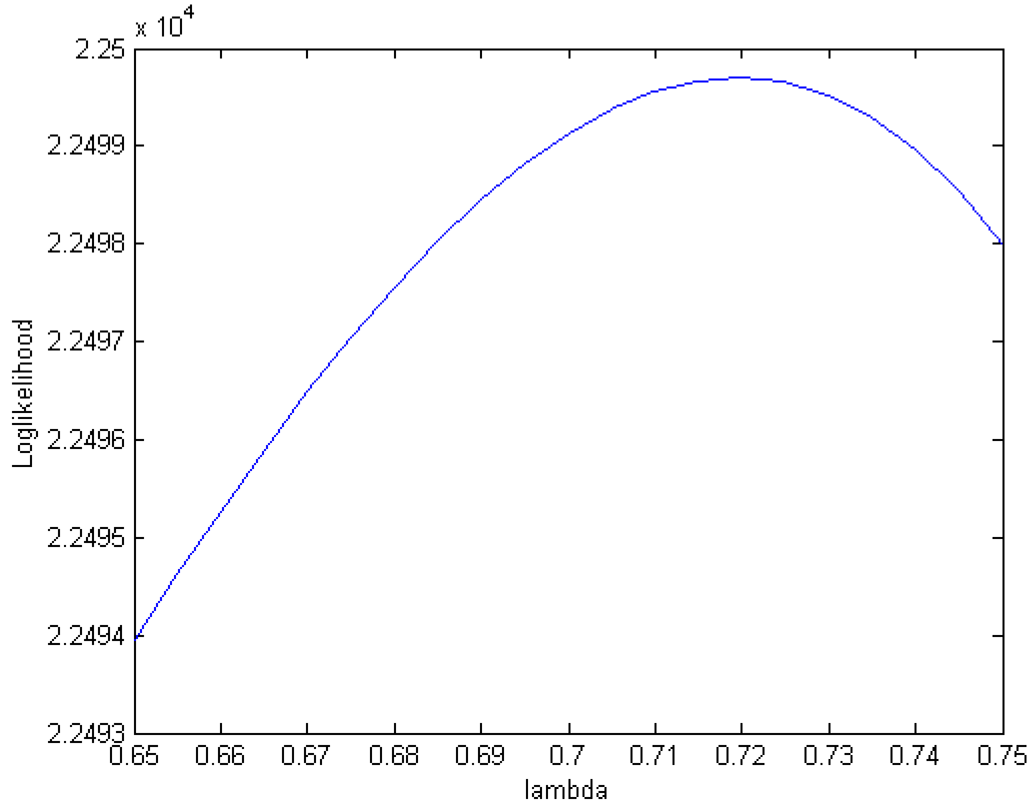


Figure 1B. Loglikelihood Surface at various values of λ and ρ

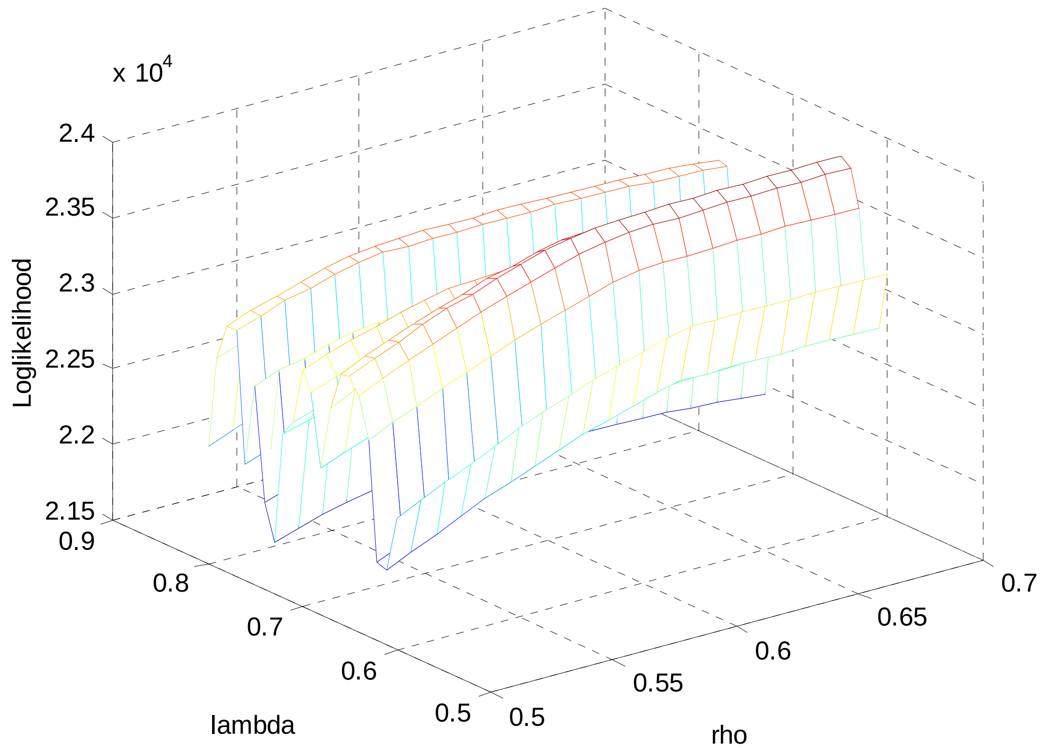


Figure 2. Illustration of a unit shock in house prices over a neighborhood

(for $\lambda = 0.78$; $\rho = 0.55$)

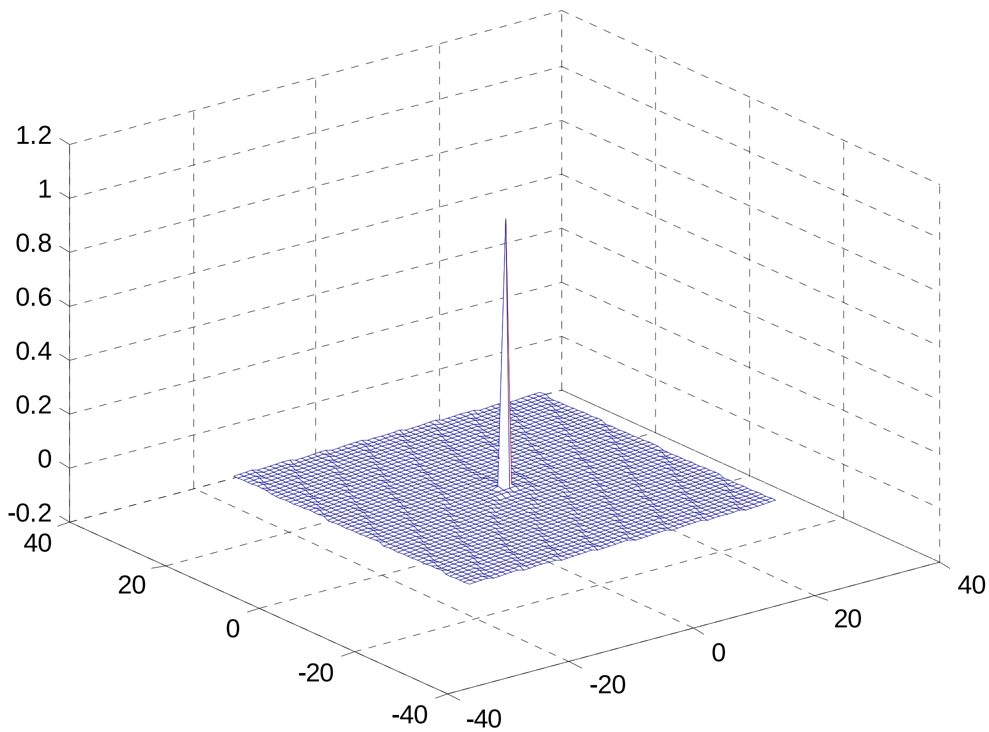


Figure 3A. Price Index Estimate for Singapore Condominiums
($\lambda = 0.78$ and $\rho = 0.55$)

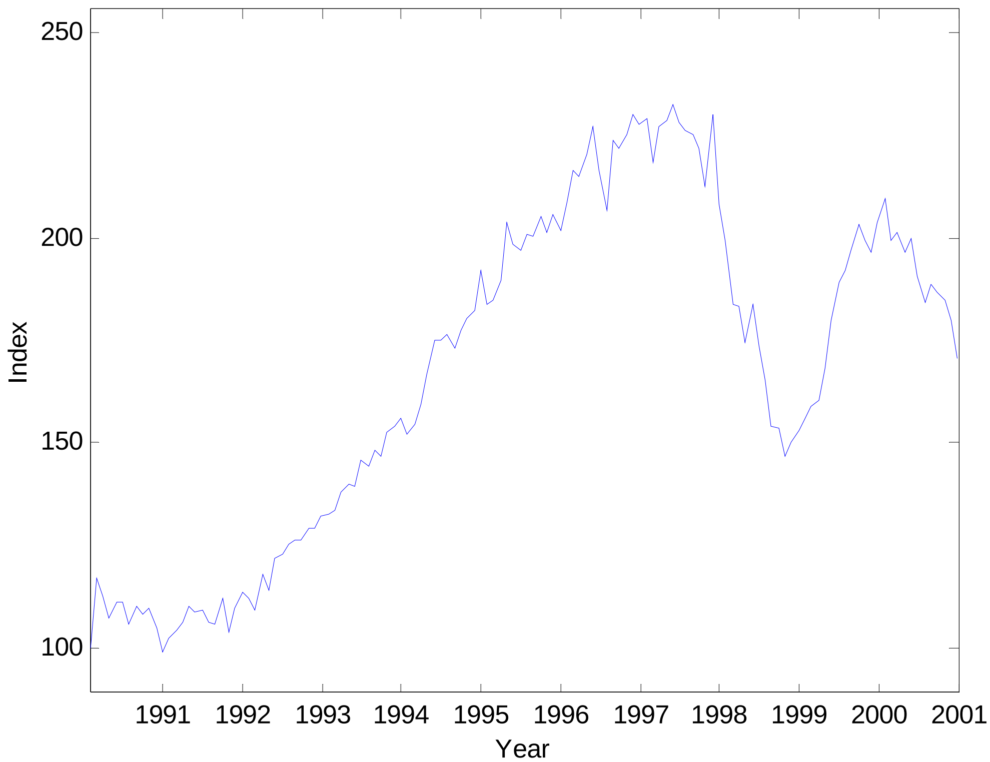


Figure 3B. Price Index Estimate for Singapore Condominiums ($\lambda = 0.72$ and $\rho = 0$)

