INTERNATIONAL REAL ESTATE REVIEW

2011 Vol. 14 No. 2: pp. 208 – 239

Estimating Estate-Specific Price-to-Rent Ratios in Shanghai and Shenzhen: A Bayesian Approach

Jie Chen
HSBC (The Hong Kong and Shanghai Banking Corporation)

Shawn Ni *
University of Missouri; Department of Economics, University of Missouri, Columbia, MO65211, U.S.A.; Office Phone: 573-882-6878; Fax: 573-882-2697; Email: nix@missouri.edu.

The price-to-rent ratio, a common yardstick for the value of housing, is difficult to estimate when rental properties are poor substitutes of owner-occupied homes. In this study, we estimate price-to-rent ratios of residential properties in two major cities in China, where urban high-rises (estates) comprise both rental and owner-occupied units. We conduct Bayesian inference on estate-specific parameters by using information of rental units to elicit priors of the unobserved rents of units sold in the same estate. We find that the price-to-rent ratios tend to be higher for low-end properties. We discuss economic explanations for the phenomenon and the policy implications.

Keywords

Housing price; Rents; Heterogeneity; Bayesian analysis

* Corresponding author

1. Introduction

Unlike most investment or consumption goods, housing units differ in attributes that are difficult to measure. Establishing the relationship between housing price and the service flow of housing is an empirical challenge because the transaction price and rent of the same property are rarely simultaneously observed and rental properties are usually qualitatively different from owner-occupied housing. This study is an empirical analysis on a cross-sectional comparison of price-to-rent ratios by taking advantage of housing market data of two major Chinese cities.

In China, an urban complex of high-rises (we call estate) typically comprises rental and owner-occupied units of a similar quality. Using about ten thousand observations on price and over seven thousand observations on rental transactions in Shanghai and Shenzhen, we have developed a Bayesian approach to estimate the relationship between prices of housing units and their unobserved rents. Our analysis on latent rents of owner-occupied units sheds new light on the pricing of housing, and complements existing approaches on time series models of housing prices, the present value model, hedonic pricing models, and pricing based on repeated sales.

A voluminous amount of literature has been devoted to the time series behavior of price-to-rent ratio (e.g., Ayuso & Restoy (2006), McCue & Kling (2006), Dokko et al. (1991).) In particular, Ayuso & Restoy (2006) estimate a VAR that contains growth rates of aggregate price and rent, along with other macro-variables. Their model is useful for examining how the time series of aggregate housing price and rent relate to macroeconomic variables, but not designed for addressing the cross-sectional pattern of price-rent ratio. For a survey of the literature on time series analyses of housing price, see Himmelberg & Sinai (2005). A number of authors (e.g., Mankiw & Weil (1989), Clark (1995), Meese & Wallace (1994), and Clayton (1996)) find that the present value of aggregate housing price indexes in North America is sensitive to the rent index data. The researchers also recognize the difficulty in measuring rent index because rental properties may differ in quality from owner-occupied homes.[1] As an alternative to the rent-based pricing model, a hedonic pricing method is often used to estimate the market value of real estate features, such as lot size, number and size of rooms, number of

[1] There are a number of novel approaches to imputing rent index. Meese & Wallace (1994) use generated hedonic price indices and two different rent indices. One is the rent component of the U.S. Consumer Price Index, and the other index is a rent series based on “asking rents for two-bedroom condos pulled from local newspapers”. Clark (1995) uses neighborhood data instead of city data to minimize measurement errors. Clayton (1996) proxies imputed rents with a function of housing market fundamentals such as net immigration of households and the inventory of newly completed but unsold homes.

bathrooms, house age, and environmental characteristics.[2] A limitation of the hedonic pricing approach is that it does not directly link price to the latent rent of the property and does not reflect unobservable quality differences in the hedonic features for different estates. Case & Shiller (1989) focus on price changes of repeated home sales to circumvent the need to estimate the effect of unobserved heterogeneity on real estate pricing. This approach is effective in capturing the time series of housing price, but does not reveal cross-sectional differences in housing values. Our research complements these approaches by combining hedonic features with rent and price to study how the price of a housing unit relates to its latent rent.

Our empirical model consists of price and rent equations. In the rent equation, the property fixed effects on rent are estimated from data of rental units. The price equation links the price of a unit sold to its latent rent and other factors. A prominent feature of our rent-based pricing model lies in its cross-estate heterogeneity, which requires inference for thousands of parameters. Because of the limited observations of the price and rent data (about 27 price and 19 rent observations on average for each estate), asymptotic distribution can be misleading for our problem. In this study, we conduct finite sample inference through a Bayesian approach, which produces probability distribution of parameter conditioning on the observed sample. The Bayesian approach is particularly effective for models with heterogeneity where the number of parameters is large relatively to the number of observations. In a Bayesian framework, we can use housing market information to elicit priors of parameters for each estate to produce sharper inference. The Bayesian approach is often used to draw finite sample inference of latent parameters and missing data. We use the information obtained from the rental units to elicit priors of the latent rent of a unit sold. The posterior of the latent rent is jointly simulated with other parameters in the model. We use Bayes factors as the criterion to select from competing models. We find strong evidence that support heterogeneous (estate-specific) models of price and rent.

Recently, an innovative study by Hui et al. (2010) applies a Bayesian method to pricing of Hong Kong's real estate market. Similar to the markets of Shanghai and Shenzhen, the Hong Kong residential market mainly consists of high-rise estates with multiple units of similar features in each estate. Hui et al. (2010) develop a Bayesian hierarchical model that makes use of unit information in the same estate for efficient estimation of pricing factors of each unit. Hui et al. (2010) show the Bayesian approach is effective in drawing inference of a large number of parameters, as we also do in this study. However, our study has a different research objective from that by Hui et al.

[2] Some examples of hedonic pricing studies on data of various countries include Goodman (1988), Can (1990), Downes & Zabel (2002), Vanderford et al. (2005), Bao & Wan (2004), Mats (2002), Witte et al.(1979), Ridker & Henning (1967), Kim (1992)), Schwartz et al. (2003), Hughes & Sirmans (1992), Basu & Thibodeau (1998), and Coulson & Bond (1990), among others.

(2010). They focus on advancing the literature on hedonic pricing by only using the transaction price data; we focus on a cross-sectional comparison of the price-rent ratios by using both the transaction price and rental data.

We address two questions that concern the cross-sectional distribution of the price of housing units relative to their fundamentals. First, “How large is the dispersion in the price-to-rent ratio across housing units?” We find that the cross-estate standard deviation of estimated price-to-rent ratio is substantial: about forty four percent for Shanghai and twenty five percent for Shenzhen. The second empirical question is “How is the price-to-rent ratio correlated with features of the property?” We find that the price-to-rent ratio is higher for low-end housing. The cross estate correlation of estate-fixed effects of price-to-rent ratio and that of rent is -0.88 for Shanghai and -0.59 for Shenzhen. Economic explanations for the high price-to-rent ratio of low-end housing considered in the paper include better growth prospects of estates in newly developed locations and stronger demand for low-end housing due to a variety of housing market features and government subsidies.

