Investment Returns and Price Discovery in the Market for Owner-Occupied Housing*

by

John M. Quigley†
University of California, Berkeley

Christian L. Redfearn‡
University of Southern California

Abstract

This paper examines the dynamics of owner-occupied housing prices both at the level of the individual dwelling and in aggregate. Using a unique data set, a model of individual dwelling prices is estimated that represents features of housing markets more faithfully than competing models. Statistical tests strongly reject the hypothesis that individual housing prices follow a random walk in favor of the alternative hypothesis that housing prices are mean reverting. This result also holds in aggregate, and provides an explanation for the "inertia" reported in housing return series. The paper then demonstrates that real and excess returns are forecastable. Finally, it considers empirically the extent to which the transactions costs associated with home ownership preclude profitable speculation in owner-occupied housing markets.

*Paper prepared for the Fifth annual conference of the Asian Real Estate Society, July 26-30, 2000. A previous revision of this paper was presented at the NBER Workshop on Real Estate Markets, November 1999. We are grateful to Owen Lamont and Tom Rothenberg for commentary and discussion.

†`quigley@econ.berkeley.edu`, phone:1-510-643-7411, fax:1-510-643-7357

‡`redfearn@econ.berkeley.edu`, phone:1-510-643-3507, fax:1-510-643-7357

1 Introduction

Many of the features that distinguish owner-occupied housing from other goods imply it is traded in a specialized and peculiar market. The heterogeneity of housing, its indivisibility and durability, its fixed location, and the large capital requirement for home purchase represent substantial deviations from the textbook ideal of the perfect market. Housing is both a consumption good and a store of wealth, but its market looks neither like conventional goods nor capital markets. Housing markets are decentralized, characterized by high transactions and search costs, the complete absence of futures markets, and a dearth of professional traders willing to buy and sell houses in the way equities and commodities are routinely traded. Given the many frictions associated with housing transactions, it should be no surprise to find housing price behavior at odds with that predicted by simple models of housing markets.[1] Indeed, since Case and Shiller (1990) reported that both real and excess returns were forecastable, several other researchers (Guntermann and Norrbin (1991), Gatzlaff (1994), and Malpezzi (1999), among others) have documented predictable returns in housing markets, typically by demonstrating that some aggregate price series exhibit "inertia" in returns. However, little is known about the dynamics of individual housing prices or whether the "inertia" found in aggregate prices has foundations the level of the dwelling.

This paper examines the dynamics of owner-occupied housing markets using a body of data uniquely suited to this task. The data consist of *each and every house sale* in Sweden during a thirteen-year period. We develop a model of housing prices that more closely represents market features that are specific to housing markets, incorporating a more general, and more appropriate, structure of prices at the level of the individual dwelling. This supports a direct test of the hypothesis that individual dwelling prices follow a random walk against the alternative hypothesis of mean reversion. We then link these results to movements in aggregate measures of housing prices and their dynamic properties. The model is more general than other widely-used methods of measuring aggregate housing prices.[2] Our

[1] Grossman and Stiglitz (1980) show that positive information costs alone preclude an informationally efficient market.

[2] For example, the "standard method"–used extensively in academic and professional research and reported by government agencies (i.e., OFHEO)–is shown to be a special case of the model developed below.

framework supports explicit tests of the assumptions implicit in the conventional models.

Maximum likelihood results clearly indicate that, at the individual dwelling level, house prices *do not* follow a random walk. It is evident that the error variances in prices are an increasing function of time between sales, but the variances do not conform to the behavior implied by a random walk. The high, but not perfect, serial correlation in the error structure at the individual dwelling level implies that changes in aggregate housing prices will also be serially correlated in a predictable way. The empirical results indicate that shocks to prices persist over several years, offering a powerful explanation for the observed predictability in aggregate housing prices.

Finally, the paper addresses whether there exists an investment rule that consistently yields extra-normal returns, given this understanding of the structure of aggregate returns. Using data on repeat sales of the same, unchanged, dwelling and standard bootstrap techniques, the value of this information to different economic agents is examined. The results of this section show that–while returns are somewhat predictable–the institutional structure of housing markets precludes consistent excess returns. In fact, the simulations suggest that economic agents in housing markets compete away the excess profits to approximately the level of the transactions costs of buying and selling housing for the typical participant. The results should be generalizable to other housing markets–as detailed below, the costs associated with entry and exit to the markets studied are similar to those in other markets.

Section 2 develops a general model of housing prices that permits an explicit test for a random walk in individual housing prices. This section demonstrates that a widely employed method of measuring housing prices, proposed by Bailey, Muth, and Nourse (1963) and extended by Case and Shiller (1987), is a special case of the general model. The data are described in Section 3. Section 4 presents a set of empirical results: tests for random walk in house prices against the alternative of a mean reverting process, and the link between individual pricing errors and aggregate price movements. Tests for the predictability of aggregate returns appear in Section 5, as do evaluations of the profitability of plausible investment rules available to market actors. Section 6 concludes.

2 A Simple Model of Housing Prices

Price discovery in housing markets cannot occur as it does in many other contexts. Unlike routinely purchased goods, housing transacts infrequently. Most other goods can be transported to the market offering the highest return; this is impossible in the case of housing. The fixed location of housing implies that even dwellings with identical physical structures may differ in price simply because the price incorporates a complicated set of site-specific amenities and costs. Furthermore, the stock of housing is characterized by diversity–dwellings vary widely across structural attributes, style, and vintage. In short, "comparison shopping" in housing markets is more problematic than it is in many other goods markets; it is substantially more difficult to determine the market price of a dwelling when every other previously sold dwelling is necessarily an imperfect substitute.

As a result, housing markets are characterized by a costly process matching heterogeneous agents on both sides of a transaction involving a heterogeneous good. The expensive and time-consuming search in which buyers and sellers engage implies that, in housing markets, sales prices are determined by a small number of participants. Certainly, the observed sale price of any dwelling may deviate substantially from one that would be obtained if information in housing markets were costless.

In practice, buyers, sellers, and their agents estimate the "market price" of a home by utilizing the information embodied in the set of previously sold dwellings. The usefulness of any one of these sales as a reference depends its similarity across several dimensions: physical, spatial, and temporal. Because housing trades infrequently, the arrival of new information about market values is slow. Indeed, from an informational standpoint, the closest comparable sale across these various dimensions may be the previous sale of the same dwelling.

The effort 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 market value, but also of unobserved buyer and seller characteristics (Quan and Quigley 1991). For any given sale, all that is known is that an offer was received that was higher the owner's reservation price.

The model presented below addresses the price dynamics of housing markets at the level of the individual dwelling by accounting for two sources of stochastic error. The first is white noise at the time of sale, reflecting unobserved buyer and seller characteristics in a thin market. The urgency to purchase on the part of the buyer and the holding costs incurred by the the seller are two examples of factors that can cause the price of a particular dwelling to deviate from the price that would be obtained in a thick market, i.e., a market with many bidders and comparable dwellings for sale.

The second type of pricing error characterizes the price dynamics in housing markets at the level of the individual dwelling. In a perfect housing market, "errors" in price would reflect new information about the market's valuation of the dwelling and would fully and permanently be incorporated into the dwelling's price. In this case housing prices would follow a random walk. In an imperfect market, pricing errors could persist because market participants either fail to perceive them or lack the ability to exploit them.