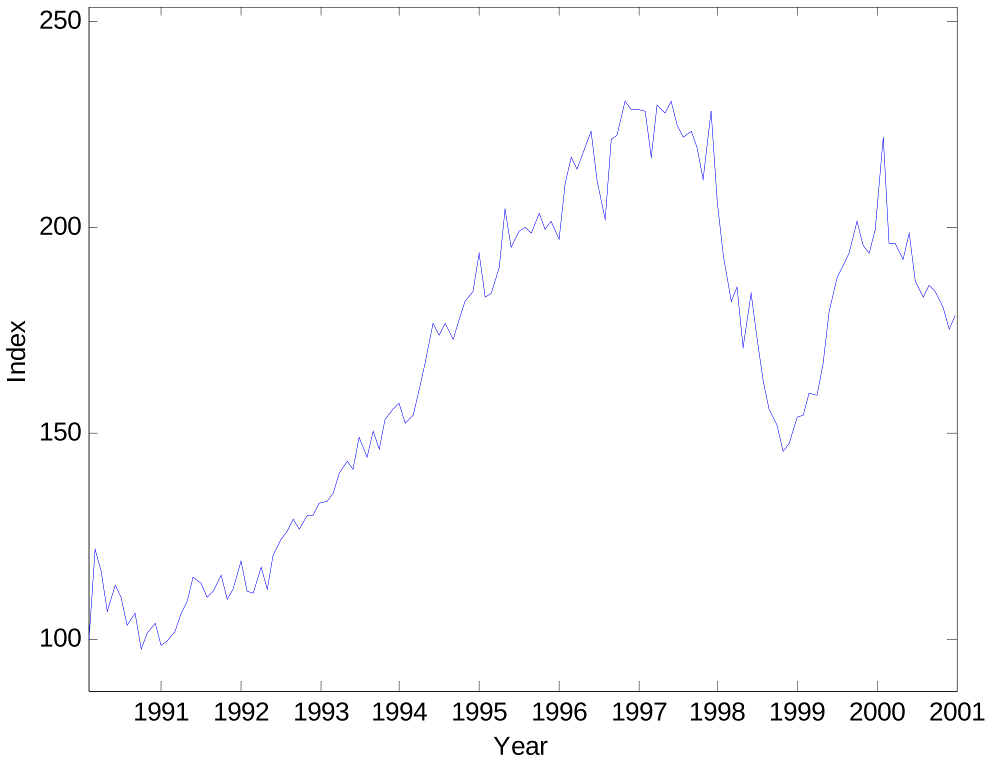


Figure 3C. Price Index Estimate for Singapore Condominiums ($\lambda = 1$ and $\rho = 0$)

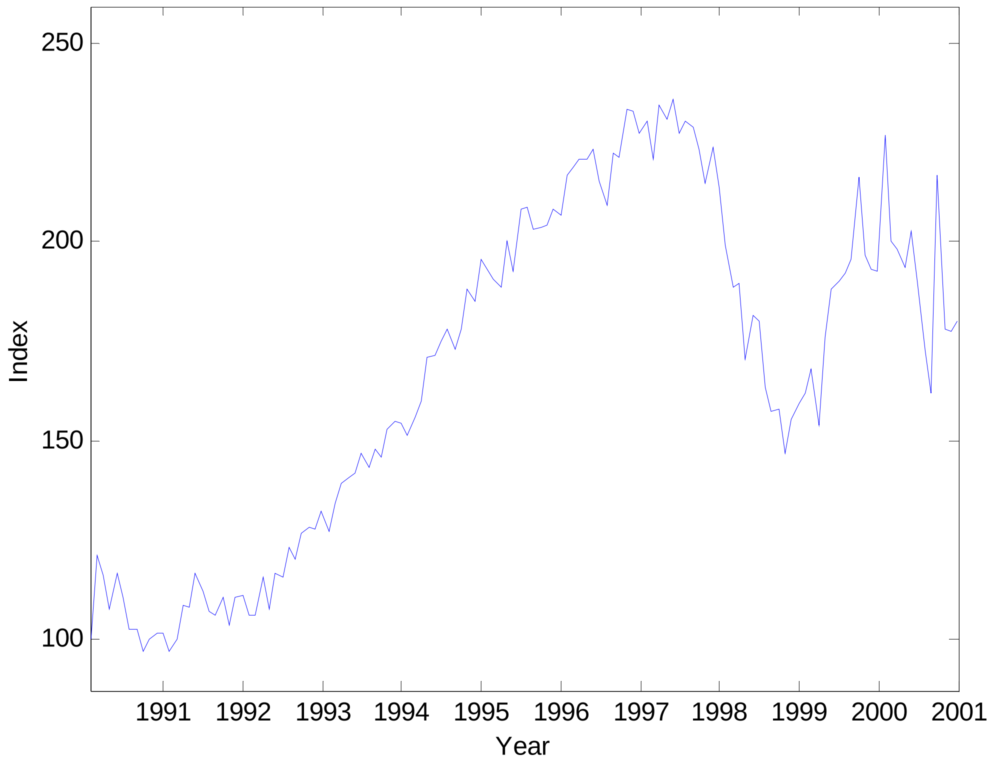


Figure 4A. Estimated Monthly Investment Returns of Condominium Housing in Singapore, 1990 – 2000.

(λ = 0.78 and ρ = 0.55)

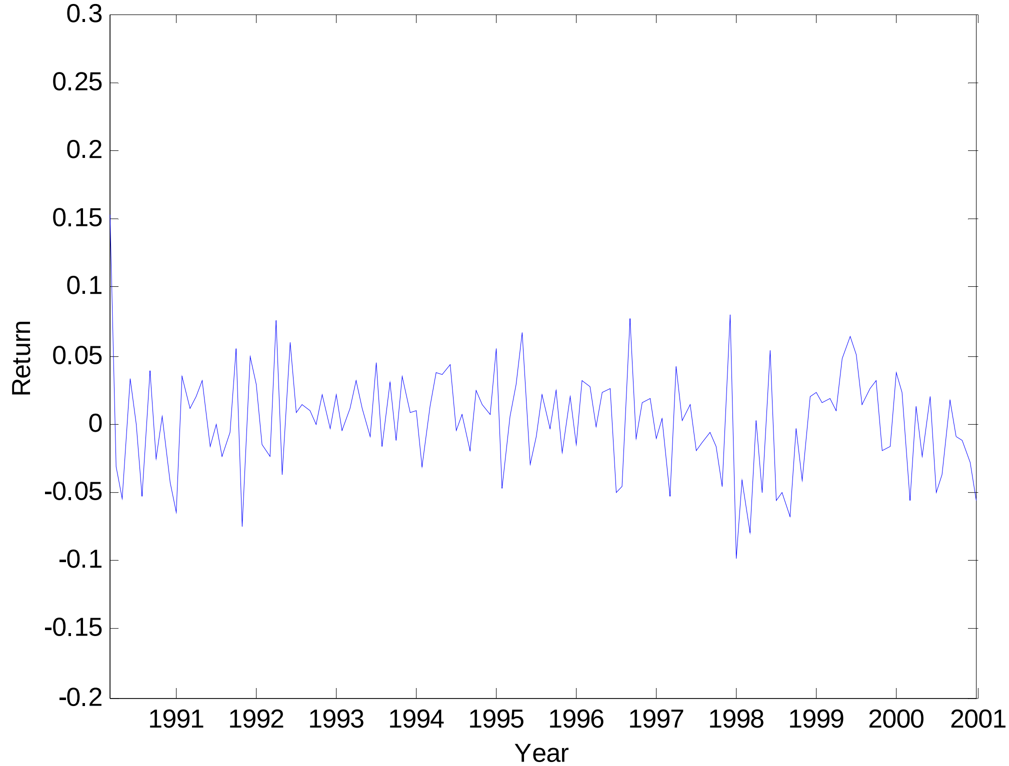


Figure 4B. Estimated Monthly Investment Returns of Condominium Housing in Singapore, 1990 – 2000.