The rest of the paper is organized as follows. Section 2 lays out the models for latent rent and the pricing models. It also includes a Markov chain Monte Carlo (MCMC) algorithm for posterior simulation. Section 3 presents empirical results on Bayesian model selection and estimation. Section 4 offers economic explanations to the phenomenon that low-end estates tend to carry a high price-to-rent ratio. Section 5 concludes.

2. A Bayesian Framework of Real Estate Pricing

2.1 Pricing Real Estate Based on Latent Rent

For estate i (i=1,…,I) apartment unit sold j, (j=1,…,J_i), the price equation is

$$P_{ij} = c_i + \alpha_i \hat{R}_{ij} + y_{ij}^{'} \beta + \varepsilon_{ij} \tag{1}$$

where P_{ij} is the logarithm of the price of apartment j in estate i, $\hat{R}_{ij}$ is the logarithm of the latent rent of unit j in estate i, $y_{ij}^{'}$ is the row-vector of factors that influence real estate pricing, for instance, macroeconomic variables, such as mortgage rates. β is the column vector of unknown parameters. The pricing error ε_{ij} is assumed to be normal N (0, σ_i^2).

The rent equation is given by:

$$R_{ik} = \mu_i + x_{ik}^{'} \theta + v_{ik} \tag{2}$$

where R_{ik} is the logarithm of the observed rent of rental unit k (k =1,…,K_i) in estate i (i=1, …, I), and $x_{ik}^{'}$ is the row-vector of seasonal and unit-specific factors that influence rent (e.g. the size of the unit, number of bedrooms and bathrooms, condition of the unit, etc.). We assume that the error term v_{ik} is normal $N(0, \tau_i^2)$.

The price and rent equations contain three types of heterogeneity in regression coefficients: μ_i, c_i and α_i. The estate-fixed effect μ_i captures the location, environmental characteristics, the quality of property management service, and other intangibles beyond the hedonic factors included in the rent equation. The parameter c_i reflects the estate-fixed effect on price conditioning on the latent rent and other factors. The parameter α_i concerns heterogeneity in the price elasticity with respect to the latent rent. There are eight combinations of model restrictions to be compared. For instance, a heterogeneous pricing model without an estate-fixed effect for rent is defined by $\mu_i=\mu$ while c_i and α_i are estate-specific for (i=1,…,I). In the most restrictive homogenous pricing model, the estate-fixed effect is absent and all parameters are constant across estates. We will show that the empirical result strongly favors the presence of estate-specific fixed effects.

We noted earlier that we observe rental and sale prices of units in the same estate, but few units are sold and rented at the same time. We need to develop methodologies for estimation of the latent rents of units sold by using the information from rent data in the same estate, for which we conduct a Bayesian analysis.

We assume that the unobserved rent of unit j sold in estate i follows the same distribution as rental units given in Equation (2):

$$\hat{R}_{ij} = \mu_i + \hat{x}_{ij}'\theta + v_{ij} \tag{3}$$

An alternative way to view the assumption is that we set the prior for the unobserved rent of a unit sold conditional on its hedonic features and model parameters as $\hat{R}_{ij} \sim N(\mu_i + \hat{x}_{ij}'\theta, \tau_i^2)$. The price equation (1) depicts the likelihood function of parameters, including the unobserved rent. Vector $\hat{x}_{ij}$ captures observable hedonic features of the unit j sold in estate i. The posterior of the latent rent $\hat{R}_{ij}$ will be simulated along with other parameters.

A common frequentist solution to unobservable regressors is the instrumental variables (IV) approach. By the IV approach, we can treat the rent equation as the first-stage regression and the pricing equation as the second stage regression, based on the instruments of the observed rents of units rented in the same estate and hedonic factors of the units sold. In contrast to the IV approach, the Bayesian approach conducts joint inference for the parameters in the rent and price equations. The information in the price data contributes to the posterior of the imputed rent. Zellner (1970) argues that a Bayesian analysis for regression models is preferable when the sample size is small. Although some advances have been made in finite sample hypothesis testing in error-in-variable regressions with homogenous parameters (e.g., Dufour &

Jasiak (2001)), the frequentist finite sample distributions of heterogeneous parameters are complicated. The Bayesian approach is convenient for dealing with heterogeneous parameters and measurement errors. Zellner (1970) uses non-informative priors for Bayesian analyses in a regression model with unobservable independent variables. In the present study, we follow the strategy taken by Rossi & Allenby (1993) and elicit informative priors for a large number of estate-specific parameters by using information borrowed from data of the whole sample. This approach is designed for sharper inference of estate-specific parameters.

2.2 Prior Setting

We assume the following prior setting: $c_i \sim N(\bar{c}, \bar{v}_c^2)$, $\alpha_i \sim N(\bar{\alpha}, \bar{v}_\alpha^2)$, $\beta \sim N(\bar{\beta}, \mathrm{B})$, $\sigma_i^2 \sim IG(s_\sigma, v_\sigma)$, $\mu_i \sim N(\bar{\mu}, \bar{v}_\mu^2)$, $\theta \sim N(\bar{\theta}, \Theta)$, $\tau_i^2 \sim IG(s_\tau, v_\tau)$. Numerous books, for example Dey & Rao (2005), Geweke (2005), and O'Hagan & West (2010), address the general issues of prior elicitation. The problem we have at hand is how to set priors for estate-specific parameters in a regression model by using the information in all estates in the sample. The specific approach of prior selection is closely related to the study by Rossi & Allenby (1993) which estimates consumer-specific parameters in a regression model of marketing research, where priors for each consumer are set based on the sample information. The prior setting for parameters in a heterogeneous parameter model is as follows. Following Rossi & Allenby (1993), we elicit priors for estate-specific parameters by using the posterior of the homogenous (constant) parameter model (where no parameter is estate specific) under flat prior. Normal prior means of parameter (c_i, α_i, μ_i) are set as their posterior means in the constant parameter model. Posterior variances are set as the number of estates (287 for Shanghai and 81 for Shenzhen) times the posterior variances of the corresponding parameters in the homogenous model. Inverse gamma priors for σ_i^2 and τ_i^2 are set as *IG* (3, 0.3) and *IG* (3, 0.2), which have means of 0.15 and 0.1 and standard deviations of 0.15 and 0.1. We consider alternative sets of hyperparameters and find that the reported empirical results are robust to the prior setting.

2.3 Decomposing the Price-to-Rent Ratio by Factors

To explore the determinants of the estimated price-to-rent ratio, we decompose the price-to-rent ratio into estate-specific, unit-specific, and macroeconomic factors.