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

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

where V_{it} is the log of the observed sales price of dwelling i at time t, P_t is the log of aggregate housing prices. Q_{it} is the log of housing quality. Housing quality is parameterized by X_{it}, the set of relevant dwelling attributes, and β, a vector of coefficients from which implicit prices can be derived for each attribute. The stochastic component is a composite error,

$$(2) \quad \xi_{it} = \varepsilon_{it} + \eta_{it},$$

reflecting the two sources of uncertainty in the model discussed above: that which occurs at the time of sale, η_{it}, and that which persists over time, ε_{it}. η_{it} is white-noise, with mean 0 and variance σ^2_η . As discussed above, the persistence of pricing errors reflects the process by which the housing market incorporates new information about the market price of a dwelling. The arrival of new information in the form of other dwelling sales will eventually eliminate the previous pricing error. We model this persistence as an autocorrelated process:

$$(3) \quad \varepsilon_{it} = \lambda\varepsilon_{i,t-1} + \mu_{it},$$

where μ_{it} is distributed with mean 0 and variance σ_μ^2.

If $\lambda < 1$, the first two moments of ξ_{it} are finite and given by

$$(4) \quad E[\xi_{it}] = E[\varepsilon_{it} + \eta_{it}] = E[\lambda\varepsilon_{i,t-1} + \mu_{it} + \eta_{it}] = \sum_{s=0}^{\infty} \lambda^s E[\mu_{t-s}] + E[\eta_{it}] = 0,$$

and, using $E[\mu_{it}\eta_{j\tau}] = 0 \, \forall \, \{i, j, t, \tau\}$ and $E[\mu_{it}\mu_{j\tau}] = 0 \, \forall \, \{i \neq j, t \neq \tau\}$,

$$(5) \quad E[(\xi_{it})^2] = E[(\varepsilon_{it} + \eta_{it})^2] = \sum_{s=0}^{\infty} \lambda^s E[\mu_{t-s}^2] + E[\eta_{it}^2] = \sigma_\mu^2 \sum_{s=0}^{\infty} \lambda^s + \sigma_\eta^2 = \frac{\sigma_\mu^2}{(1-\lambda^2)} + \sigma_\eta^2.$$

Because housing sales are infrequent, the covariance is more usefully defined for general intervals between sales, i.e.,

$$
\begin{aligned}
(6) \quad V\left[\xi_{it}, \xi_{i\tau}\right] &= E\left[\xi_{it}\xi_{i\tau}\right] \\
&= E\left[(\varepsilon_{it} + \eta_{it})(\varepsilon_{i\tau} + \eta_{i\tau})\right] \\
&= E\left[\varepsilon_{it}\varepsilon_{i\tau}\right] + E\left[\eta_{it}\eta_{i\tau}\right] \\
&= E\left[(\lambda^{t-\tau}\varepsilon_{i\tau} + \sum_{j=0}^{\tau} \mu_{t-j})\varepsilon_{i\tau}\right] + 0 \\
&= \lambda^{t-\tau} E\left[\varepsilon_{i\tau}^2\right] \\
&= \lambda^{t-\tau}\left(\frac{\sigma_\mu^2}{1-\lambda^2}\right)
\end{aligned}
$$

The error structure developed in this section is identified by multiple observations of sales on the same dwelling. Note also that, if the null hypothesis that the autocorrelation coefficient λ equals one holds, the derivation of the moments of the error structure presented above are not meaningful, as the infinite series in equations (5) and (6) do not converge. Thus, if prices follow a random walk, the unconditional variance of ε_{it}, and therefore of ξ_{it}, does not exist.

However, with multiple sales the hedonic model, equation (1), can be differenced

$$(7) \quad V_{it} - V_{i\tau} = X_{it}\beta + P_t - X_{i\tau}\beta - P_\tau + \xi_{it} - \xi_{i\tau}.$$

If dwellings are unchanged between sales,[3] the model simplifies to

$$(8) \quad V_{it} - V_{i\tau} = P_t - P_\tau + \xi_{it} - \xi_{i\tau}.$$

[3]This assumption is discussed below.

This can be estimated with the regression

$$(9) \quad V_{it} - V_{i\tau} = D_{it\tau} + \Xi_{it\tau},$$

where $D_{it\tau}$ is a matrix of dummy variables indicating time of sale; $D_{it\tau}$ takes -1 at the time of the first sale, +1 at the second, and 0 elsewhere. $\Xi_{it\tau}$ is the differenced stochastic terms, $\xi_{it} - \xi_{i\tau}$.

Note that in this form, the unconditional variance of the stochastic term exists even if the error process follows a random walk. Indeed, when $\lambda = 1$, the variance of the error term $\Xi_{it\tau}$ is linear in the time between sales, that is, for sales at t and τ,

$$(10) \quad E[(\Xi_{it})^2 | t - \tau < \infty] = \sigma_\mu^2 (t - \tau).$$

The covariance matrix associated with the regression based on multiple sales, equation (9), is block diagonal with each block representing the set of repeat sales of an unchanged dwelling. The general form of the covariance matrix, $\mathbf{\Psi}$, is

$$(11) \quad \mathbf{\Psi} = E[\Xi_{it\tau}, \Xi_{j\gamma}] = \left(\frac{\sigma_\mu^2}{1-\lambda^2}\right)\left(\lambda^{t-g} - \lambda^{t-\gamma} - \lambda^{\tau-g} + \lambda^{\tau-\gamma}\right) + \sigma_\eta^2 \left(I_{tg} - I_{t\gamma} - I_{\tau g} + I_{\tau\gamma}\right)$$

The indicator variables, I_{jk} equal 1 if $j = k$ and 0 otherwise. Consider the three types of elements which comprise the covariance matrix. The diagonal elements, the variances of each draw of $\Xi_{it\tau}$ (where $t = g$ and $\tau = \gamma$) are given by

$$(12) \quad V[(\xi_{it} - \xi_{i\tau}), (\xi_{it} - \xi_{i\tau})] = \left(\frac{2\,\sigma_\mu^2}{1-\lambda^2}\right)\left(1 - \lambda^{t-\tau}\right) + 2\,\sigma_\eta^2.$$

Whenever unchanged dwellings are sold three or more times, there exist "adjacent" paired sales.[4] These are given by

$$(13) \quad V[(\xi_{it} - \xi_{i\tau}), (\xi_{i\tau} - \xi_{i\gamma})] = \left(\frac{\sigma_\mu^2}{1-\lambda^2}\right)\left(\lambda^{t-\tau} - \lambda^{t-\gamma} - 1 + \lambda^{\tau-\gamma}\right) - \sigma_\eta^2.$$

Whenever dwellings are sold four or more times, there exist "non-adjacent" paired sales. In these cases, there is no individual sale common to either paired sale. These elements are defined by

$$(14) \quad V[(\xi_{it} - \xi_{i\tau}), (\xi_{ig} - \xi_{i\gamma})] = \left(\frac{\sigma_\mu^2}{1-\lambda^2}\right)\left(\lambda^{t-g} - \lambda^{t-\gamma} - \lambda^{\tau-g} + \lambda^{\tau-\gamma}\right)$$

[4] Adjacent paired sales are those which share an individual sale. That is, if dwelling i sells three times (t, τ, γ), then the first paired sales are at times t and τ, while the second results from the two sales at τ and γ. They are "adjacent" in the sense that both pairs share the observation at time τ.

Note that, if indeed λ does equal one, then the model developed above collapses into the "weighted repeat sales" model proposed by Case and Shiller (1987) and widely employed in academic research[5]

The weighted repeat sales model uses assumptions about the error structure in house prices to generate efficient parameter estimates of the effect of time on aggregate housing price levels. Case and Shiller argue that there is a drift in housing "value," that it follows a Gaussian random walk, and that the variance of housing prices is therefore linear in the time between sales. The weighted repeat sales procedure, as typically implemented, makes this explicit assumption about the form of heteroskedasticity in addition to an implied assumption concerning the covariance between any two paired sales. Specifically, it is assumed that the covariances are zero *everywhere*. Neither of these assumptions has been tested formally. However, because the "weighted repeat sales" method is a special case of the more general model developed above, a joint test of its maintained hypotheses can be developed by testing the null hypothesis that $\lambda = 1$ against $\lambda \neq 1$.