($\lambda = 0.72$ and $\rho = 0$)

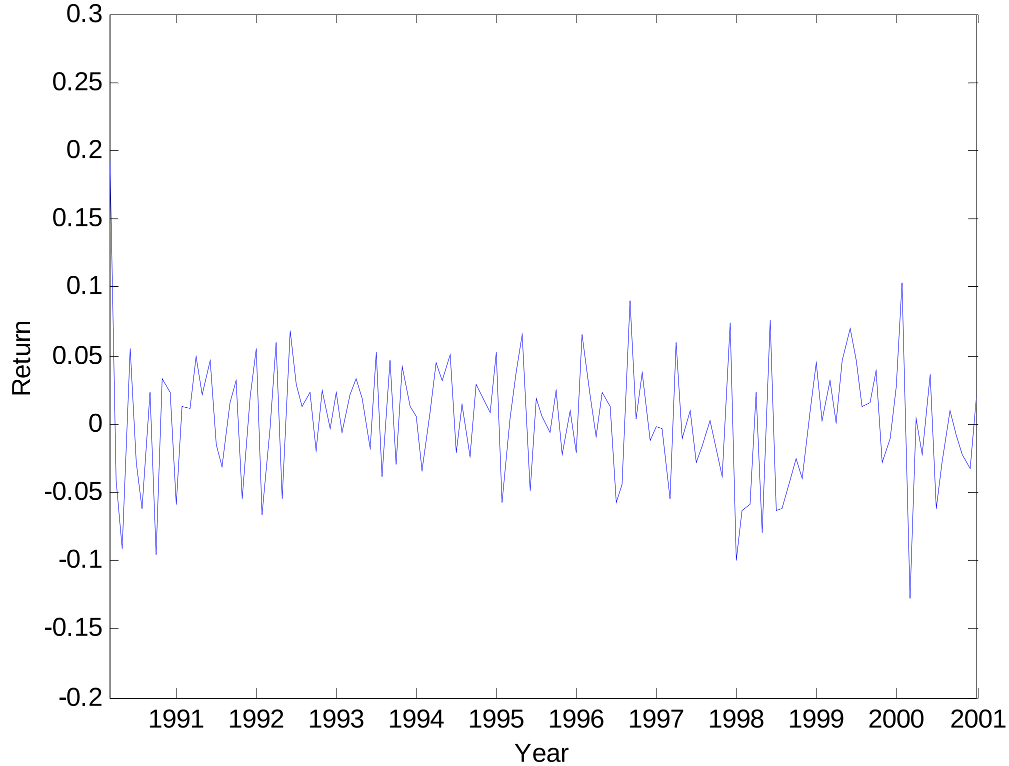


Figure 4C. Estimated Monthly Investment Returns of Condominium Housing in Singapore, 1990 – 2000.

(λ = 1 and ρ = 0)

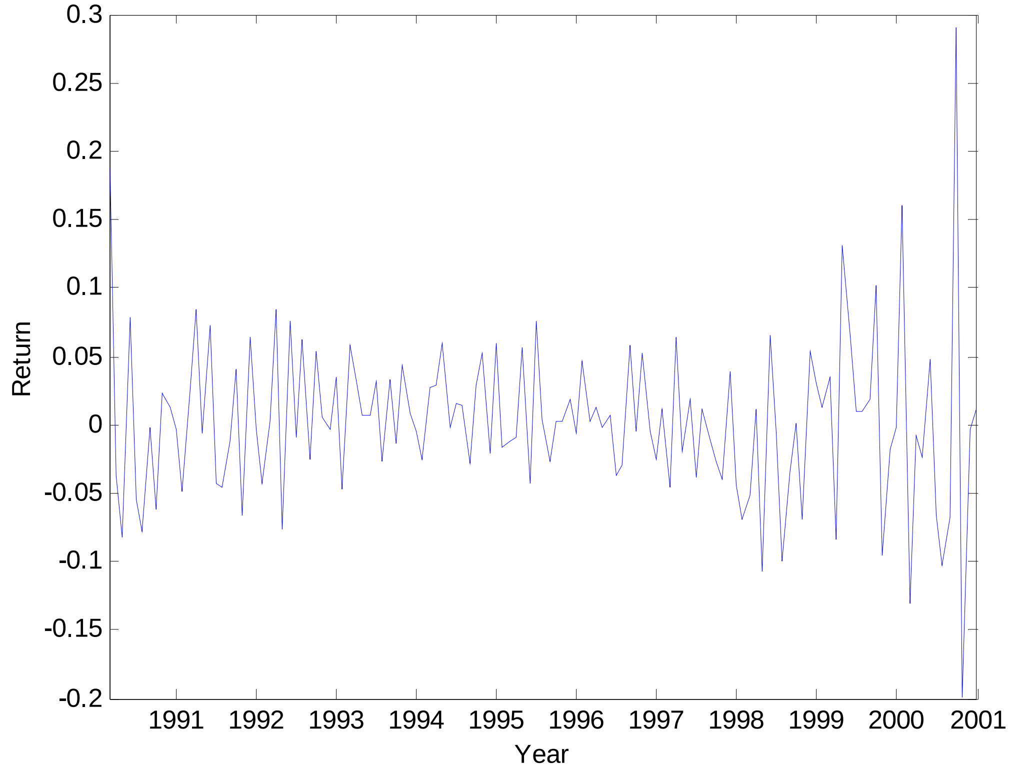
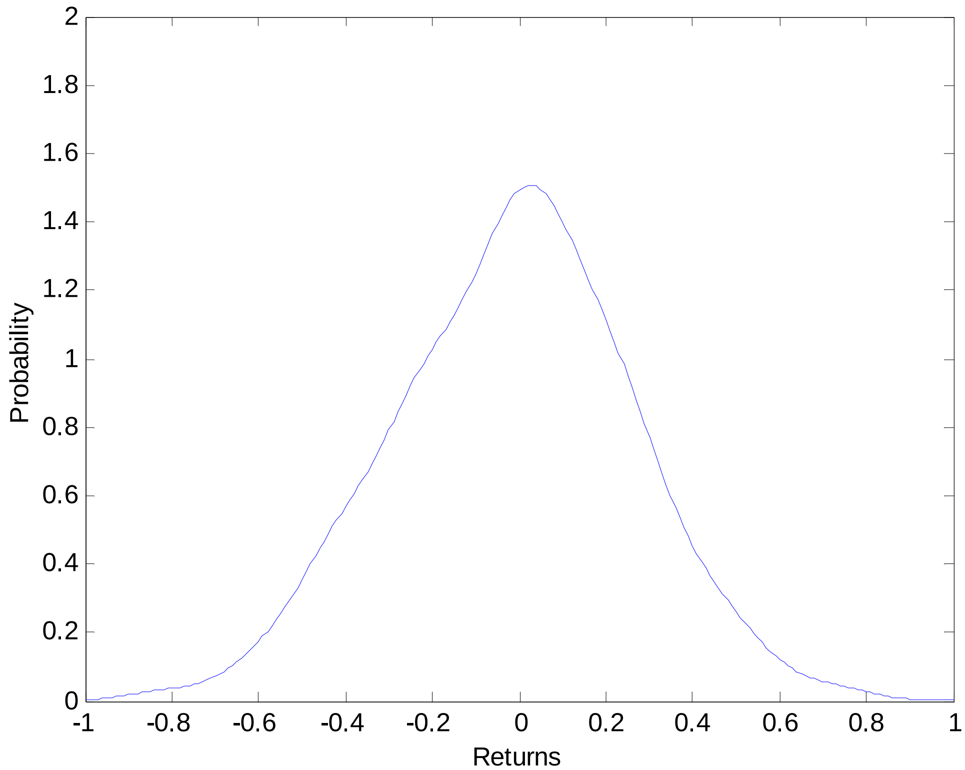
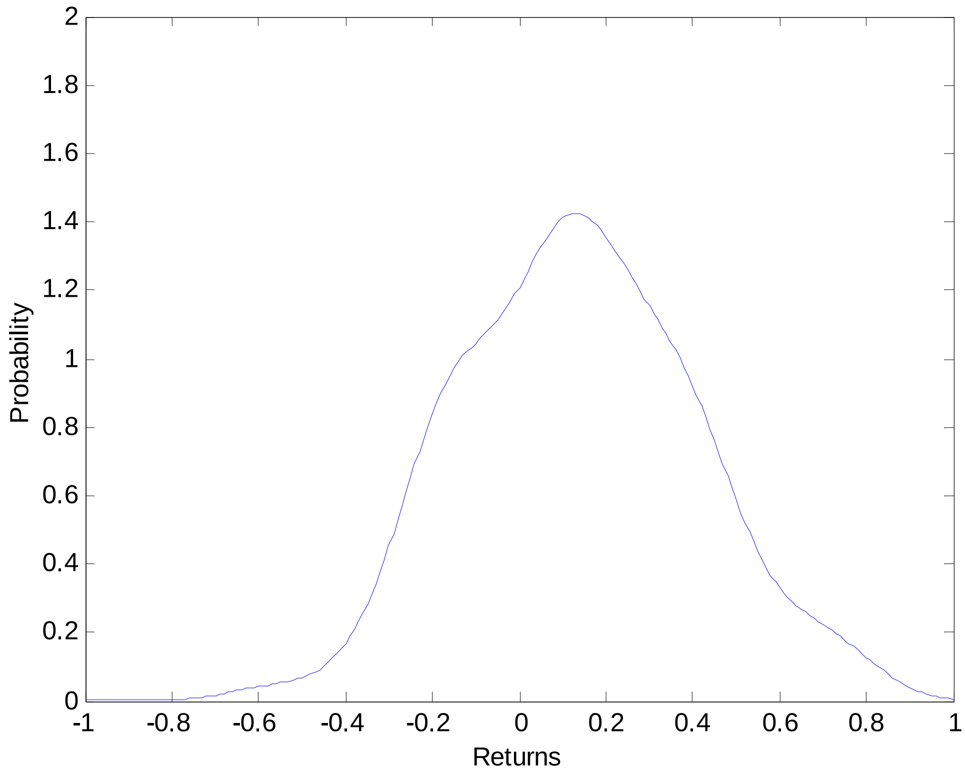