Taking the difference of the logarithm of price in (1) and logarithm of imputed rent in (3), we have the logarithm of the price-to-rent ratio:

$$P_{ij} - \hat{R}_{ij} = \left[c_i + (\alpha_i - 1)\mu_i\right] + (\alpha_i - 1)x_{ij}'\theta + y_{ij}'\beta + \left(\varepsilon_{ij} + (\alpha_i - 1)v_{ij}\right). \quad (4)$$

Note that since the latent rent is unobserved, so is the price-to-rent ratio. In (4) the price-to-rent ratio is explicitly decomposed. The term, $c_i + (\alpha_i - 1)\mu_i = f_i$, is the estate-specific factor. Empirically, this turns out to be the dominating factor of the price-rent ratio. The second term, $(\alpha_i - 1)x'_{ij}\theta = u_{ij}$, is a unit-specific factor that captures the effect of hedonic features of the unit on the price-to-rent ratio.

One important factor of housing price is the value of land user right. In China, land is state owned, but land user right is tradable and transferable. According to the government regulation, owners of commercialized residential housing have a limited time of land use (usually 70 years from the date of the initial commercialization of the real estate or the initial development of the land). When the lease on the land expires, its owner-occupier is expected to either extend the lease by paying a renewal fee, or revert the housing to the state. The effect of the land use policy on housing price is largely unexplored empirically. The Shenzhen data contain the year each estate became commercialized (from which we can calculate the remaining years of the user right). The land user effect is measured by $e_{ij} = \delta\, l_{ij}$, where l_{ij} is the remaining years of the land user right at the time unit j in estate i is sold. $y'_{ij}\beta$ is the sum of the macroeconomic and seasonal factors and the user right factor. The seasonal and macroeconomic effect is measured by part of the unit-specific effect $m_{ij} = y'_{ij}\beta - e_{ij}$.

The remaining portion of the price-to-rent ratio, $\xi_{ij} = \varepsilon_{ij} + (\alpha_i - 1)v_{ij}$, is a pricing error that can not be attributed to the factors mentioned above. With this notion, from (4), the price-rent ratio $\lambda_{ij} = P_{ij} - \hat{R}_{ij}$ can be written as:

$$\lambda_{ij} = f_i + \mu_{ij} + m_{ij} + e_{ij} + \xi_{ij}. \tag{5}$$

In the following, we will discuss how to simulate numerical distributions of the posterior of model parameters. The posteriors of the price-to-rent ratio and its components, $(\lambda_{ij}, f_i, \mu_{ij}, m_{ij}, e_{ij}, \xi_{ij})$, can be computed from the simulated posteriors of model parameters.

2.4 Posterior Simulation

Let

$$P_i = \begin{pmatrix} P_{i,1} \\ \vdots \\ P_{i,J_i} \end{pmatrix}, \hat{R}_i = \begin{pmatrix} \hat{R}_{i,1} \\ \vdots \\ \hat{R}_{i,J_i} \end{pmatrix}, R_i = \begin{pmatrix} R_{i,1} \\ \vdots \\ R_{i,K_i} \end{pmatrix}, y_i = \begin{pmatrix} y'_{i,1} \\ \vdots \\ y'_{i,Ji} \end{pmatrix}, x_i = \begin{pmatrix} x'_{i,1} \\ \vdots \\ x'_{i,K_i} \end{pmatrix}, \hat{x}_i = \begin{pmatrix} \hat{x}'_{i,1} \\ \vdots \\ \hat{x}'_{i,Ji} \end{pmatrix}. \tag{6}$$

The variables with 'hat' $(\hat{R}_i, \hat{x}_i)$ pertain to the units sold, while those without 'hat' pertain to rental units in estate i. Denote data $D=\{P_i, R_i, y_i, x_i\}$, $(i=1,\ldots N)$, and the collection of μ_i, c_i and α_i, for all i=1,…I by {μ, c, α}. The joint posterior is:

$$\pi(\alpha,\beta,\mu,\theta,\sigma,\tau,c|D)$$
$$\propto L(P|\hat{R},\alpha,c,\beta,\sigma)\pi(\hat{R}|\mu,\theta,\tau,\hat{x})L(R|\theta,\mu,\tau)\pi(\beta)\pi(\theta)\pi(\sigma)\pi(\tau)\pi(\alpha)\pi(\mu)\pi(c)$$

Unlike the conventional regression model, the presence of the endogenous latent rent $\hat{R}_{ij}$ in the price equation implies that the marginal posteriors of the regression coefficients are not standard distributions. In a Bayesian analysis, when the posterior does not follow a standard distribution, researchers usually use numerical draws from the posterior to compute quantities of interest. A common approach to numerical draws is MCMC Gibbs sampling, based on the posterior of each parameter conditional on the data and other parameters in the model.

In the appendix, we present the conditional posteriors and a Gibbs sampling algorithm. We focus on the simulation of the posterior for Model 1, the most general model in Table 3. The algorithms for more restricted models, such as Models 2 to 8, are similar and omitted (we only need to replace vector parameters c_i or α_i for i=1,…N by scalar parameters c or α.)

Using the simulated posteriors, eight combinations of specifications on (μ, c, α) will be compared through a Bayesian model selection.

2.5 Bayesian Model Selection

We compare competing models on the basis of posterior probability of the model given the data **D**, *pr* ($M|D$). *pr* ($M|D$) is the product of prior probability *pr*(M) and the marginal likelihood of the data given the model. The marginal likelihood is obtained by integrating out model parameters in the posterior. The choice between two competing models M_1 and M_2 depends on the ratio $\frac{pr(M_1|D)}{pr(M_2|D)}=\frac{pr(M_1)}{pr(M_2)}\times\frac{pr(D|M_1)}{pr(D|M_2)}$, that is, posterior odds = prior odds × Bayes factor. When the prior probabilities of the competing models are equal ($pr(M_1) = pr(M_2) = 0.5$), the Bayes factor $B_{12}=\frac{pr(D|M_1)}{pr(D|M_2)}$ greater than unity suggests that the data are in favor of Model 1 over Model 2. The strength of the evidence is given by the size of the Bayes factor. Kass & Raftery (1995) suggest a guideline in interpreting Bayes factors: if the natural logarithm of the Bayes factor is between 1 and 3, the evidence is “positive”; between 3 and 5, the evidence is “strong”; and above 5, the evidence is “very strong.” In each model, the posterior is simulated by 10000 MCMC cycle after 10000 “burn in” runs. We report the model selection result in Table 3 and conduct an empirical analysis of the price-rent ratio based on the selected model. We calculate the marginal likelihood of each model by analytically integrating parameters whenever possible and then using the harmonic mean for numerical approximation of the remaining portion of the integration.

3. Empirical Results

3.1 Data

The data of two major cities in China, Shanghai and Shenzhen, from January 2003 to December 2005, were provided by Centaline Property Consultants Ltd. China. The data for Shanghai include 5258 rental transactions (each with an observation of monthly rent) and 7740 resale transactions of 287 estates. The data for Shenzhen consist of 1760 rental transactions and 2179 resale transactions of 81 estates. These transactions are made on the secondary market so that there are considerable variations in the time of transactions across apartment units within each estate. Figures 1 to 4 plot the histograms of the estates in Shanghai and Shenzhen, given the number of observations of price and rent.