3 Data

We rely upon data compiled by Statistics Sweden consisting of *all* arms-length sales of single family housing in Sweden from 1981 to 1993. The data are unique both in their breadth– *each* housing sale in Sweden is recorded - and in their detail, with an extensive array of

[5]To see this, reconsider the structure of the repeat sales model developed above. Specifically require that $\lambda = 1$ in the repeat sales regression, that is

$$
\begin{aligned}
(15) \quad V_{it} - V_{i\tau} &= P_t - P_\tau + \varepsilon_t - \varepsilon_\tau + \eta_{it} - \eta_{i\tau} \\
&= P_t - P_\tau + (\lambda^{t-\tau} - 1)\varepsilon_\tau + \sum_{s=0}^{t-\tau-1} \lambda^s \mu_{t-s} + \eta_{it} - \eta_{i\tau}. \\
&= P_t - P_\tau + \sum_{s=0}^{t-\tau-1} \mu_{t-s} + \eta_{it} - \eta_{i\tau}.
\end{aligned}
$$

This equation is the same regression employed, either directly and indirectly, in many applications and is the same technique utilized by the Office of Federal Housing Enterprise Oversight (OFHEO) in the construction of their aggregate price indexes. Variants on this index are marketed by private firms to portfolio investors. The variance structure of the weighted repeat sales method can be derived from the general model developed above. If λ is required to be 1, the covariance matrix is diagonal with elements $\sigma_\mu^2(t-\tau) + 2\sigma_\eta^2$. The model can be estimated by generalized least squares.

physical characteristics reported. The research reported below distinguishes among eight administrative regions—four metropolitan regions (Gothenburg, Malmö, Stockholm, and Uppsala) and four non-metropolitan regions (the South, the Central, the North, and the Far North). The data are discussed at length in Englund, Quigley, and Redfearn (1998), but several features of the data central to this analysis warrant discussion.

First, the detailed set of characteristics that describe the physical structure make it possible to verify the central assumption of repeat sales models of housing prices, which is that twice-sold dwellings are identical in attributes at each time of sale. In practice, this assumption is difficult to verify, but the attributes reported in the data set make this possible. The data contain not only primary characteristics such as lot size and living area, but also additional variables that offer a comprehensive description of the dwelling, including the number of garages, kitchen quality, type of wall, roof, and floor, and the presence of amenities such as sauna, fireplace, and furnished basement. The list is extensive—in excess of thirty housing characteristics. This level of detail facilitates the detection of structural changes that would violate the assumption of constant quality.[6] The research reported here uses only those repeat sales of homes whose measured characteristics are unchanged between the first and the second sales.

Second, the population of dwellings is a panel so that units that sold more than once can be distinguished from those that did not. The panel nature of the data identifies the appropriate specification of the model developed in Section 2.

Table 1 previews the central empirical question addressed in this paper: the relationship between the length of the time between sales and the distribution of prices. The table shows the average and standard deviation of the total return from price appreciation for dwellings on which there are two or more observed sales. Both statistics exhibit strong tendencies to decline as the number of sales increases. Of course, the average time between sales declines with the number of sales, since the sample period is fixed. Because the majority of the sales during the sample period are drawn from a period of rising nominal housing prices, it is not

[6]Englund, Quigley, and Redfearn (1999a) examine in detail the validity of the constant quality assumption. They find that is does not hold in general, and that failure exclude altered dwellings can lead to substantial bias in measured aggregate house prices.

Table 1: Distribution of Price Changes for Paired House Sales

Average Price Appreciation Between Sales, Std. Dev. in Parentheses

Number of	Metropolitan Regions				Non-Metropolitan Regions			
Paired Sales	Stockholm	Uppsala	Gothenburg	Malmö	South	Central	North	Far North
1	0.375	0.256	0.283	0.298	0.236	0.241	0.217	0.240
	(0.51)	(0.37)	(0.46)	(0.60)	(0.35)	(0.38)	(0.48)	(0.37)
2	0.320	0.210	0.215	0.245	0.202	0.186	0.157	0.187
	(0.49)	(0.32)	(0.32)	(0.38)	(0.28)	(0.26)	(0.25)	(0.26)
3	0.273	0.174	0.167	0.190	0.135	0.143	0.103	0.165
	(0.38)	(0.25)	(0.25)	(0.29)	(0.21)	(0.23)	(0.17)	(0.22)
4	0.197	0.153	0.084	0.147	0.109	0.133	0.103	0.121
	(0.29)	(0.22)	(0.17)	(0.24)	(0.27)	(0.17)	(0.13)	(0.13)
5	0.174	0.116	-	0.077	-0.006	0.053	0.020	0.118
	(0.21)	(0.17)	-	(0.14)	(0.08)	(0.08)	(0.06)	(0.17)
6	-	0.094	0.132	-	-	-	0.021	0.074
	-	(0.15)	(0.28)	-	-	-	(0.07)	(0.11)
7	-	0.125	-	-	-	-	-	-
	-	(0.12)	-	-	-	-	-	-
Observations on Paired Sales	7714	11052	11409	11017	5972	6287	2626	3594

surprising that total appreciation is higher when the average time between sales is longer. However, the standard deviation of price appreciation shows consistent reduction in the volatility in prices between sales as the interval between sales decreases. This is consistent with the hypothesis that heteroskadasticity in housing price appreciation is a function of the time interval between sales of the same dwelling.

4 Do House Prices Follow a Random Walk?

The model presented above supports an explicit test the hypothesis that individual housing prices follow a random walk against the alternative hypothesis that prices follow a mean reverting process. The test is implemented by maximizing the concentrated log-likelihood function

$$(16) \quad logL = -n * (\mathbf{e}'_{\mathbf{gls}} \mathbf{\Psi}\ \mathbf{e}_{\mathbf{gls}}) - det|\mathbf{\Psi}|,$$

where n is the number of observations and $\mathbf{e}_{\mathbf{gls}}$ is the vector of residuals from a generalized least squares regression. $\mathbf{\Psi}$ is the estimated covariance matrix described by (11), a function of λ. Well-defined probabilistic statements about λ can be made using likelihood-ratio tests.

As discussed in Section 2, the covariance matrix, defined in (11), is block diagonal with the dimension of each block determined by the number of sales of an unchanged dwelling. The elements of the covariance matrix depend on the time between sales, the value λ, and the variances of the pricing errors defined in (1), σ^2_μ and σ^2_η. Consistent estimates of these error variances are obtained from the regression

$$(17) \quad e_{it\tau}e_{ig\gamma} = \begin{cases} 2\sigma^2_\eta & + \quad h(t,\tau,g,\gamma)\sigma^2_\mu & \text{if } t{=}g,\ \tau = \gamma & \text{(diagonal elements)}, \\ -1\sigma^2_\eta & + \quad h(t,\tau,g,\gamma)\sigma^2_\mu & \text{if } \tau = g, & \text{(off-diagonal elements)}, \\ 0\sigma^2_\eta & + \quad h(t,\tau,g,\gamma)\sigma^2_\mu & \text{otherwise} & \text{(off-off diagonal elements)}. \end{cases}$$

The dependent variables, $e_{it\tau}$ and $e_{ig\gamma}$ are obtained from the appropriate elements of the outer product $\mathbf{e} \cdot \mathbf{e}'$, where $\mathbf{e}$ is the vector of residuals from a first-step regression. $h(t,\tau,g,\gamma)$ is an element of the covariance matrix, given by (11), defining the expected covariance between two paired sales of the same dwelling. σ^2_η and σ^2_μ are the estimated regression coefficients.

The value of λ that maximizes the log-likelihood function, $\tilde{\lambda}$, is easily obtained from this procedure, as is the value of the log-likelihood function for other values of λ.