Figure 5A.
Distribution of Returns for Investor who knows λ = 0.78 and ρ = 0.55.
Investment Rule : Always Invest



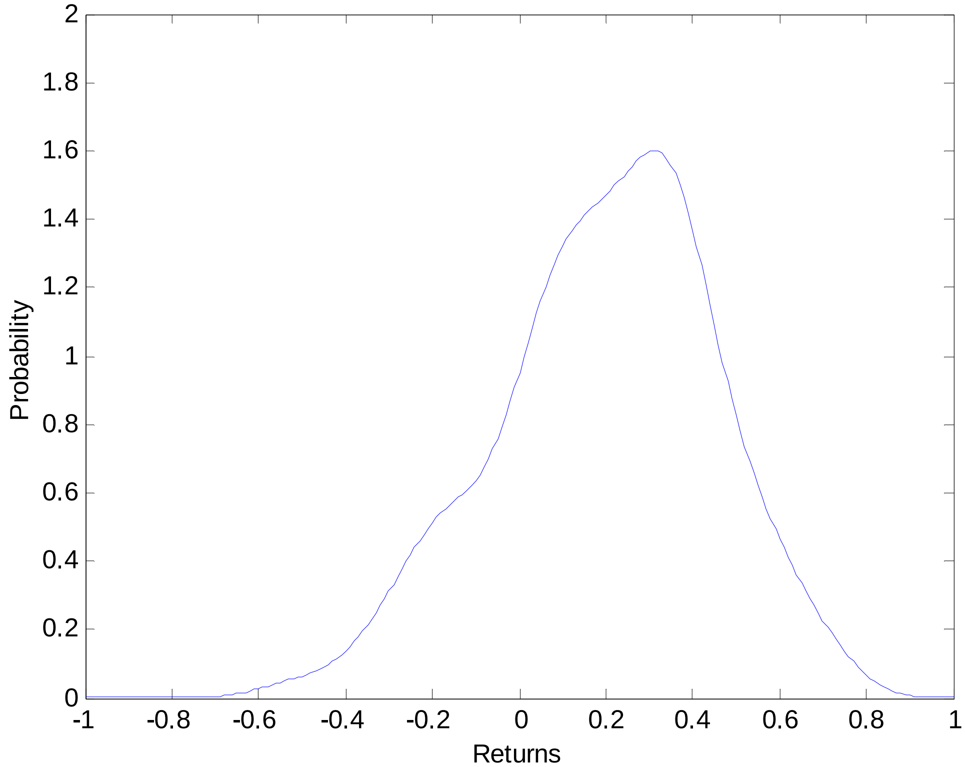
Mean Return : -1.28 %
Standard Deviation: 27.24 %

Figure 5B.
Distribution of Returns for Investor who knows λ = 0.78 and ρ = 0.55.
Investment Rule : 0%



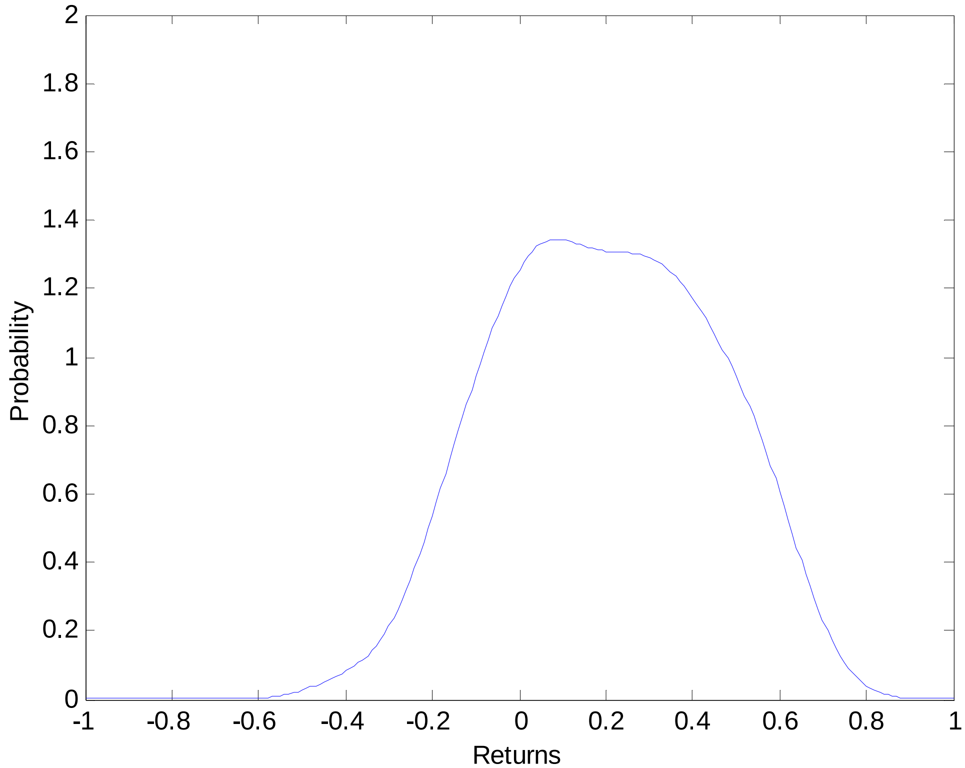
Mean Return : 13.69 %
Standard Deviation: 26.56 %

Figure 5C.
Distribution of Returns for Investor who knows λ = 0.78 and ρ = 0.55.
Investment Rule : 5%



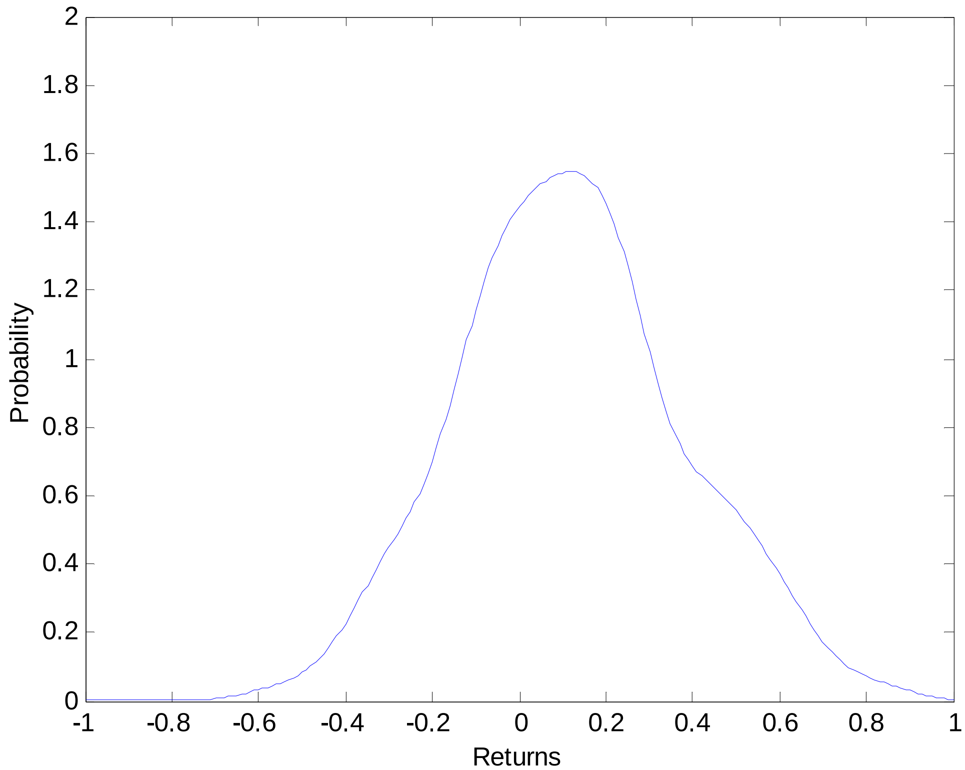
Mean Return : 20.22 %
Standard Deviation: 24.77 %

Figure 5D.
Distribution of Returns for Investor who knows λ = 0.78 and ρ = 0.55.
Investment Rule : 10%