In this study, transaction price and monthly rent are in RMB (in constant 2003 Yuan) per square meter. Tables 1 and 2 report the annual averages of the price and rent. The Shanghai data include the following hedonic features: the floor level of the apartment, number of bedrooms, number of living rooms, number of bathrooms, and the size of the apartment. In addition, measurements of the quality of the interior decoration and furniture, kitchen cabinets, and accessories, are grouped into three categories: none, simple, or luxurious. The Shenzhen data are more limited; the size of the apartment is the only observable feature for all units. For both markets, in the rent equation, we control for month and year dummies. Besides the latent rent, control variables in the price equation include mortgage rate at the time of purchase for Shanghai, and mortgage rate and the remaining years of the land user right for Shenzhen.

3.2 Results of Model Selection

Table 3 lists the log marginal likelihoods of eight competing models. For both the residential housing markets of Shenzhen and Shanghai, Bayes factors favor Model 4 over the other models. The constant parameter Model 8 is decisively rejected by the data. This suggests that estate specific heterogeneity plays an important role in explaining the price-to-rent ratio. The empirical results for the rest of the paper pertain to Model 4 (α_i=α for all estates).

Figure 1 Histogram of Estates Given the Number of Sale Transactions: Shanghai

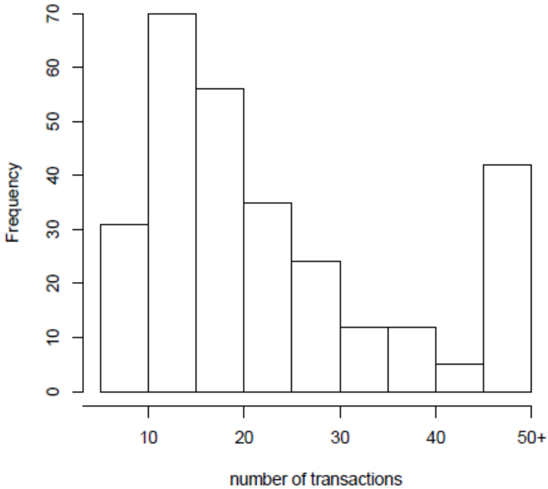


Figure 2 Histogram of Estates Given the Number of Rental Transactions: Shanghai

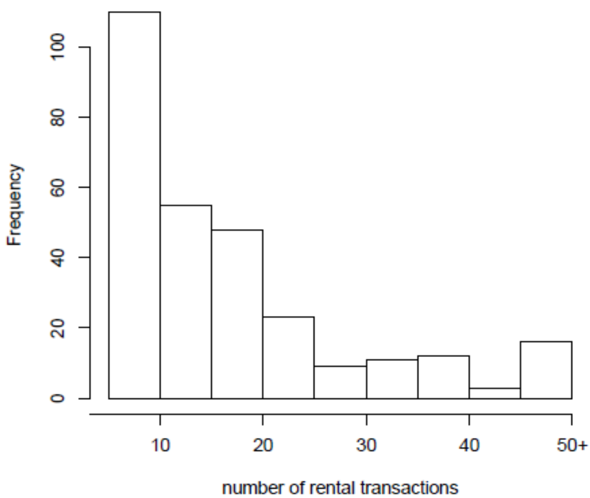


Figure 3 Histogram of Estates Given the Number of Sale Transactions: Shenzhen

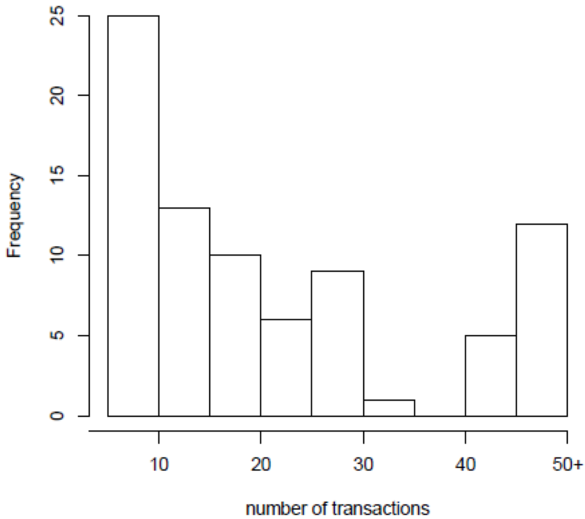


Figure 4 Histogram of Estates Given the Number of Rental Transactions: Shenzhen

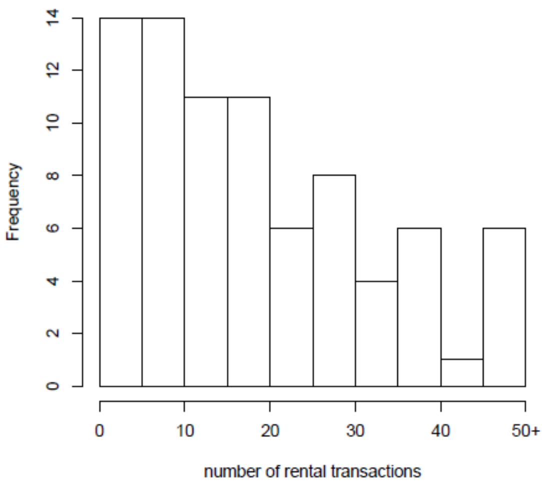


Table 1 Summary Statistics of the Transaction Price/Rent of Shanghai

Year	Average Sale Price (RMB/Square Meter)	Average Rent (RMB/Square Meter per Month)
2003	8919.89 (2980.22)	50.25(21.57)
2004	11643.87(4277.54)	48.23(26.71)
2005	13109.27(5586.36)	39.63(25.45)

Note: Numbers in parenthesis are standard deviations of sale prices/rents. The price and rent are in constant RMB (with 2003 as the base year).

Table 2 Summary Statistics of the Transaction Price/Rents of Shenzhen

Year	Average Sale Price (RMB/Square Meter)	Average Rent (RMB/Square Meter per Month)
2003	5454.59 (2145.24)	39.34 (19.31)
2004	5674.19 (3915.85)	33.05 (17.68)
2005	7569.32 (5301.91)	37.82 (17.12)

Note: Numbers in parenthesis are standard deviations of sale prices/rents. The price and rent are in constant RMB (with 2003 as the base year).

Table 3 Bayesian Model Selection, Log (Marginal Likelihoods) of Shanghai and Shenzhen Data

	c_i	μ_i	α_i	Log-m.l (Shanghai)	Log-m.l (Shenzhen)
Model 1	estate-specific	estate-specific	estate-specific	-956.19	41.93
Model 2	constant	estate-specific	estate-specific	-835.93	158.86
Model 3	estate-specific	constant	estate-specific	-5405.79	-1746.9
Model 4	estate-specific	estate-specific	constant	**853.4**	**349.13**
Model 5	constant	constant	estate-specific	-5002.47	-1594.97
Model 6	constant	estate-specific	constant	-541.63	-80.5
Model 7	estate-specific	constant	constant	-1746.9	-1621.14
Model 8	constant	constant	constant	-5914.02	-2288.8

Note: A parameter is labeled 'estate-specific' ('constant') when it is assumed to be different (the same) across estates. The marginal likelihood of Model *i* (*i* =1,…, 8) is computed by integrating out all parameters in the posterior under Model *i*. A difference of 5 between the log marginal likelihoods of two models is considered strong evidence in favor of the model with the larger marginal likelihood. By this criterion, Model 4 is the best model.