Figure 1 illustrates the nature of the maximization problem for one of the metropolitan areas. The upper panel shows the sample error variances and the relative frequency of observations by elapsed time, in quarter years, between sales. Two things are clear from this panel. First, the error variances are an increasing function of the time interval between sales. Second, the large majority of observations are paired sales in an interval shorter than than six years. The very high sample variances observed in the longer intervals are based on small numbers of observations.

The covariance matrix described in (11) defines the predicted variance of the price appreciation as a function of two variables: the elapsed time between sales and the correlation of the errors. This relationship is illustrated in the lower panel of Figure 1. It again plots the sample error variances shown in the upper panel, but also the predicted variances using several assumed values of λ.

Clearly, the three lines with positive serial correlation fit the data far better than does the line implied by no serial correlation. For a random walk, the variance is linear in the elapsed time between sales; this would occur if the market permanently incorporated pricing innovations.

The lines representing intermediate values of serial correlation increase asymptotically to a level given by the standard definition of the (unconditional) variance of an autocorrelated process. That is, if the errors follow (3), and $|\lambda| < 1$, then

$$(18) \quad E[(\varepsilon_{it})^2] = \frac{\sigma_\mu^2}{(1-\lambda^2)}.$$

This implies that for any paired sale, the unconditional variance of price appreciation increases with the correlation coefficient. That is, pricing errors for a particular dwelling persists as a function of the arrival and incorporation of new information about housing prices.

Results of the estimation procedure for the both metropolitan and non-metropolitan regions are plotted in Figure 2. The results offer strong visual support against the random

Figure 1:

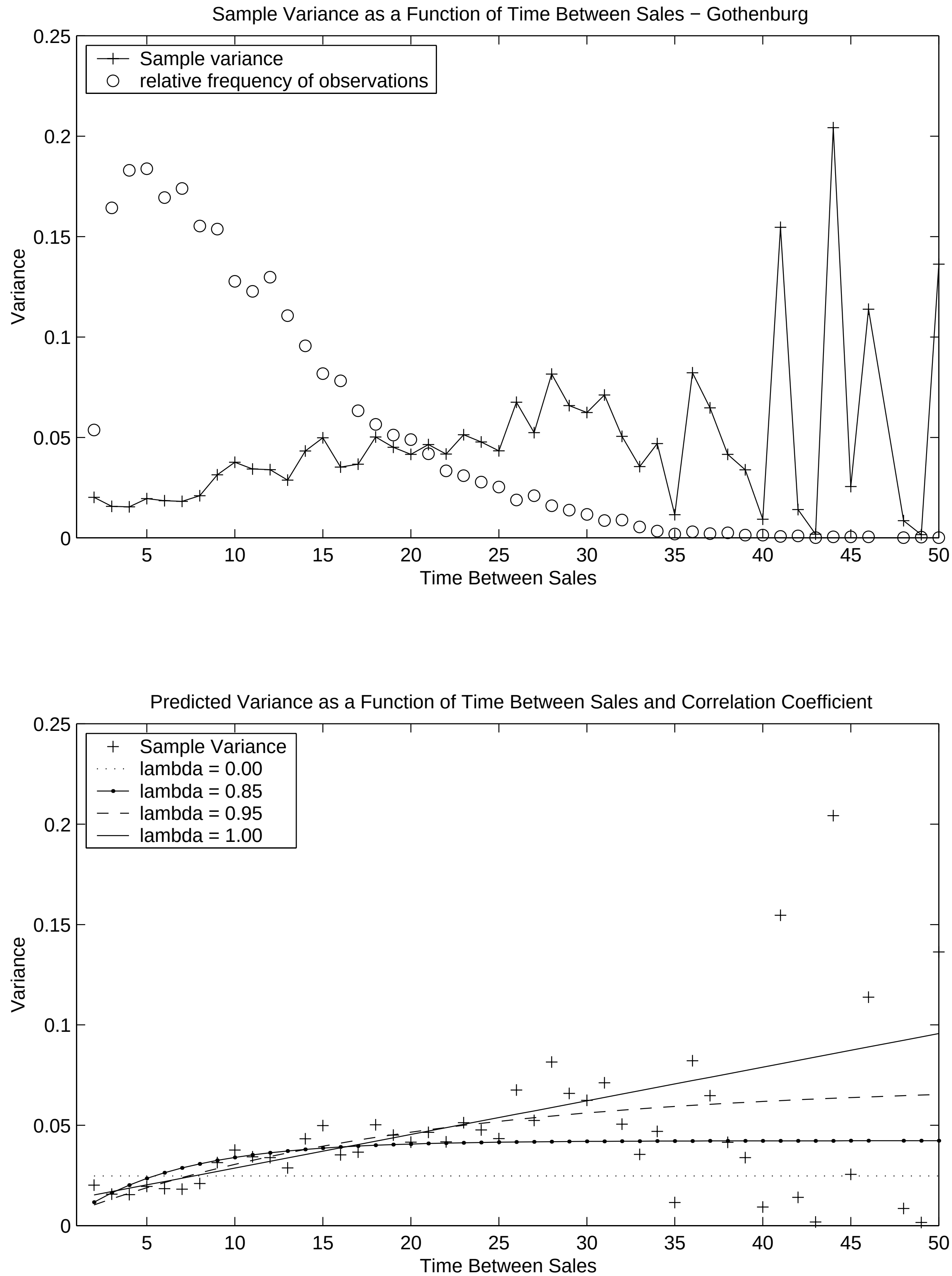

Sample Variance as a Function of Time Between Sales − Gothenburg
Sample variance
relative frequency of observations
Variance
Time Between Sales
Predicted Variance as a Function of Time Between Sales and Correlation Coefficient
Sample Variance
lambda = 0.00
lambda = 0.85
lambda = 0.95
lambda = 1.00
Variance
Time Between Sales