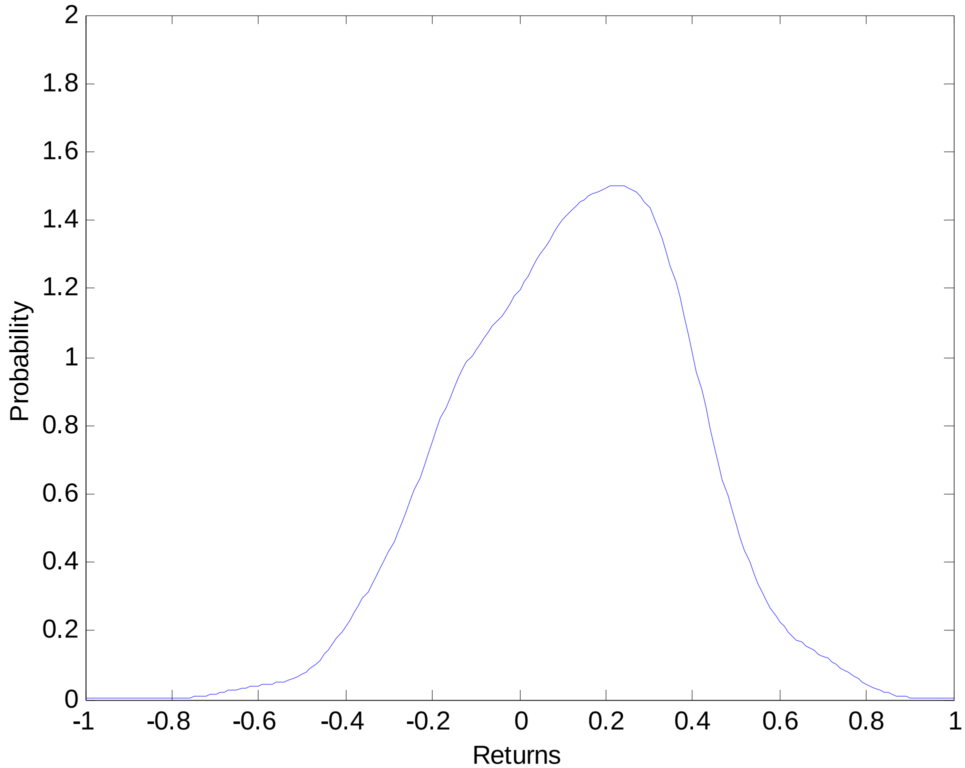
Mean Return : 19.71 %
Standard Deviation: 23.94 %

Figure 6A.
Distribution of Returns for Investor who estimates λ = 0.72 and ρ = 0.
Investment Rule : 0%



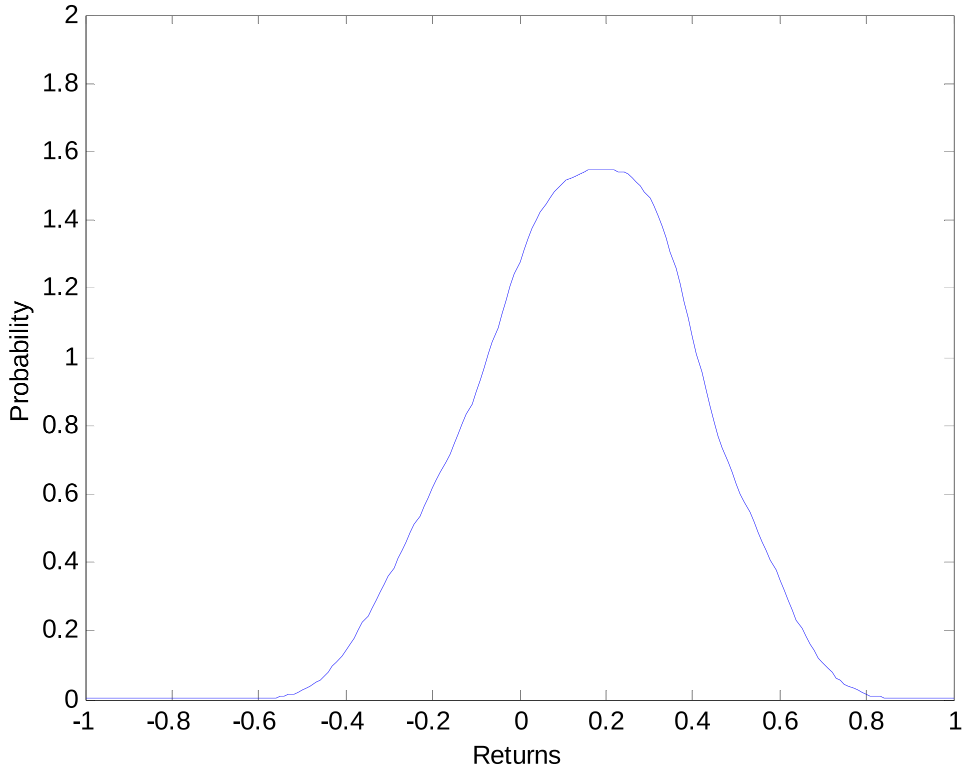
Mean Return : 11.92 %
Standard Deviation: 25.81 %

Figure 6B.
Distribution of Returns for Investor who estimates λ = 0.72 and ρ = 0.
Investment Rule : 5%



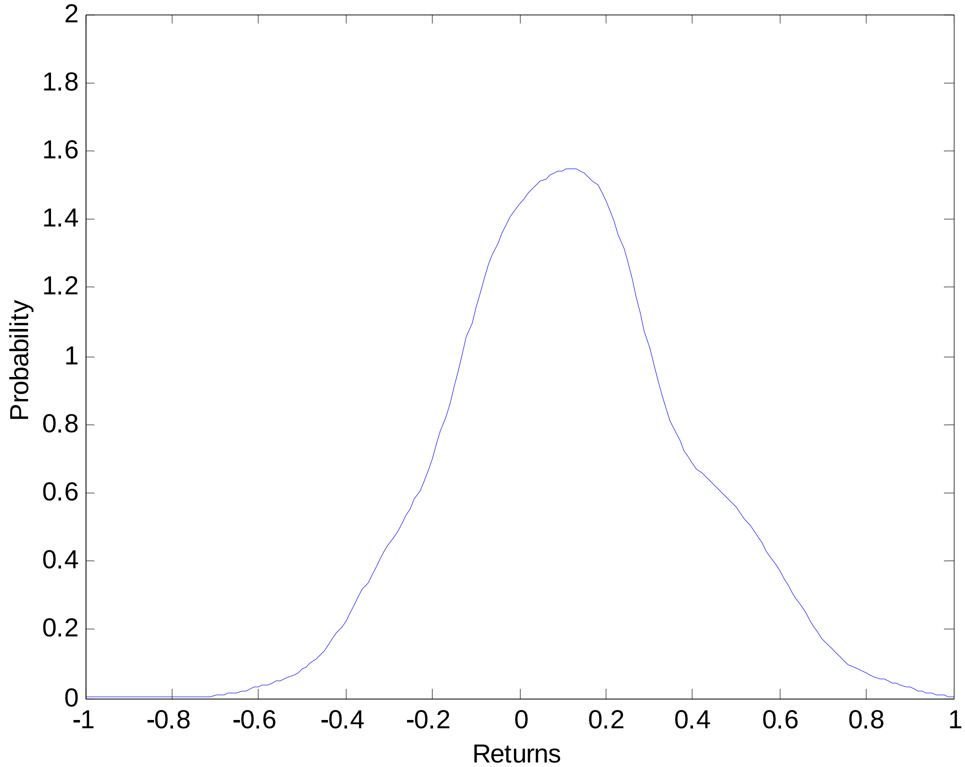
Mean Return : 12.53 %
Standard Deviation: 24.55 %

Figure 6C.
Distribution of Returns for Investor who estimates λ = 0.72 and ρ = 0.
Investment Rule : 10%



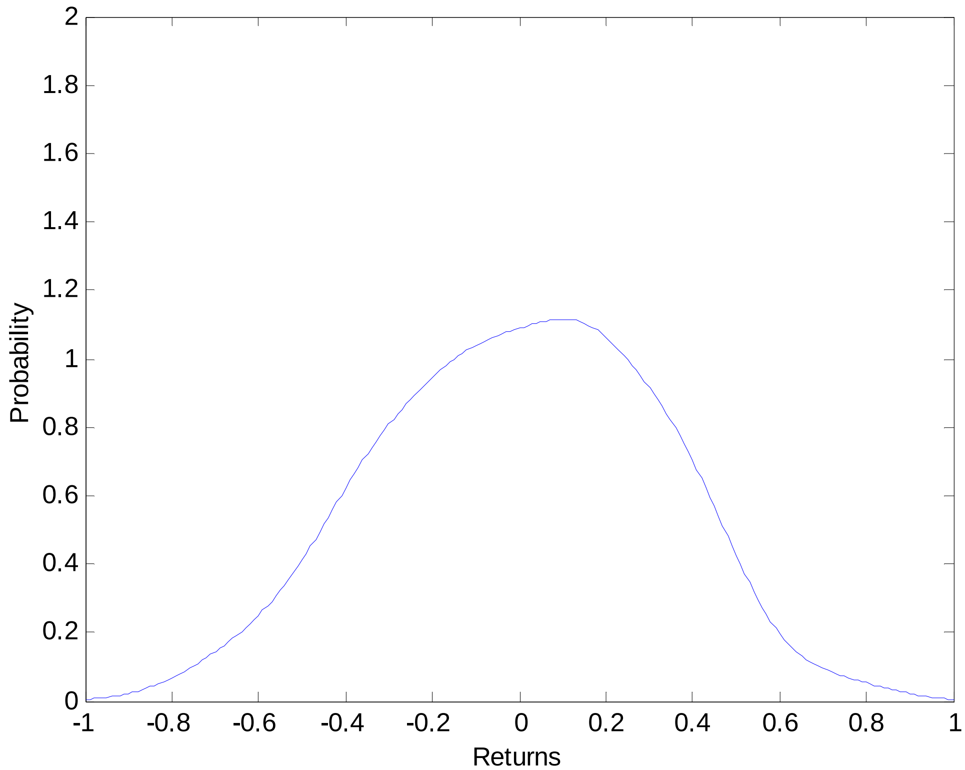
Mean Return : 15.41 %
Standard Deviation: 22.85 %

Figure 7A.
Distribution of Returns for Investor who assumes λ = 1 and ρ = 0.
Investment Rule : 0%



Mean Return : **-1.28 %**
Standard Deviation: **27.24 %**

Figure 7B.
Distribution of Returns for Investor who assumes $\lambda = 1$ and $\rho = 0$.
Investment Rule : 5%



Mean Return : 0.25 %
Standard Deviation: 31.09 %

Appendix

Difficulties of ML Estimation when $\rho \neq 0$: An Illustration

This appendix illustrates the difficulties encountered in maximizing the likelihood function in equation (14) in the presence of spatial as well as temporal autocorrelation. We consider a simple example. Suppose there are three houses in the sample; house A has been sold twice at $t = t_1, t_2$, B three times at $t = t_1, t_2, t_3$ and C four times at $t = t_1, t_2, t_3, t_4$. Let $\lambda = 0.5$, $\rho = 0.5$, $\sigma_\mu^2 = 0.06$, $\sigma_\eta^2 = 0.1$, $t_1 = 0$, $t_2 = 1$, $t_3 = 3$ and $t_4 = 5$. Also let the distance between house A and house B be 1, between house B and house C be 2 and between house A and house C be 3.