3.3 Posterior Properties: Latent Rent

Tables 4 and 5 report the cross-estate averages of the posterior mean and standard deviations of parameters in Model 4. Posterior mean, our benchmark estimator, is the Bayesian estimator under quadratic loss. The posterior mean of estate specific intercept μ_i ranges from 2.5 to 5.5, which indicates a substantial quality difference in the neighborhood characteristics, environmental features, and accessibility of various estates.

Structural attributes identified in previous studies of hedonic pricing, such as the floor level and number of bedrooms, are available in Shanghai. The number of floor levels of an apartment has an ambiguous effect on price. This may be because for apartments in a high-rise housing complex, a higher floor level means better air quality and view. On the other hand, because most of the low-rise apartment buildings are not equipped with elevators, a high floor level may be a negative factor. Tables 4 and 5 also report the estimates of other parameters in the model.

Table 4 Bayesian Estimates Averaged over Estates, Shanghai

Parameter	Average of Posterior Mean	Average of Posterior STD
Price Model		
c_i	8.5018	0.0568
α **(constant)**	0.3114	0.0089
β_i: ***real mortgage rate***	-0.0966	0.0018
σ_i^2	0.0274	0.0096
Rent Model		
μ_i	3.7379	0.0866
θ		
floor level	0.0051	0.0005
number of bedroom	-0.106	0.0096
number of livingroom	-0.0751	0.011
number of bedroom	0.0064	0.0059
decoration: simple	0.0216	0.0081
decoration: luxury	0.1375	0.0091
size	-0.0001	0.0001
τ_i^2	0.0735	0.0232

Note: The results pertain to Model 4 (selected based on the Bayes factor.). The posterior of parameter of each estate is simulated using the MCMC algorithm stated in the paper. From the numerical distribution, we compute the posterior mean and posterior standard deviation for each parameter. The first column is obtained by averaging the posterior mean over all estates. The second column is obtained by averaging the posterior standard deviation over all estates. The estimates of seasonal and year dummies are not reported.

Table 5 Bayesian Estimates Averaged over Estates, Shenzhen

Parameter	Average of Posterior Mean	Average of Posterior STD
Price Model		
c_i	7.5209	0.1139
α **(constant)**	0.1718	0.0305
β_{1i}: ***real mortgage rate***	-0.048	0.0035
β_{2i}: ***land use right***	0.0109	0.0018
σ_i^2	0.0369	0.0131
Rent Model		
μ_i	3.5883	0.067
θ		
size	-0.0012	0.0001
τ_i^2	0.0665	0.0226

Note: See the note of Table 4.

The number of bedrooms is negatively correlated with the rent (per square meter) while the number of living rooms and restrooms is positively correlated with rent. In addition, we find that compared with an unfurnished apartment, a plainly furnished apartment does not raise the rent by much, while a luxuriously decorated one on average raises rent by about 14 percent. Real mortgage rate is found to be negatively correlated with housing prices. A one percent increase in real mortgage rate drives the housing price down by 9.7 percent in Shanghai and 4.8 percent in Shenzhen. The overall fit of the model for the two markets are comparable. The posterior means of the variance of the pricing error σ_i^2 are similar for Shanghai and Shenzhen.

Table 6 Decomposition of the Price-to-Rent Ratio, Shanghai

Parameter	Posterior Mean Average	Posterior Std Average
λ_{ij}	5.9056(0.4364)	0.2392(0.0532)
f_i	5.9499(0.3815)	0.0814(0.0278)
u_{ij}	0.2185(0.1251)	0.0335(0.0060)
m_{ij}	-0.2628(0.1260)	0.0050(0.0024)

Note: The posterior of parameter of each estate is simulated by using the MCMC algorithm stated in the paper. From the numerical distribution, we compute the posterior mean and posterior standard deviation for each term in Equation (5). The first column is obtained by averaging the posterior mean of these terms over all estates. The second column is obtained by averaging the posterior standard deviation of the terms over all estates. Numbers in parenthesis in the first (second) column are standard deviations of the posterior mean (posterior standard deviation) across all estates.

Table 7 Decomposition of the Price-to-Rent Ratio, Shenzhen

Parameter	Posterior Mean Average	Posterior Std Average
λ_{ij}	5.3195(0.2478)	0.2614(0.1078)
f_i	4.5992(0.1716)	0.1403(0.0208)
u_{ij}	0.0939(0.0549)	0.0195(0.0034)
m_{ij}	-0.0513(0.0585)	0.0039(0.0041)
e_{ij}	0.6775(0.0374)	0.1152(0.0064)

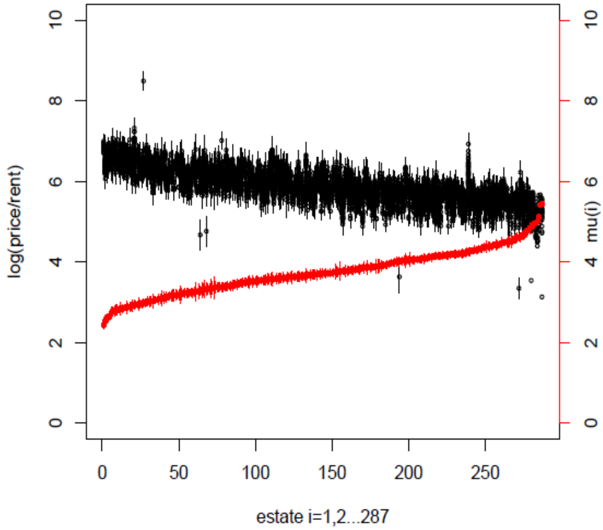
Note: See the note of Table 6.

Tables 6 and 7 report the sample averages and standard deviations of the posterior means and posterior standard deviations of the factors that influence the price-rent ratio. The tables illustrate the prominent role of the estate-fixed effect. The unit factor and mortgage rate play more significant roles in Shanghai than Shenzhen. The average estimate of 0.678 for land-user right (which is only observed for Shenzhen) accounts for roughly thirteen percent of the price-to-rent ratio for units sold in Shenzhen. The cross-estate variation in the land user right is small because the housing reform only started in the late 1990s and the transactions in the sample occurred within a short period of time. One may incorporate the land-user right into the estate fixed effect in the price-to-rent ratio f_i. The correlation between f_i and estate fixed effect in rent c_i is not significantly altered if we do so.