Figure 2: Normalized Likelihood Functions for the Eight Regions

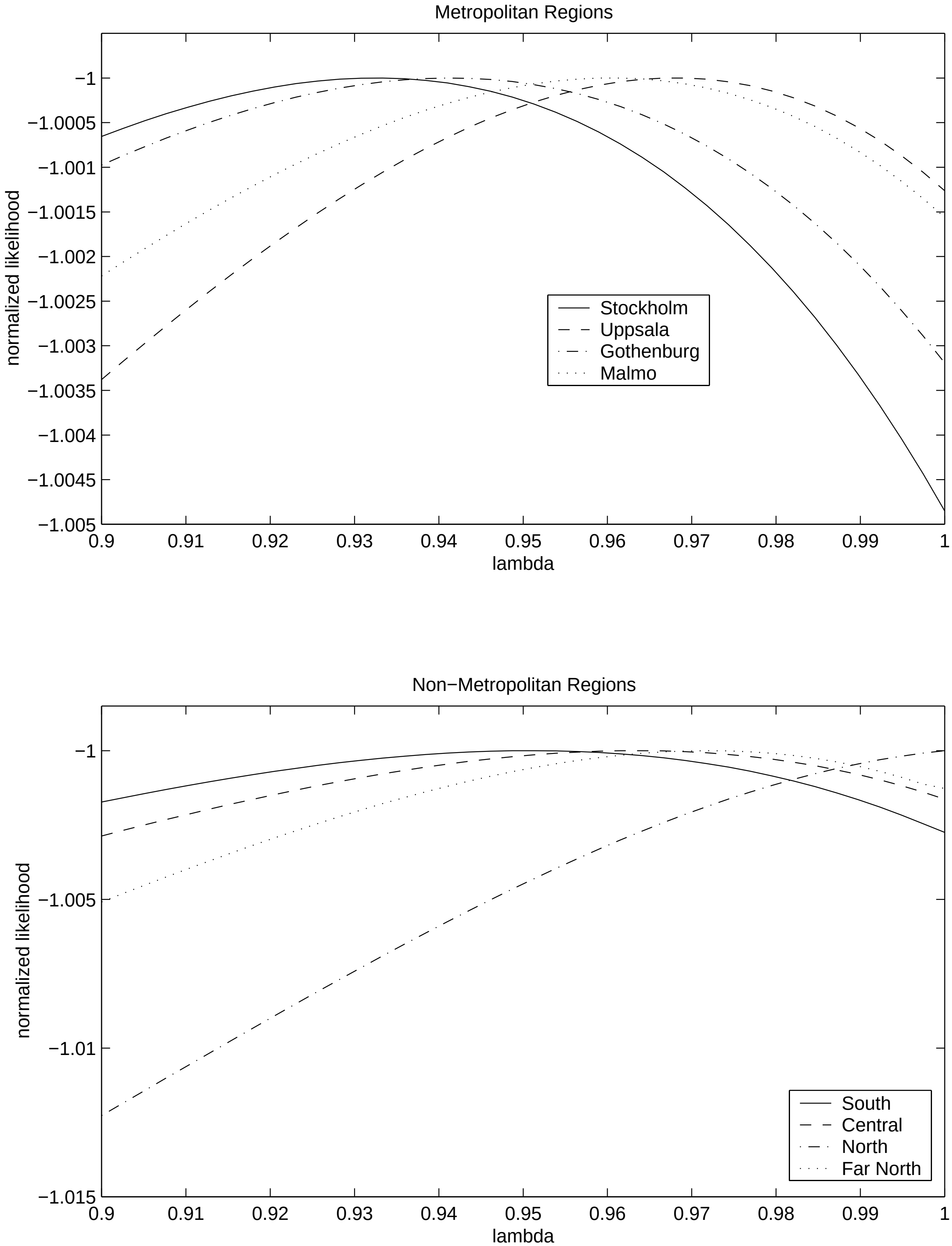


Table 2: Tests for a Random Walk in Housing Prices

		95% confidence interval	
Region	$\tilde{\lambda}$	Lower Bound	Upper Bound
Stockholm	0.933	0.923	0.944
Uppsala	0.969	0.962	0.974
Gothenburg	0.941	0.933	0.949
Malmö	0.962	0.951	0.969
South	0.951	0.941	0.962
Central	0.964	0.954	0.974
North	1.000	0.992	1.000
Far North	0.972	0.962	0.982

walk, with $\tilde{\lambda}$ apparently less than 1 in all but one region. (Time is measured in quarter-year intervals.)

Formal tests of these conclusions are reported in Table 2. For seven of the eight regions the upper bound of the 95 percent confidence interval does not include 1. For these seven, the maximum likelihood value of λ ranges from 0.93 to 0.97. These values suggest that the housing market removes pricing errors only quite slowly. A quarterly serial correlation coefficient of 0.95 implies that the market will eliminate only one-fifth of a prior pricing error over the course of year. That is, after three years just over half of the initial discrepancy will remain. The discrepancy will not be reduced to ten percent of its initial value for approximately ten years.

It is straightforward to consider the effect of these individual lags upon aggregate measures of housing prices. Consider an economy of identical and unchanging dwellings. Let housing quality be normalized to one so that the log of quality, $X_{it}\beta$, is zero for all i and t. Let an estimate of the housing price level be the mean price of the dwellings sold in each period. That is, let

$$(19) \quad \widehat{P_t} = n^{-1} \sum_i V_{it} = n^{-1} \left(\sum_i P_t + \sum_i \varepsilon_{it} + \sum_i \eta_{it} \right)$$

The true housing price index, P_t, is a constant and the mean of the white-noise errors, η_{it}, is zero. Denote the average of the the autocorrelated errors as

$$(20) \quad \overline{\varepsilon}_t = n^{-1} \sum_i \varepsilon_{it}.$$

The estimator of aggregate prices is

$$(21) \quad \widehat{P_t} = P_t + \overline{\varepsilon}_t.$$

The unconditional expectation of $\widehat{P_t}$ is unbiased, since $E\left[\overline{\varepsilon}_t\right] = 0$. However, with an aggregate shock to prices one period earlier, $\overline{\varepsilon}_{t-1} = e > 0$, the aggregate price estimator is no longer unbiased. The mismeasurement is

$$(22) \quad \widehat{P_t} = P_t + \overline{\varepsilon}_t = P_t + \lambda \overline{\varepsilon}_{t-1} = P_t + \lambda e.$$

Moreover, aggregate pricing errors will lead to correlations in aggregate returns to housing for many periods if, as indicated above, λ is close to one.

"Inertia" in aggregate prices has become an accepted empirical regularity in housing markets. Case and Shiller (1990) argue that the predictability in housing returns arises from the failure of housing markets to incorporate predictable interest rates over the course of their sample period. Gatzlaff (1994) finds that forecastability is substantially reduced, but not eliminated, if anticipated inflation is controlled for. Thus, predictable fundamentals may not be incorporated efficiently into prices. However, the results presented in this section suggest that persistence at the individual dwelling level can cause persistence at the aggregate level. If household expectations about housing prices are adaptive, unexpected shocks to housing prices can lead to predictable changes in aggregate housing prices because the market slowly incorporates information about the shock. This also accounts for the empirical regularity that housing returns are correlated.

5 The Predictability of the Returns to Housing

This section examines housing market price dynamics in two ways. First, it tests formally the proposition that aggregate returns are forecastable. Previous research has found that both real and excess returns are predictable. The results presented here confirm this but use a more appropriate data set and technique to conduct the tests. The relevant question, however, is not whether returns are predictable, but rather whether there is a strategy that consistently realizes excess profits. The second half of the section focuses on this question using simulations in which the costs of entry and exit are explicitly considered for several different economic agents.

5.1 Aggregate Returns

The first series considered is the nominal return due to capital appreciation in house prices. Price appreciation, π_t, is simply

$$(23) \quad \pi_t = \frac{P_t}{P_{t-1}} - 1,$$

where P_t is the index of aggregate housing prices.

The second series is real returns, which includes not only the return due to price appreciation but also the "dividend" from the implicit rent paid to owner-occupiers. This total return is discounted by the change in cost of living index (less shelter). The real return at time t, r_t, is given by

$$(24) \quad r_t = \frac{P_t + R_t}{P_{t-1}} - \frac{CPI_t}{CPI_{t-1}},$$

where R_t and CPI_t are,respectively, the implicit rent and cost of living indexes at time t.

The aggregate housing price index employed in this section is constructed using the hybrid model developed in our companion paper, Englund, Quigley, and Redfearn (1998). The hybrid method uses all available sales information to construct the aggregate price index. Not only are the large majority of housing transactions during the sample period single sales, many of those which do sell more than once are altered between sales and, for the purposes of measuring aggregate prices, are no longer considered repeat sales. The usable data sets differ by approximately a factor of five for the Swedish data used in this paper.[7]

[7] Sample selectivity is also a problem in the repeat sales data sets. Gatzlaff and Haurin (1997), Gatzlaff and Haurin (1998), and Englund, Quigley, and Redfearn (1999a) find that the sample of sold dwellings are not a random sample of all dwellings and that the appreciation rates of those homes selected into the observed sample is not reflective of the stock as a whole.

Moreover, the appropriate approach for testing the predictability of returns is to approximate, as closely as possible, the information set available to the investor at each point in time. The small sample problems inherent in repeat sales models are exacerbated over short intervals.[8] In order to avoid this problem, typically the aggregate housing price index is estimated for the entire sample period and it is assumed that investors observe prices only up to the time period of their decision. However, this approach does not address the fact that the indexes are conditional on all repeat sales over the *entire* sample period. This means that when forecasting changes forward from time t, the estimated price index is calculated using sales information from $t+1$ on. The repeat sales indexes used to evaluate the forecastability of housing returns are, in fact, not the indexes available to agents at time t.