To compute $E[(e_{it} - e_{is})(e_{k\tau} - e_{k\varsigma})]$, the elements of the matrix **Σ**, we need $E[(\xi_{it} - \xi_{is})(\xi_{i\tau} - \xi_{i\varsigma})]$ and $\boldsymbol{\pi}_i'\boldsymbol{\pi}_k$.
First, the values of $E[(\xi_{it} - \xi_{is})(\xi_{i\tau} - \xi_{i\varsigma})]$ using above numbers, are, from Equation (10):

$$E\left[\left(\xi_2^A - \xi_1^A\right)^2\right] = 2\left(1 - \lambda^{t_2 - t_1}\right)\left(\frac{\sigma_\mu^2}{1-\lambda^2}\right) + 2\sigma_\eta^2 = 0.28\,.$$

$$E\left[\left(\xi_2^B - \xi_1^B\right)^2\right] = 2\left(1 - \lambda^{t_2 - t_1}\right)\left(\frac{\sigma_\mu^2}{1-\lambda^2}\right) + 2\sigma_\eta^2 = 0.28$$

$$E\left[\left(\xi_3^B - \xi_2^B\right)^2\right] = 2\left(1 - \lambda^{t_3 - t_2}\right)\left(\frac{\sigma_\mu^2}{1-\lambda^2}\right) + 2\sigma_\eta^2 = 0.32$$

$$E\left[\left(\xi_2^C - \xi_1^C\right)^2\right] = 2\left(1 - \lambda^{t_2 - t_1}\right)\left(\frac{\sigma_\mu^2}{1-\lambda^2}\right) + 2\sigma_\eta^2 = 0.28$$

$$E\left[\left(\xi_3^C - \xi_2^C\right)^2\right] = 2\left(1 - \lambda^{t_3 - t_2}\right)\left(\frac{\sigma_\mu^2}{1-\lambda^2}\right) + 2\sigma_\eta^2 = 0.32$$

$$E\left[\left(\xi_4^C - \xi_3^C\right)^2\right] = 2\left(1 - \lambda^{t_4 - t_3}\right)\left(\frac{\sigma_\mu^2}{1-\lambda^2}\right) + 2\sigma_\eta^2 = 0.32$$

$$E\left[\left(\xi_2^B - \xi_1^B\right)\left(\xi_3^B - \xi_2^B\right)\right] = \left(\lambda^{t_3 - t_2} - 1 - \lambda^{t_3 - t_1} + \lambda^{t_2 - t_1}\right)\left(\frac{\sigma_\mu^2}{1-\lambda^2}\right) - \sigma_\eta^2 = -0.13$$

$$E\left[\left(\xi_2^C - \xi_1^C\right)\left(\xi_3^C - \xi_2^C\right)\right] = \left(\lambda^{t_3 - t_2} - 1 - \lambda^{t_3 - t_1} + \lambda^{t_2 - t_1}\right)\left(\frac{\sigma_\mu^2}{1-\lambda^2}\right) - \sigma_\eta^2 = -0.13$$

$$E\left[\left(\xi_3^C-\xi_2^C\right)\left(\xi_4^C-\xi_3^C\right)\right]=\left(\lambda^{t_4-t_3}-1-\lambda^{t_4-t_2}+\lambda^{t_3-t_2}\right)\left(\frac{\sigma_\mu^2}{1-\lambda^2}\right)-\sigma_\eta^2=-0.145$$

$$E\left[\left(\xi_2^C-\xi_1^C\right)\left(\xi_4^C-\xi_3^C\right)\right]=\left(\lambda^{t_4-t_2}-\lambda^{t_3-t_2}-\lambda^{t_4-t_1}+\lambda^{t_3-t_1}\right)\left(\frac{\sigma_\mu^2}{1-\lambda^2}\right)=-0.0075$$

Let the weight on a neighboring house be the inverses of distance, i.e.,

$$W=\begin{bmatrix}0 & 1 & 1/3\\ 1 & 0 & 1/2\\ 1/3 & 1/2 & 0\end{bmatrix}.\ \text{Thus, } \left(I-\rho W\right)^{-1}=\begin{bmatrix}1.5169 & 0.8764 & 0.4719\\ 0.8764 & 1.5736 & 0.5393\\ 0.4719 & 0.5393 & 0.2135\end{bmatrix}.$$

For $\boldsymbol{\pi}_i'\boldsymbol{\pi}_j$'s, we have

$\boldsymbol{\pi}_A'\boldsymbol{\pi}_A=3.2916$,
$\boldsymbol{\pi}_A'\boldsymbol{\pi}_B=2.9625$,
$\boldsymbol{\pi}_A'\boldsymbol{\pi}_C=1.7611$,
$\boldsymbol{\pi}_B'\boldsymbol{\pi}_B=3.5334$,
$\boldsymbol{\pi}_B'\boldsymbol{\pi}_C=1.9164$ and
$\boldsymbol{\pi}_C'\boldsymbol{\pi}_C=1.9861$.

Now when there is no spatial dependency, $\rho=0$,

$\boldsymbol{\Pi}=\left(\mathbf{I}-\rho\mathbf{W}\right)^{-1}=\mathbf{I}$, which implies

$\boldsymbol{\pi}_i'\boldsymbol{\pi}_k=1$, when $i=k$ and 0, otherwise.

Then,

$$E\left[\left(e_{it}-e_{is}\right)\left(e_{k\tau}-e_{k\varsigma}\right)\right]=\boldsymbol{\pi}_i'\left\{E\left[\left(\xi_{it}-\xi_{is}\right)\left(\xi_{i\tau}-\xi_{i\varsigma}\right)\right]\times\mathbf{I}\right\}\boldsymbol{\pi}_k$$

$$=E\left[\left(\xi_{it}-\xi_{is}\right)\left(\xi_{i\tau}-\xi_{i\varsigma}\right)\right]\text{, when } i=k \text{ and 0, otherwise.}$$

When ρ=0, innovations among houses are uncorrelated.

The block diagonal elements of $\boldsymbol{\Sigma}$ are

$$\mathbf{V}_A=E\left[\left(\xi_2^A-\xi_1^A\right)^2\right],$$

$$\mathbf{V}_B=\begin{bmatrix}E\left[\left(\xi_2^B-\xi_1^B\right)^2\right] & E\left[\left(\xi_2^B-\xi_1^B\right)\left(\xi_3^B-\xi_2^B\right)\right]\\ E\left[\left(\xi_2^B-\xi_1^B\right)\left(\xi_3^B-\xi_2^B\right)\right] & E\left[\left(\xi_3^B-\xi_2^B\right)^2\right]\end{bmatrix}\text{, and}$$

$$\mathbf{V}_C = \begin{bmatrix} E\left[\left(\xi_2^C - \xi_1^C\right)^2\right] & E\left[\left(\xi_2^C - \xi_1^C\right)\left(\xi_3^C - \xi_2^C\right)\right] & E\left[\left(\xi_2^C - \xi_1^C\right)\left(\xi_4^C - \xi_3^C\right)\right] \\ E\left[\left(\xi_2^C - \xi_1^C\right)\left(\xi_3^C - \xi_2^C\right)\right] & E\left[\left(\xi_3^C - \xi_2^C\right)^2\right] & E\left[\left(\xi_3^C - \xi_2^C\right)\left(\xi_4^C - \xi_3^C\right)\right] \\ E\left[\left(\xi_2^C - \xi_1^C\right)\left(\xi_4^C - \xi_3^C\right)\right] & E\left[\left(\xi_3^C - \xi_2^C\right)\left(\xi_4^C - \xi_3^C\right)\right] & E\left[\left(\xi_4^C - \xi_3^C\right)^2\right] \end{bmatrix},$$

then,

$$\mathbf{\Sigma} = \left[E\left[\left(e_{it} - e_{is}\right)\left(e_{k\tau} - e_{k\varsigma}\right)\right]\right] = \begin{bmatrix} \mathbf{V}_A & 0 & 0 \\ 0 & \mathbf{V}_B & 0 \\ 0 & 0 & \mathbf{V}_C \end{bmatrix}.$$

$$(1)\ \ \mathbf{\Sigma} = \begin{bmatrix} 0.28 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.28 & -0.13 & 0 & 0 & 0 \\ 0 & -0.13 & 0.32 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.28 & -0.13 & -0.0075 \\ 0 & 0 & 0 & -0.13 & 0.32 & -0.145 \\ 0 & 0 & 0 & -0.0075 & -0.145 & 0.32 \end{bmatrix}.$$