Figures 5 and 6 plot the posterior mean and the posterior 10 and 90 percentiles of the price-to-rent ratio of each unit sold and fixed effect μ_i of each estate of Shanghai and Shenzhen. The horizontal axis corresponds to estates of each city sorted by the posterior mean of μ_i. These figures exhibit two distinct features. First, the posterior distributions of the price-to-rent rent ratio and that of estate quality (μ_i) are quite tight, i.e., the estate specific fixed effects are estimated with high precision relative to the cross sectional difference. Second, the price-to-rent ratio of low-end housing is higher than that of high-end housing, especially for Shanghai.

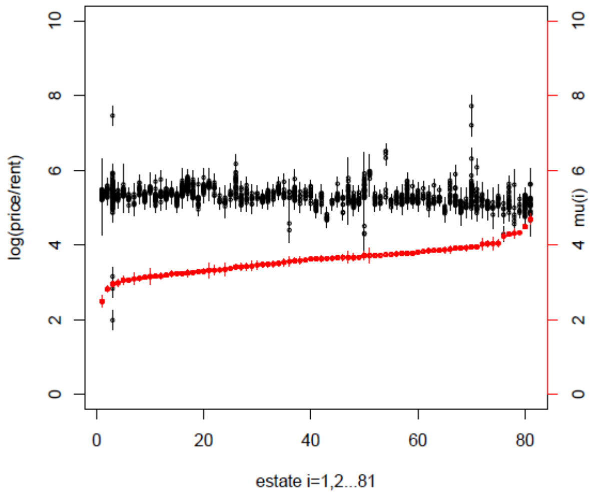
Figures 7 and 8 show a positive cross-estate correlation between the fixed effect in rent, μ_i, *with* the fixed effect in price, c_i. There is a substantial portion of cross-estate variation in the pricing factor c_i that is not explained by μ_i. Across estates, c_i is more strongly correlated with μ_i in the Shanghai market than the Shenzhen market. This is because the data of Shanghai include richer hedonic features. For the Shenzhen sample, most of the variation in rent is explained by the estate-fixed effect rather than unit-specific features. Consequently, the estimate of α (price elasticity with respect to rent) is larger for the Shanghai sample.

Figure 5 Price-to-Rent Ratio λ_{ij} and Estate Fixed Effect μ_i (Shanghai)



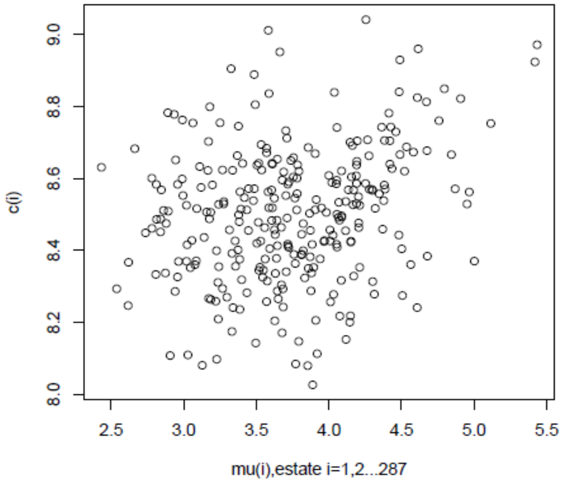
Note: Figure 5 (Shanghai) plots the posterior mean and the posterior 10 and 90 percentiles of the price-to-rent ratio of each unit sold λ_{ij} (in black) and those of the estate-fixed effect μ_i in rent (in red). The estates on the horizontal axis are sorted by the posterior mean of μ_i.

Figure 6 Price-to-Rent Ratio λ_{ij} and Estate Fixed Effect μ_i (Shenzhen)



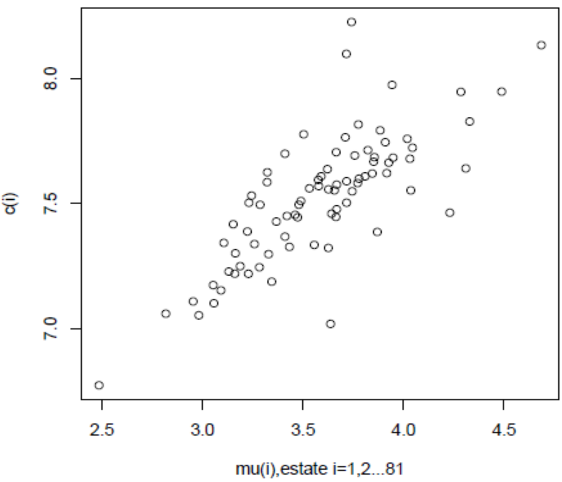
Note: Figure 6 (Shenzhen) plots the posterior mean and the posterior 10 and 90 percentiles of the price-to-rent ratio of each unit sold λ_{ij} (in black) and those of the estate-fixed effect μ_i in rent (in red). The estates on the horizontal axis are sorted by the posterior mean of μ_i.

Figure 7 Correlation between estate fixed effect in pricing c_i and estate fixed effect in rent μ_i (Shanghai)



Note: Figure 7 plots the posterior mean of c_i for estate i against μ_i. c_i appears in the price equation (1) and μ_i appears in the rent equation (2). The posterior of parameter of estate is simulated by using the MCMC algorithm stated in the paper. The cross-estate correlation between the posterior means of c_i and μ_i is 0.2549 for Shanghai.

Figure 8 Correlation between estate fixed effect in pricing c_i and estate fixed effect in rent μ_i (Shenzhen)



Note: Figure 8 plots the posterior mean of c_i for estate i against μ_i. c_i appears in price equation (1) and μ_i appears in the rent equation (2). The posterior of parameter of estate is simulated by using the MCMC algorithm stated in the paper. The cross-estate correlation between the posterior means of c_i and μ_i is 0.7685 for Shenzhen.

3.4 ‘Within’ Estates and ‘Between’ Estates Decomposition of Price-to-Rent Ratios

In the following, we will examine the ‘within’ estates and ‘between’ estate variations of the posterior means and posterior standard deviations of (λ_{ij}, f_i, μ_{ij}, m_{ij}, e_{ij}), which characterize the cross-unit and cross-estate distributions of the price-rent ratio and its determinants. Comparisons of the patterns of these variations shed new light on the housing markets of the two cities.

An essential difference between the cities of Shanghai and Shenzhen lies in their histories. Shanghai is a city that is several hundreds of years old, with well developed historic districts and recently developed suburban districts. Shenzhen, a major city with more than eight million residents, on the other hand, was a small fishing town three decades ago and developed when it was designated as a Special Economic Zone in 1979. One would expect larger dispersion in estate-fixed effect in Shanghai than Shenzhen.