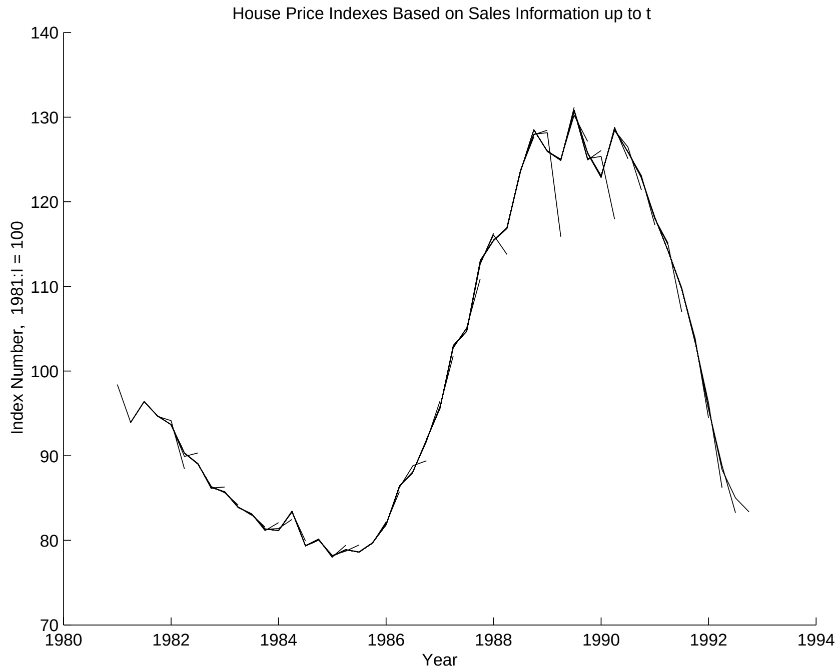
If there were sufficient observations, repeat sales indexes, and the returns series dependent on them, could be estimated at each point in time. In general, the small fraction of sales that are repeats precludes period-by-period estimation. Use of the hybrid method avoids this problem. It is possible, using the Swedish data described above, to estimate an aggregate price index and each of the returns defined above at each point in time, accurately reflecting the information available when the housing investment decision is made. Figure 3 plots the last eight quarters of aggregate price indexes estimated quarter by quarter from 1982:I forward using the hybrid method. Despite the fact that the hybrid method is substantially less vulnerable to *ex post* revisions, the figure shows that the estimated indexes do not take the same value at the time of estimation as they do when taken from an estimated index that spans the sample period.

Calculation of the real return series requires an estimate of the implicit rent that accrues to owners who live in their home. [9]

[8] This problem is discussed and measured in Englund, Quigley, and Redfearn (1999b). They find the confidence interval around measured prices is vastly larger for the repeat sales method when compared with those of the hybrid method.

[9] Previous research has employed local rent indexes as a proxy for the implicit rents, but this is less appropriate in the Swedish context. The rental market is highly regulated; rents are based on construction costs rather than on supply and demand in spot markets. Moreover, at the regulated price, rental markets do not clear: there are lengthy waits for apartments in Stockholm. Instead, the proxy used in this research is obtained by assuming that the implicit rent is equated to the real return on the value of the home.

Figure 3:



After constructing the two series described above, their forecastability is tested by estimating an autoregressive process for each. For example, the regression for price appreciation is given by

$$(25)\quad \pi_t = \kappa + \beta_1\pi_{t-1} + \beta_2\pi_{t-2} + \beta_3\pi_{t-3} + \beta_4\pi_{t-4} + \omega_t,$$

where ω_t is white noise. The regression for real returns is analogous. Forecastability is accepted if the set of coefficients, β, are jointly non-zero.

Tables 3 and 4 shows the results of the regressions

for each of the return series for each of the four metropolitan regions. The table indicates that, with the exception of Gothenburg, all three return series are "predictable." In Stockholm, Uppsala, and Malmö the models explain between twenty and forty percent of the variance in returns. The models use 39 quarters of data and estimate four coefficients, implying F-statistics of 4.02, 2.69, and 2.14 for one-, five-, and ten-percent level of significance, respectively. The null hypothesis that the set of coefficients on the lagged returns is zero is rejected for both series in Stockholm, Uppsala, and Malmö. These results confirm the previous research–aggregate housing returns are predictable.

Table 3: Forecasting Aggregate Housing Returns: Nominal

Dependent Variable: nominal return, π_t
Observations: n= 46

Region	I	II	III	IV	V	VI	VII	VIII
R^2	0.42	0.38	0.22	0.24	0.15	0.12	0.41	0.27
$F - test$	7.57**	6.32**	2.93*	3.35*	1.92	1.43	7.28**	3.96**
Intercept	0.002	0.003	0.005	0.002	0.003	0.009	0.001	0.002
	(0.38)	(0.58)	(0.67)	(0.34)	(0.38)	(1.20)	(0.10)	(0.22)
π_{t-1}	0.252	−0.131	−0.272	−0.028	0.030	−0.239	−0.516	−0.583
	(1.90)	(0.84)	(2.50)	(0.24)	(0.19)	(1.26)	(4.06)	(3.31)
π_{t-2}	0.315	0.284	0.003	0.279	0.104	0.175	0.051	0.098
	(2.12)	(2.19)	(0.02)	(2.64)	(0.70)	(0.99)	(0.44)	(0.58)
π_{t-3}	0.168	0.180	0.305	0.176	0.223	0.251	0.422	0.490
	(1.05)	(1.23)	(1.79)	(1.14)	(1.49)	(1.16)	(3.49)	(1.46)
π_{t-4}	0.055	0.418	0.450	0.305	0.287	0.142	0.545	0.454
	(0.36)	(2.25)	(3.48)	(1.56)	(1.14)	(0.60)	(4.16)	(1.56)

note: ** indicates rejection at the 1% level, * at the 5% level

5.2 Is that a $20 Dollar Bill on the Porch?

Housing markets lack the financial instruments needed to exploit the aggregate relationships estimated in the previous section. Of the set of options available in financial markets, only the buy-and-hold strategy exists in housing markets–investment in housing requires taking ownership of a particular property. This appears to be substantially more risky than indicated by aggregate return, the standard deviation of price change reported in Table 1. The tests developed in this section use iterative cross-validation techniques (bootstrap) to determine the value of a simple investment rule.

The investment rule is straightforward: buy if the model predicts returns greater than zero. For real returns, this means that the combined return due to capital gains and implicit rent need only to surpass inflation. For excess returns, the combined return has to exceed the home owner's cost of capital. The profitability of this rule is examined over three investment horizons, 4, 8, and 12 quarters.

Table 4: Forecasting Aggregate Housing Returns: Real

Dependent Variable: real return, r_t Observations: n= 46								
Region	I	II	III	IV	V	VI	VII	VIII
R^2	0.45	0.38	0.23	0.27	0.16	0.18	0.44	0.27
$F - test$	8.70**	6.36**	3.06*	3.81*	1.98	2.24	8.19**	3.85**
Intercept	−0.002	−0.001	−0.004	−0.003	−0.003	−0.002	−0.008	−0.007
	(0.41)	(0.33)	(0.71)	(0.56)	(0.69)	(0.33)	(1.55)	(1.05)
r_{t-1}	0.175	−0.230	−0.315	−0.034	0.002	−0.330	−0.586	−0.615
	(1.26)	(1.50)	(2.78)	(0.28)	(0.01)	(1.83)	(4.43)	(3.83)
r_{t-2}	0.238	0.210	0.018	0.307	0.094	0.154	0.028	0.101
	(1.77)	(1.44)	(0.12)	(2.62)	(0.61)	(0.98)	(0.22)	(0.67)
r_{t-3}	0.245	0.290	0.343	0.197	0.285	0.360	0.466	0.477
	(1.59)	(1.82)	(2.11)	(1.33)	(1.93)	(1.77)	(4.01)	(1.45)
r_{t-4}	0.187	0.488	0.406	0.296	0.256	0.193	0.535	0.421
	(1.42)	(3.18)	(2.88)	(1.43)	(1.16)	(0.79)	(3.99)	(1.47)

note: ** indicates rejection at the 1% level, * at the 5% level

The relationship between past and future returns is estimated by exploiting the panel nature of the data. For each paired sale, both the lagged aggregate returns and realized return *for that particular dwelling* are known. Those paired sales which are sold 4, 8, and 12 quarters apart are extracted from the repeat sales data described above. These observations are then merged with lagged aggregate real and excess returns. The subsample is split–ninety percent of the merged data are used to estimated the relationship between lagged returns and realized returns, ten percent is used to test for excess profitability given this relationship.