When there is spatial dependency, i.e., $\rho \neq 0$, then innovations in houses are all correlated, and $\mathbf{\Sigma}$ is no longer block diagonal. Indeed, we have,

$$\mathbf{\Sigma} = \begin{bmatrix} 0.28\boldsymbol{\pi}'_A\boldsymbol{\pi}_A & 0.28\boldsymbol{\pi}'_A\boldsymbol{\pi}_B & -0.13\boldsymbol{\pi}'_A\boldsymbol{\pi}_B & 0.28\boldsymbol{\pi}'_A\boldsymbol{\pi}_C & -0.13\boldsymbol{\pi}'_A\boldsymbol{\pi}_C & -0.0075\boldsymbol{\pi}'_A\boldsymbol{\pi}_C \\ 0.28\boldsymbol{\pi}'_B\boldsymbol{\pi}_A & 0.28\boldsymbol{\pi}'_B\boldsymbol{\pi}_B & -0.13\boldsymbol{\pi}'_B\boldsymbol{\pi}_B & 0.28\boldsymbol{\pi}'_B\boldsymbol{\pi}_C & -0.13\boldsymbol{\pi}'_B\boldsymbol{\pi}_C & -0.0075\boldsymbol{\pi}'_B\boldsymbol{\pi}_C \\ -0.13\boldsymbol{\pi}'_B\boldsymbol{\pi}_A & -0.13\boldsymbol{\pi}'_B\boldsymbol{\pi}_B & 0.32\boldsymbol{\pi}'_B\boldsymbol{\pi}_B & -0.13\boldsymbol{\pi}'_B\boldsymbol{\pi}_C & 0.32\boldsymbol{\pi}'_B\boldsymbol{\pi}_C & -0.145\boldsymbol{\pi}'_B\boldsymbol{\pi}_C \\ 0.28\boldsymbol{\pi}'_C\boldsymbol{\pi}_A & 0.28\boldsymbol{\pi}'_C\boldsymbol{\pi}_B & -0.13\boldsymbol{\pi}'_C\boldsymbol{\pi}_B & 0.28\boldsymbol{\pi}'_C\boldsymbol{\pi}_C & -0.13\boldsymbol{\pi}'_C\boldsymbol{\pi}_C & -0.0075\boldsymbol{\pi}'_C\boldsymbol{\pi}_C \\ -0.13\boldsymbol{\pi}'_C\boldsymbol{\pi}_A & -0.13\boldsymbol{\pi}'_C\boldsymbol{\pi}_B & 0.32\boldsymbol{\pi}'_C\boldsymbol{\pi}_B & -0.13\boldsymbol{\pi}'_C\boldsymbol{\pi}_C & 0.32\boldsymbol{\pi}'_C\boldsymbol{\pi}_C & -0.145\boldsymbol{\pi}'_C\boldsymbol{\pi}_C \\ -0.0075\boldsymbol{\pi}'_C\boldsymbol{\pi}_A & -0.0075\boldsymbol{\pi}'_C\boldsymbol{\pi}_B & -0.145\boldsymbol{\pi}'_C\boldsymbol{\pi}_B & -0.0075\boldsymbol{\pi}'_C\boldsymbol{\pi}_C & -0.145\boldsymbol{\pi}'_C\boldsymbol{\pi}_C & 0.32\boldsymbol{\pi}'_C\boldsymbol{\pi}_C \end{bmatrix}$$

$$(2)\ \mathbf{\Sigma} = \begin{bmatrix} 0.92 & 0.83 & -0.39 & 0.49 & -0.23 & -0.01 \\ 0.83 & 0.99 & -0.46 & 0.54 & -0.25 & -0.01 \\ -0.39 & -0.46 & 1.13 & -0.25 & 0.61 & -0.28 \\ 0.49 & 0.54 & -0.25 & 0.56 & -0.26 & -0.01 \\ -0.23 & -0.25 & 0.61 & -0.26 & 0.64 & -0.29 \\ -0.01 & -0.01 & -0.28 & -0.01 & -0.29 & 0.64 \end{bmatrix}$$

The inverse of the sparse block diagonal matrix, illustrated in (1), can be computed rather easily, even when $\mathbf{\Sigma}$ is large. The inverse of the general matrix, illustrated in (2), can be computationally burdensome.

Appendix Table A1.
Generalized Least Squares Estimates of Price Index Equation (13)

	Spatially Correlated and Mean Reversion	Mean Reversion	Random Walk
Constant	0.0723	0.0885	0.0783
	(20.2685)	(24.0443)	(29.8365)
Feb,1990	0.1558	0.1951	0.1903
	(1.0329)	(1.4461)	(0.7027)
Mar,1990	0.1209	0.1502	0.1501
	(0.8252)	(1.1889)	(0.5804)
Apr,1990	0.0713	0.0647	0.0732
	(0.5057)	(0.5397)	(0.2912)
May,1990	0.1063	0.1219	0.1534
	(0.7670)	(1.0476)	(0.6255)
Jun,1990	0.1078	0.0952	0.1007
	(0.7579)	(0.7824)	(0.3975)
Jul,1990	0.0549	0.0333	0.0221
	(0.3770)	(0.2651)	(0.0841)
Aug,1990	0.0971	0.0592	0.0235
	(0.6488)	(0.4568)	(0.0882)
Sep,1990	0.0787	-0.0286	-0.0308
	(0.5245)	(0.2327)	(0.1197)
Oct,1990	0.0922	0.0127	0.0003
	(0.6301)	(0.1034)	(0.0012)
Nov,1990	0.0498	0.0355	0.0127
	(0.3309)	(0.2742)	(0.0477)
Dec,1990	-0.0107	-0.0185	0.0129
	(0.0745)	(0.1546)	(0.0513)
Jan,1991	0.0256	-0.0048	-0.0341
	(0.1662)	(0.0360)	(0.1257)
Feb,1991	0.0451	0.0154	-0.0012
	(0.3126)	(0.1258)	(0.0048)
Mar,1991	0.0606	0.0605	0.0793
	(0.4353)	(0.5163)	(0.3228)
Apr,1991	0.0979	0.0879	0.0779

	(0.7135)	(0.7690)	(0.3229)
May,1991	0.0851	0.1380	0.1540
	(0.6160)	(1.1978)	(0.6336)
Jun,1991	0.0866	0.1258	0.1131
	(0.6260)	(1.0934)	(0.4689)
Jul,1991	0.0635	0.0956	0.0690
	(0.4563)	(0.8233)	(0.2856)
Aug,1991	0.0576	0.1110	0.0562
	(0.4118)	(0.9482)	(0.2301)
Sep,1991	0.1152	0.1448	0.0982
	(0.8251)	(1.2364)	(0.4003)
Oct,1991	0.0402	0.0895	0.0313
	(0.2874)	(0.7597)	(0.1276)
Nov,1991	0.0922	0.1114	0.0974
	(0.6649)	(0.9552)	(0.3994)
Dec,1991	0.1292	0.1741	0.1016
	(0.9205)	(1.4721)	(0.4149)
Jan,1992	0.1140	0.1077	0.0585
	(0.8000)	(0.8965)	(0.2364)
Feb,1992	0.0888	0.1028	0.0592
	(0.6337)	(0.8692)	(0.2394)
Mar,1992	0.1637	0.1610	0.1419
	(1.1530)	(1.3559)	(0.5726)
Apr,1992	0.1323	0.1128	0.0714
	(0.9497)	(0.9653)	(0.2921)
May,1992	0.1975	0.1861	0.1522
	(1.4453)	(1.6271)	(0.6310)
Jun,1992	0.2065	0.2154	0.1430
	(1.5159)	(1.8875)	(0.5954)
Jul,1992	0.2251	0.2323	0.2091
	(1.6548)	(2.0379)	(0.8695)
Aug,1992	0.2325	0.2529	0.1822
	(1.7074)	(2.2157)	(0.7585)
Sep,1992	0.2328	0.2342	0.2371
	(1.7160)	(2.0608)	(0.9908)
Oct,1992	0.2574	0.2616	0.2457
	(1.9007)	(2.3065)	(1.0279)

Nov,1992	0.2562	0.2602	0.2438
	(1.8904)	(2.2918)	(1.0197)
Dec,1992	0.2791	0.2842	0.2801
	(2.0575)	(2.4990)	(1.1714)
Jan,1993	0.2819	0.2863	0.2415
	(2.0776)	(2.5166)	(1.0103)
Feb,1993	0.2889	0.3026	0.2952
	(2.1311)	(2.6637)	(1.2357)
Mar,1993	0.3233	0.3378	0.3305
	(2.3950)	(2.9889)	(1.3842)
Apr,1993	0.3376	0.3589	0.3399
	(2.5014)	(3.1782)	(1.4239)
May,1993	0.3318	0.3445	0.3500
	(2.4590)	(3.0512)	(1.4663)
Jun,1993	0.3777	0.3978	0.3825
	(2.7987)	(3.5245)	(1.6025)
Jul,1993	0.3657	0.3636	0.3595
	(2.7068)	(3.2151)	(1.5059)
Aug,1993	0.3945	0.4094	0.3912
	(2.9205)	(3.6196)	(1.6379)
Sep,1993	0.3830	0.3795	0.3775
	(2.8342)	(3.3547)	(1.5810)
Oct,1993	0.4218	0.4256	0.4251
	(3.1221)	(3.7618)	(1.7804)
Nov,1993	0.4328	0.4413	0.4360
	(3.2014)	(3.8999)	(1.8260)
Dec,1993	0.4460	0.4507	0.4346
	(3.2970)	(3.9801)	(1.8200)
Jan,1994	0.4185	0.4209	0.4130
	(3.0941)	(3.7171)	(1.7293)
Feb,1994	0.4345	0.4335	0.4449
	(3.2099)	(3.8250)	(1.8630)
Mar,1994	0.4671	0.4734	0.4690
	(3.4597)	(4.1892)	(1.9649)
Apr,1994	0.5101	0.5121	0.5358
	(3.7785)	(4.5339)	(2.2449)
May,1994	0.5586	0.5687	0.5398