Table 8 ‘Within’ and ‘Between’ Variations in the Posterior Means of Price-to-Rent Ratios, Shanghai

	Between-Estate Variation of Posterior Mean	Within-Estate Variation of Posterior Mean
λ_{ij}	0.1578	0.0332
f_i	0.146	NA
u_{ij}	0.0043	0.0113
m_{ij}	0.0036	0.0123

Notes: The tables report the ‘within estates’ and ‘between estates’ variations of the posterior mean of each term in Equation (5). Let z_{ij} be a generic notation for a quantity of interest of estate i ($i=1,\dots,N$) and unit sold j ($j = 1,\dots,J_i$). Denote the within estate average $z_{i.} = \frac{1}{J_i}\sum_j z_{ij}$ and the whole sample average $z_{.} = \frac{1}{N}\sum_i z_{i.}$. The ‘within’ and ‘between’ variations of quantity of z_{ij} are given by $\frac{1}{N}\sum_i \frac{1}{J_i}\sum_j \left(z_{ij} - z_{i.}\right)^2$ and $\frac{1}{N}\sum_i \left(z_{i.} - z_{.}\right)^2$. The posterior of parameter of each estate is simulated by using the MCMC algorithm stated in the paper.

Table 9 'Within' and 'Between' Variations in the Posterior Means of Price-to-Rent Ratios, Shenzhen

	Between-Estate Variation of Posterior Mean	Within-Estate Variation of Posterior Mean
λ_{ij}	0.0271	0.0362
f_i	0.032	NA
u_{ij}	0.0011	0.0019
m_{ij}	0.0007	0.0028
e_{ij}	0.0014	0.0001

Note: See note in Table 8.

Tables 8 and 9 show the cross-unit ('within'-estate) and 'between'-estate distributions of the point estimate of the price-rent ratio. Consistent with the history of development in the cities, the magnitude of cross-estate ('between') price-rent ratios and estate fixed factors are much larger for Shanghai than Shenzhen. In contrast, the cross-unit ('within') estate distributions are similar for Shanghai and Shenzhen. We also consider the price-rent ratio and its determinants from a different perspective, by examining the posterior means and standard deviations of 'within' and 'between' variations of (λ_{ij}, f_i, μ_{ij}, m_{ij}, e_{ij}), the point estimate and posterior uncertainty (precision) of cross-unit and cross-estate variations of the price-rent ratio and its determinants. These unreported estimates are similar to those in Tables 8 and 9.

4. Explaining the Negative Correlation between Rent and Price-to-Rent Ratio

The negative correlation between the estate-specific fixed effect in the price-rent ratio, f_i, and the fixed effect of rent, μ_i, (plotted in Figures 9 and 10) warrants further discussion. We consider three explanations in Subsections 4.1 to 4.3 for the negative correlation.

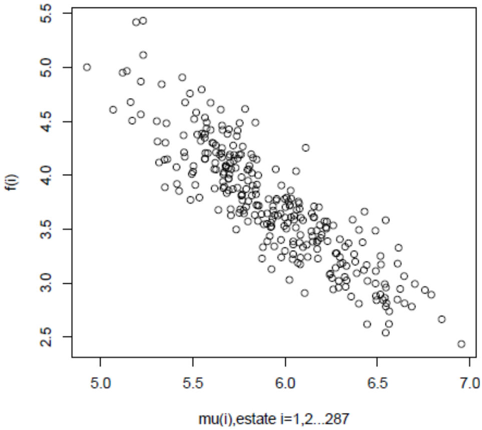
4.1 Estimation Errors in Rent Fixed Effect (μ_i)

First, note that the negative cross-estate correlation between $f_i = c_i + (\alpha_i - 1)\mu_i$ and μ_i may be partly due to the estimation errors in μ_i. Suppose the 'true' parameter is $f_i^* = c_i + (\alpha_i - 1)\mu_i^*$ and μ_i^*, and suppose $\mu_i = \mu_i^* + \upsilon_i$, where the error υ is uncorrelated with μ or c. Then, the cross-estate covariance is:

$$\begin{aligned}\operatorname{cov}(f_i, \mu_i) &= E\left(f_i - \bar{f}\right)\left(\mu_i - \bar{\mu}\right)\\ &= E\left\{[c_i + (\alpha - 1)\left(\mu_i^* + \upsilon_i\right) - \bar{c} - (\alpha - 1)\bar{\mu}]\left(\mu_i^* - \bar{\mu} + \upsilon_i\right)\right\}\\ &= E\left[\left(f_i^* - \bar{f}\right)\left(\mu_i^* - \bar{\mu}\right)\right] + (\alpha - 1)E\upsilon_i^2\end{aligned}$$

Since the estimate of α is smaller than unity, the covariance has a downward bias $(\alpha-1)E\upsilon_i^2$. We will show that for two reasons, the magnitude of this bias is small for the present problem.

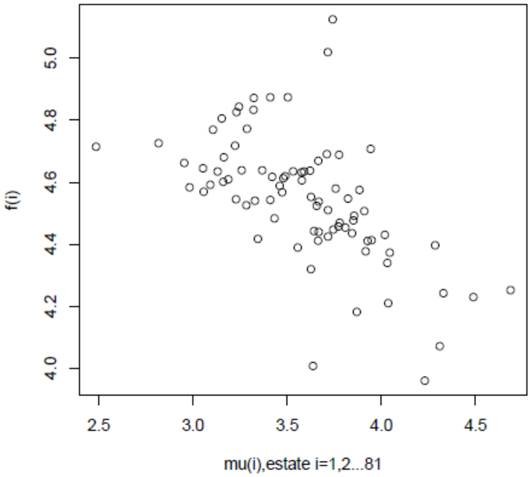
Figure 9 Correlation between Estate Fixed Effect in Price-to-Rent Ratio f_i and Estate Fixed Effect in Rent μ_i (Shanghai)



Note: Figure 9 plots the posterior mean of f_i for estate i against μ_i. f_i is the fixed effect of price-to-rent ratio in Equation (5) and μ_i appears in the rent equation (2). The posterior of estate-specific parameters is simulated by using the MCMC algorithm stated in the paper. The cross-estate correlation between the posterior means of f_i and μ_i is -0.8774 in Figure 9 (Shanghai data).

First, if we use the posterior variance of μ_i to proxy that of the estimation error υ_i, then based on the statistics reported in Tables 4 and 6, and the fact that the cross-estate standard deviation μ_i is 0.53 for the Shanghai data (not reported in the tables), $\frac{(\alpha-1)E\upsilon_i^2}{sd(f_i)sd(\mu_i)}$ is roughly $\frac{(0.31-1)\times 0.09^2}{0.38\times 0.53}\approx -0.03$, much lower than the sample correlation -0.88 (for Shanghai). For the Shenzhen data, the cross-estate standard deviation μ_i is 0.39. Given the numbers reported in Tables 5 and 7, the contribution from the estimation error in the estate-fixed effects of rent is approximately $\frac{(\alpha-1)E\upsilon_i^2}{sd(f_i)sd(\mu_i)}=\frac{(0.17-1)\times 0.07^2}{0.17\times 0.39}\approx -0.06$, negligible compared to -0.59, the negative correlation between the posterior mean of f_i and that of μ_i. Note that due to the lack of hedonic features of housing units in the Shenzhen data, unit-specific rent plays a smaller role in explaining the price in Shenzhen than Shanghai. In particular, the estimate of α (0.172) for the Shenzhen sample is half the magnitude of that of Shanghai (0.311). For a given estimation error in the estate-fixed effect μ_i, the downward bias in the cross-estate correlation between f_i and μ_i for Shenzhen should be larger than that for Shanghai. Yet the obtained sample correlation for Shenzhen (-0.59) is higher than that of Shanghai (-0.88). This suggests that besides the estimation error in μ_i, there are fundamental reasons that the price-to-rent ratio is negatively correlated with latent rent.