For each of the test intervals, the quarterly geometric mean of realized returns in the ninety-percent subsample are regressed on lagged returns to test for predictability over a specific investment horizon. The investment rule can then be evaluated. For each of the remaining observations in the ten-percent subsample–those not used in the regression–the forecasted return is calculated. If the predicted return is greater than zero, the dwelling is “purchased.” Realized returns for “purchased” properties are then compared with those which are not.

The number of observations in any one interval is in the hundreds so that the 90/10 split leaves few observations on which to judge profitability. To overcome this problem, the cross-validation technique just described is repeated. Fifty iterations are executed for each interval for each of the four regions. The results of this process are displayed in several tables below.

Table 5 shows the distribution of forecasts for real and excess returns. It contains the number of sampled dwellings that are forecast to have positive and non-positive returns and the actual return observed over the period between the two sales. It is immediately clear that the combined return due to capital gains and implicit rent exceeded inflation for most of the sample period as almost every sampled dwelling is forecast to have positive real returns. This is not true of the forecasted excess returns. While a clear majority of sampled dwellings is forecast to have positive returns, a sizeable minority is not. The difference in realized excess returns is striking. The average excess return is one to three percentage points higher *per quarter*. This indicates that the predictable relationship observed at the aggregate level holds when measured at the level of the individual dwelling.

However, predictability alone does not indicate the existence of arbitrage opportunities. In order to determine the potential to exploit the relationship between past and future returns, it is necessary to account for the substantial costs of buying, owning, and selling housing. These costs are explained in detail in Söderberg (1995). Table 6 summarizes the major costs of participating in the housing market for four hypothetical home buyers.

The first is referred to as a market timer. This individual faces no transaction costs as they are considered sunk costs–he is living at home with his parents and waiting for the right time to buy. The second buyer is the mover, who incurs the costs of entry and exit but continues to roll any accumulated capital gains back into housing and thus avoids paying capital gains taxes. The speculator is an individual who invests in housing because of the total returns and will move out and sleep in his office when higher returns appear elsewhere. The last of the four buyers is the investor who acquires housing but does not occupy it, and is therefore subject to the rental income he earns from the property.

Table 5: Forecasting Individual Housing Returns

(Returns Reported as Percent per Quarter)

	Realized Real Return Given				Realized Excess Return Given			
	$E[r_{t+i} > 0]$		$E[r_{t+i} < 0]$		$E[ER_{t+i} > 0]$		$E[ER_{t+i} < 0]$	
	obs	mean	obs	mean	obs	mean	obs	mean
A. Interval = 4 Quarters								
Stockholm	666	2.41%	17	1.04%	387	1.89%	296	-1.23%
Uppsala	1069	2.77	-	-	639	1.77	457	-0.15
Malmö	1192	4.09	-	-	1140	2.45	52	1.17
Gothenburg	1288	3.31	-	-	1271	1.54	17	0.30
B. Interval = 8 Quarters								
Stockholm	636	2.64			459	1.29	177	-0.18
Uppsala	1009	2.66	-	-	631	1.32	378	0.11
Malmö	1138	2.76	-	-	768	1.49	370	0.00
Gothenburg	1027	2.21	-	-	603	1.15	424	-0.68
C. Interval = 12 Quarters								
Stockholm	442	3.44			442	1.89	-	-
Uppsala	811	2.79	-	-	750	1.35	61	-2.08
Malmö	720	2.62	-	-	679	1.04	41	-0.06
Gothenburg	743	2.87	-	-	743	1.12	-	-

Table 6: Transactions Costs

Costs reported as percent of house price, unless otherwise indicated

Cost	Type of Buyer: Market Timer	Mover	Speculator	Investor
Registration, Inspection, Assessment	none	2%	2%	2%
Tax on Rental Income	none	none	none	40% (of rent)
Broker's Fee at Sale	none	5%	5%	5%
Capital Gains Tax	none	none	30% (of gain)	30% (of gain)

Clearly, these are simplifications, but the four buyers described above provide enough detail to understand the approximate magnitude of the restraints on competition in housing markets. For example, the tax on rental income is faced by owners who do not occupy their dwelling. This institutional feature of housing markets penalizes the "investor" relative to the other buyers, and represents a substantial barrier to entry for economic agents who are strong forces for efficiency in other markets.

Based on the investment rule employed above ("purchase" if expected returns are positive, realized profits from trade are calculated and presented in Tables 7 and 8. Table 7 shows that, even after incorporating the appropriate costs, the real returns to housing are generally positive for all four types of buyer. The increasingly higher costs facing the mover, the speculator, and the investor, significantly reduce the realized real returns for each. In the shortest interval, the investor experiences negative real returns, but as the interval increases real returns rise, reflecting the amortization of the fixed costs incurred at purchase and sale.

Table 8 shows the realized excess returns for the same four types of buyers. The table

Table 7: Real Returns for Different Types of Home Buyers

"Realized" Real Return Given $E[r_{t+i} > 0]$
(Real Return Reported as Percent per Quarter)

	Market Timer	Mover	Live-in Speculator	Outside Investor
A. Interval = 4 Quarters				
Stockholm	2.41%	0.60%	-0.02%	-0.72%
Uppsala	2.77	0.94	0.19	-0.49
Malmö	4.09	2.17	0.97	0.32
Gothenburg	3.31	1.43	0.50	-0.18
B. Interval = 8 Quarters				
Stockholm	2.64	1.65	0.89	0.27
Uppsala	2.66	1.65	0.87	0.25
Malmö	2.76	1.77	0.98	0.36
Gothenburg	2.21	1.26	0.68	0.04
C. Interval = 12 Quarters				
Stockholm	3.44	2.67	1.57	1.00
Uppsala	2.79	2.07	1.24	0.65
Malmö	2.62	1.92	1.13	0.55
Gothenburg	2.87	2.14	1.27	0.69

Table 8: Excess Returns for Different Types of Home Buyers

"Realized" Excess Return Given $E[r_{t+i} > 0]$
(Excess Return Reported as Percent per Quarter)

	Market Timer	Mover	Live-in Speculator	Outside Investor
A. Interval = 4 Quarters				
Stockholm	1.89%	0.01%	-0.99%	-1.68%
Uppsala	1.77	-0.11	-1.11	-1.72
Malmö	2.46	0.54	-0.67	-1.32
Gothenburg	1.54	-0.34	-1.27	-1.95
B. Interval = 8 Quarters				
Stockholm	1.29	0.29	-0.57	-1.18
Uppsala	1.32	0.30	-0.59	-1.18
Malmö	1.50	0.48	-0.46	-1.05
Gothenburg	1.15	0.15	-0.65	-1.26
C. Interval = 12 Quarters				
Stockholm	1.84	1.07	-0.03	-0.59
Uppsala	1.35	0.62	-0.27	-0.86
Malmö	1.04	0.33	-0.48	-1.06
Gothenburg	1.20	0.47	-0.39	-0.98

differs sharply from the realized real returns in Table 7. After accounting for the costs associated with investment in housing, excess returns are not as large as real returns. The table shows that both the speculator and the investor lose money when buying and selling housing over these three time horizons. The market timer, as in Case and Shiller (1989), is able to make consistent excess returns by waiting for a positive forecast.