	(4.1432)	(5.0423)	(2.2619)
Jun,1994	0.5590	0.5525	0.5596
	(4.1409)	(4.8913)	(2.3445)
Jul,1994	0.5671	0.5683	0.5756
	(4.2007)	(5.0288)	(2.4114)
Aug,1994	0.5489	0.5454	0.5479
	(4.0663)	(4.8276)	(2.2956)
Sep,1994	0.5732	0.5746	0.5768
	(4.2448)	(5.0857)	(2.4163)
Oct,1994	0.5910	0.5984	0.6327
	(4.3742)	(5.2896)	(2.6499)
Nov,1994	0.6011	0.6107	0.6156
	(4.4524)	(5.4064)	(2.5792)
Dec,1994	0.6528	0.6602	0.6721
	(4.8308)	(5.8353)	(2.8154)
Jan,1995	0.6076	0.6040	0.6571
	(4.4900)	(5.3301)	(2.7521)
Feb,1995	0.6137	0.6096	0.6462
	(4.5219)	(5.3602)	(2.7053)
Mar,1995	0.6394	0.6437	0.6337
	(4.7338)	(5.7066)	(2.6557)
Apr,1995	0.7110	0.7134	0.6945
	(5.2592)	(6.3037)	(2.9086)
May,1995	0.6854	0.6679	0.6552
	(5.0720)	(5.9184)	(2.7451)
Jun,1995	0.6780	0.6880	0.7327
	(5.0134)	(6.0811)	(3.0690)
Jul,1995	0.6978	0.6919	0.7352
	(5.1572)	(6.1082)	(3.0790)
Aug,1995	0.6941	0.6859	0.7085
	(5.1364)	(6.0623)	(2.9673)
Sep,1995	0.7189	0.7107	0.7105
	(5.3085)	(6.2658)	(2.9737)
Oct,1995	0.6993	0.6899	0.7134
	(5.1713)	(6.0904)	(2.9871)
Nov,1995	0.7206	0.7010	0.7343
	(5.3237)	(6.1835)	(3.0742)

Dec,1995	0.7027	0.6778	0.7253
	(5.1972)	(5.9881)	(3.0372)
Jan,1996	0.7360	0.7458	0.7737
	(5.4448)	(6.5930)	(3.2396)
Feb,1996	0.7720	0.7750	0.7839
	(5.7070)	(6.8448)	(3.2828)
Mar,1996	0.7654	0.7606	0.7922
	(5.6618)	(6.7207)	(3.3174)
Apr,1996	0.7904	0.7854	0.7922
	(5.8594)	(6.9563)	(3.3185)
May,1996	0.8211	0.8034	0.8044
	(6.0808)	(7.1100)	(3.3687)
Jun,1996	0.7718	0.7466	0.7678
	(5.6650)	(6.5290)	(3.2046)
Jul,1996	0.7267	0.7025	0.7386
	(5.3541)	(6.1729)	(3.0833)
Aug,1996	0.8052	0.7931	0.7982
	(5.9092)	(6.9342)	(3.3272)
Sep,1996	0.7960	0.7977	0.7950
	(5.8311)	(6.9524)	(3.3125)
Oct,1996	0.8109	0.8351	0.8471
	(5.9378)	(7.2746)	(3.5233)
Nov,1996	0.8323	0.8263	0.8449
	(6.0937)	(7.1921)	(3.5162)
Dec,1996	0.8234	0.8265	0.8216
	(6.0273)	(7.2052)	(3.4217)
Jan,1997	0.8294	0.8244	0.8342
	(6.0645)	(7.1604)	(3.4711)
Feb,1997	0.7814	0.7749	0.7928
	(5.7195)	(6.7468)	(3.2998)
Mar,1997	0.8192	0.8300	0.8522
	(6.0178)	(7.2674)	(3.5523)
Apr,1997	0.8267	0.8233	0.8377
	(6.0753)	(7.2133)	(3.4920)
May,1997	0.8430	0.8347	0.8581
	(6.2181)	(7.3594)	(3.5849)
Jun,1997	0.8249	0.8086	0.8208

	(6.0618)	(7.0740)	(3.4205)
Jul,1997	0.8150	0.7962	0.8350
	(6.0277)	(7.0124)	(3.4871)
Aug,1997	0.8123	0.8022	0.8281
	(5.9580)	(7.0087)	(3.4502)
Sep,1997	0.7973	0.7862	0.8038
	(5.8366)	(6.8454)	(3.3434)
Oct,1997	0.7525	0.7490	0.7653
	(5.5050)	(6.5162)	(3.1824)
Nov,1997	0.8333	0.8238	0.8049
	(6.0622)	(7.1098)	(3.3305)
Dec,1997	0.7338	0.7235	0.7599
	(5.3464)	(6.2632)	(3.1583)
Jan,1998	0.6894	0.6579	0.6879
	(4.9434)	(5.5769)	(2.8396)
Feb,1998	0.6089	0.5977	0.6346
	(4.4021)	(5.1125)	(2.6186)
Mar,1998	0.6070	0.6158	0.6404
	(4.4081)	(5.3063)	(2.6553)
Apr,1998	0.5562	0.5361	0.5332
	(4.0845)	(4.6896)	(2.2219)
May,1998	0.6077	0.6102	0.5963
	(4.4521)	(5.3231)	(2.4830)
Jun,1998	0.5513	0.5462	0.5893
	(4.0406)	(4.7686)	(2.4565)
Jul,1998	0.5016	0.4852	0.4900
	(3.6672)	(4.2231)	(2.0423)
Aug,1998	0.4324	0.4433	0.4544
	(3.1531)	(3.8505)	(1.8870)
Sep,1998	0.4279	0.4173	0.4548
	(3.1124)	(3.6074)	(1.8873)
Oct,1998	0.3843	0.3754	0.3837
	(2.8180)	(3.2943)	(1.5995)
Nov,1998	0.4056	0.3871	0.4386
	(2.9822)	(3.3957)	(1.8309)
Dec,1998	0.4267	0.4304	0.4662
	(3.1459)	(3.7891)	(1.9480)

Jan,1999	0.4441	0.4341	0.4804
	(3.2724)	(3.8169)	(2.0048)
Feb,1999	0.4647	0.4674	0.5177
	(3.4311)	(4.1186)	(2.1627)
Mar,1999	0.4715	0.4651	0.4311
	(3.4818)	(4.1004)	(1.8023)
Apr,1999	0.5208	0.5135	0.5649
	(3.8550)	(4.5412)	(2.3637)
May,1999	0.5865	0.5849	0.6322
	(4.3449)	(5.1773)	(2.6460)
Jun,1999	0.6364	0.6299	0.6414
	(4.7103)	(5.5679)	(2.6832)
Jul,1999	0.6522	0.6442	0.6526
	(4.8237)	(5.6924)	(2.7291)
Aug,1999	0.6785	0.6601	0.6705
	(5.0168)	(5.8284)	(2.8037)
Sep,1999	0.7098	0.6989	0.7720
	(5.2428)	(6.1627)	(3.2269)
Oct,1999	0.6892	0.6702	0.6753
	(5.0844)	(5.8988)	(2.8215)
Nov,1999	0.6741	0.6609	0.6578
	(4.9755)	(5.8231)	(2.7504)
Dec,1999	0.7126	0.6892	0.6564
	(5.2601)	(6.0719)	(2.7431)
Jan,2000	0.7398	0.7964	0.8203
	(5.4596)	(7.0242)	(3.4278)
Feb,2000	0.6888	0.6739	0.6945
	(5.0800)	(5.9308)	(2.8994)
Mar,2000	0.6989	0.6741	0.6833
	(5.1530)	(5.9307)	(2.8506)
Apr,2000	0.6763	0.6530	0.6605
	(4.9850)	(5.7416)	(2.7525)
May,2000	0.6927	0.6864	0.7060
	(5.1093)	(6.0394)	(2.9456)
Jun,2000	0.6436	0.6261	0.6414
	(4.7392)	(5.4978)	(2.6707)
Jul,2000	0.6122	0.6027	0.5442

	(4.5082)	(5.2867)	(2.2664)
Aug,2000	0.6352	0.6191	0.4823
	(4.6749)	(5.4318)	(2.0103)
Sep,2000	0.6253	0.6112	0.7734
	(4.6049)	(5.3657)	(3.2213)
Oct,2000	0.6148	0.5902	0.5754
	(4.5153)	(5.1643)	(2.3908)
Nov,2000	0.5879	0.5592	0.5738
	(4.1855)	(4.6844)	(2.3322)
Dec,2000	0.5345	0.5794	0.5876
	(2.9333)	(3.6423)	(1.8693)
Presale	-0.0012	-0.0011	-0.0010
	(6.7099)	(7.5041)	(4.2092)
λ	0.78	0.72	1
ρ	0.55	-	-
σ_{μ}^{2}	0.0077	0.0133	0.0009
σ_{η}^{2}	0.0054	0.0016	-
Log Likelihood	23743.95	22498.75	7678.98

Note: The numbers in parentheses are t-statistics.