The most interesting result given in Table 8 is the returns experienced by the "mover." This type of buyer is the most common; the majority of housing transactions are not first time purchases, but reflect adjustments to housing consumption. Most of these transactions do not trigger capital gains taxes because the gains are reinvested in more expensive housing.

For these individuals, the excess returns are close to zero over the shortest interval, and increase slightly as the interval increases. While this appears to be evidence that excess returns are available to typical market participants, it should be noted that the costs of moving have not been included because no data were available. It seems likely that the true excess returns will be close to zero.

These results should hearten believers in competitive markets. Housing markets are inefficient but do not offer arbitrage opportunities. For the majority of the participants housing markets, the predictability of excess returns cannot be exploited. Furthermore, it appears that market forces have driven the returns to housing to exactly the home owner's cost of capital after accounting for the fixed cost associated with buying and selling housing.

6 Conclusion

Past research on housing market efficiency has focused on the behavior of aggregate housing prices. The general consensus from the literature is that aggregate housing prices and average returns are predictable, which is a violation of market efficiency. However, it is typically argued that the high costs associated with housing transactions prevents market forces from competing away what appear to be excess profits.

The model presented in this paper addresses the question of housing market efficiency at the level of the individual dwelling. By incorporating an appropriate error structure, which reflect the prominent features of housing markets, into a general model of housing prices, individual dwelling prices are shown quite clearly to follow a mean reverting process. This result indicates that, at this level, housing markets are inefficient.

Case and Shiller's (1989) speculation that "the house-specific component of the change in log price is probably not homoskedastic but that the variance of the noise increases with the interval between sales," is correct. The results of this paper indicate that the error variance is positively correlated with the interval of time between sales. However, the results also indicate the the correlation is *not one*. Rather than following a random walk, the errors are characterized more accurately as an autocorrelated process with a quarterly correlation coefficient of about 0.9–pricing errors will persist for many years.

These results have implications for aggregate prices, but unfortunately do not offer an opportunity for great wealth. The paper demonstrates that the serial correlation that characterizes individual home prices is found also in aggregate prices. The results indicate that for market timers–those for whom entry costs are sunk and capital gains taxes are mostly irrelevant–there exists at least one strategy that earns excess returns. However, the institutional features of housing markets for owner-occupied housing prevent investors from exploiting the apparent inefficiency to earn consistent excess profits. Indeed, it appears that excess returns are predictable, but that they have been driven down to approximately the fixed costs of buying and selling housing.

B Appendix: Derivation of the Covariance Matrix

The derivation of the general form of the covariance matrix, (11) in the text, is straightforward, but tedious. Consider

$$Var[X \pm Y] = Var[X] + Var[Y] \pm 2Var[X,Y] \text{ or } Var[X,Y] = \frac{1}{2}\left(Var[X+Y] - Var[X] - Var[Y]\right)$$

Substituting $\xi_{it} - \xi_{i\tau}$ for X and $\xi_{jg} - \xi_{j\gamma}$ for Y, we get

$$Var[(\xi_{it} - \xi_{i\tau}, \xi_{jg} - \xi_{j\gamma}] = \frac{1}{2}\left(Var[\xi_{it} - \xi_{i\tau} + \xi_{jg} - \xi_{j\gamma}] - Var[\xi_{it} - \xi_{i\tau}] - Var[\xi_{jg} - \xi_{j\gamma}]\right).$$

By assumption, the covariance is 0 across units ($i \neq j$), and the stochastic terms have mean zero. Therefore, only within-unit covariation is considered—the unit subscript is dropped below. Note that the covariance matrix is block diagonal under the assumption that the errors are neither spatially nor temporally correlated across dwellings. This may be a strong assumption, but routinely made. See Goetzmann and Spiegel (1995), for a discussion of this assumption. The general covariation equation is solved in parts, using the elements of the covariance matrix, equations (5) and (6) developed above.

$$\begin{aligned}
Var[\xi_t - \xi_\tau + \xi_g - \xi_\gamma] &= E[\xi_t] + E[\xi_\tau] + E[\xi_g] + E[\xi_\gamma] + \\
&\quad 2\left(-E[\xi_t\xi_\tau] + E[\xi_t\xi_g] - E[\xi_t\xi_\gamma] - E[\xi_\tau\xi_g] + E[\xi_\tau\xi_\gamma] - E[\xi_g\xi_\gamma]\right) \\
&= 4\left(\frac{\sigma_\mu^2}{1-\lambda^2} + \sigma_\eta^2\right) + 2\left(\frac{\sigma_\mu^2}{1-\lambda^2}\right)\Big(-\lambda^{t-\tau} + \lambda^{t-g} - \lambda^{t-\gamma} - \lambda^{\tau-g} + \lambda^{\tau-\gamma} - \\
&\quad + 2\,\sigma_\eta^2\Big(-I_{t\tau} + I_{tg} - I_{t\gamma} - I_{\tau g} + I_{\tau\gamma} - I_{g\gamma}\Big)
\end{aligned}$$

$$Var[\xi_t - \xi_\tau] = E[\xi_t^2] + E[\xi_\tau^2] - 2E[\xi_t\xi_\tau] = \frac{2\,\sigma_\mu^2}{1-\lambda^2} + 2\sigma_\eta^2 - \frac{2\,\sigma_\mu^2\,\lambda^{t-\tau}}{1-\lambda^2}$$

$$Var[\xi_g - \xi_\gamma] = E[\xi_g^2] + E[\xi_\gamma^2] - 2E[\xi_g\xi_\gamma] = \frac{2\,\sigma_\mu^2}{1-\lambda^2} + 2\sigma_\eta^2 - \frac{2\,\sigma_\mu^2\,\lambda^{g-\gamma}}{1-\lambda^2}$$

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.

Case, K. E., and R. J. Shiller (1987): "Prices of Single-Family Homes since 1970: New Indexes for Four Cities," *New England Economic Review*, 0(0), 45–56.

——— (1989): "The Efficiency of the Market for Single-Family Homes," *American Economic Review*, 79(1), 125–37.

——— (1990): "Forecasting Prices and Excess Returns in the Housing Market," *American Real Estate and Urban Economics Association Journal*, 18(3), 253–73.

Englund, P., J. M. Quigley, and C. L. Redfearn (1998): "Improved Price Indexes for Real Estate: Measuring the Course of Swedish Housing Prices," *Journal of Urban Economics*, 44(2), 171–96.

——— (1999a): "Do Housing Transactions Provide Misleading Evidence About the Course of Housing Values?," *unpublished manuscript*, pp. 1–26.

——— (1999b): "The Choice of Methodology for Computing Price Indexes: Comparisons of Temporal Aggregation and Sample Definition," *Journal of Real Estate Finance and Economics*, 19(3), 91–112.

Gatzlaff, D. H. (1994): "Excess Returns, Inflation, and the Efficiency of the Housing Market," *Journal of the American Real Estate and Urban EconomicsAssociation*, 22(4), 553–81.

Gatzlaff, D. H., and D. R. Haurin (1997): "Sample Selection Bias and Repeat-Sales Index Estimates," *Journal of Real Estate Finance and Economics*, 14(1-2), 33–50.

——— (1998): "Sample Selection and Biases in Local House Value Indices," *Journal of Urban Economics*, 43(2), 199–222.

Goetzmann, W. N., and M. Spiegel (1995): "Non-temporal Components of Residential Real Estate Appreciation," *Review of Economics and Statistics*, 77(1), 199–206.

Grossman, S. J., and J. E. Stiglitz (1980): "On the Impossibility of Informationally Efficient Markets," *American Economic Review*, 70(3), 393–408.

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.

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.

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.