Figure 10 Correlation between Estate Fixed Effect in Price-to-Rent Ratio f_i and Estate Fixed Effect in Rent μ_i (Shenzhen)



Note: Figure 10 plots the posterior mean of f_i for estate i against μ_i. f_i is the fixed effect of price-to-rent ratio in Equation (5) and μ_i appears in the rent equation (2). The posterior of estate-specific parameters is simulated by using the MCMC algorithm stated in the paper. The cross-estate correlation between the posterior means of f_i and μ_i is -0.5860 in Figure 10 (Shenzhen data).

Second, the estimation error in μ_i leads to negative bias in the correlation between f_i and μ_i when α is less than unity. To shut down this source of bias, we consider a more restrictive pricing model by setting α_i = 1 in (1), with the rent equation unchanged. The corresponding price-to-rent ratio becomes

$$P_{ij} - \hat{R}_{ij} = c_i + y_{ij}^{'}\beta + \varepsilon_{ij} \tag{7}$$

Now, the estate specific factor f_i = c_i. The resultant cross-estate correlation between the posterior mean of f_i and μ_i is -0.88 for Shanghai and -0.59 for Shenzhen, not much different from the values reported earlier. We conclude that the negative cross-estate correlation between the price-to-rent ratio and latent rent is not a statistical artifact and warrants economic explanations.

4.2 Growth Potentials of Low-Rent High Price-to-Rent Ratio Estates

One explanation for the negative correlation between latent rent and price-to-rent ratio is that low-rent estates tend to have higher potential for future growth and command a higher price-to-rent ratio. The expected return to housing is $\mathrm{E}\left(\frac{future\ price}{current\ price}\right)+\frac{rent}{current\ price}$. Hence in an economy that consists of risk-neutral investors, equalization of the expected return to all estates implies that owners of high price-to-rent properties should expect fast appreciation in housing price. By using metropolitan statistical area (MSA) level data, Capozza & Seguin (1996) show that the price-to-rent ratio is useful in predicting long-run housing price appreciation, but only if the differences in the quality of rental versus owner-occupied housing are controlled.

We find ample evidence in the Shanghai data that supports this theory. Specifically, we discover that estates in newly developed locations that are away from the city center of Shanghai tend to command low rent, high price-to-rent ratio, and better growth potential. There are thirteen administrative districts in Shanghai. Sorting the districts by the average price reveals a pattern in which high-price districts are located in the old city of Shanghai with high population density and established shopping areas and restaurants. In comparison, estates in newly developed suburban districts tend to be cheaper in price and rent, but with higher price-to-rent ratio. The premium in housing price in the low price districts in part reflects the growth potential of the expanding suburban districts in Shanghai. From 2003 to 2005, the population density of the suburban districts increased while that of the historic downtown districts remained unchanged or even decreased. Growth in allocation of public goods is faster in the suburban districts of Shanghai. From 2003 to 2005, the teacher-student ratio increased in suburban districts more than it did in the old city districts. This trend in the improvement of public education in the low rent-high price/rent ratio districts points to their better prospects of price appreciation. We argue that the difference between Shanghai and Shenzhen lends support to the location-dependent growth potential theory. The negative correlation between f_i and μ_i is stronger in Shanghai than in Shenzhen partly because in Shanghai, low-rent estates are more concentrated in the lightly populated suburban districts while districts in Shenzhen are similar, and correlation between rent and growth potential is weaker.

The higher growth of remote districts explains a portion of the negative correlation between f_i and μ_i, but it is unlikely to be the only factor. The within-district correlations between f_i and μ_i are also negative (about -0.7 on average within districts of Shanghai and -0.3 within districts of Shenzhen). This suggests the presence of other factors for the negative correlation between -0.88. These factors are considered in the following.

4.3 Strong Demand of Low-End Units Induced by Factors in Housing Market and Government Policy

The notion that low quality housing is more expensive (relative to the fundamentals) than high quality housing is not new. Sweeney (1974) presents a theoretical model on equilibrium of indivisible goods with quality hierarchy. He shows that replacing a low quality unit by a high quality one creates a chain reaction that lowers the prices of all units with quality higher than the threshold level and raises the prices of all units with quality below that level. The accumulation of the new constructions then makes low-end housing more expensive relative to high-end housing.

The price of housing may be systematically related to factors other than the service flow of housing for various reasons. Genesove & Mayer (1997) show that the price of housing is affected by the owner's equity. Housing units with higher loan-to-equity ratio tend to be more expensive and take longer to sell. Another possible explanation is that low-end housing may carry higher risk premium. For the U.S. market, Sinai & Souleles (2005) show that regional prices of real estates are positively related with the volatility of rent risk in the region. In our sample, if renters of low-end housing are more averse to unexpected increases in rent than renters of high-end housing, or the growth rate in the rent of low-end housing is more uncertain, then owners of low-end housing are willing to pay a higher price. In addition, low-end housing units are more expensive given their hedonic features because they tend to be traded more frequently. The transaction cost, commissions, and transaction tax are paid more times on low-end units. Case et al. (1997) find that properties that transact more frequently tend to be at the lower-end (starters or fixer-uppers), but experience faster price appreciation. Although their finding does not directly relate the frequency of transactions to price-to-rent ratio, it is consistent with the scenario that starter units are more frequently traded at higher price-to-rent ratios.

Stein (1995) studies the effect of down-payment requirement on housing price dynamics. He shows that the requirement of a down-payment amplifies the magnitude of price fluctuations and creates a positive correlation between price movement and trading volume. The down-payment requirement also has a cross-section effect on prices. The high housing price (relative to income) of the two Chinese cities coupled with the thirty percent down-payment requirement for standard mortgages make housing unaffordable for a large fraction of residents. In markets where fast appreciation in housing price is anticipated, households often own the maximum amount of housing permitted by the ability to pay. As the income and saving of potential buyers rise above a threshold, the majority of the first-time home buyers purchase low-end units. This elevates demand for low-end housing and makes it relatively more expensive than high-end housing. The theory also predicts that a larger dispersion in quality means more expensive low-end housing. The evidence in the two cities confirms this prediction. The estate-specific factor is much more diverse in Shanghai than Shenzhen. The negative correlation of the price-to-rent ratio and latent rent is also stronger in Shanghai.

Finally, the demand for low-end housing may be disproportionately raised by government policies. Studies on similar policies in other countries yield the same conclusion. Susin (2002) finds evidence that rent vouchers for low income households in the U.S. elevate the price of low-end housing. Vigdor (2006) finds empirical evidence that U.S. government policies which relax liquidity constraint for veterans pushes up housing price. Mortgage subsidies for low end housing raise its demand for low-end housing in Shanghai and Shenzhen. According to the lending policy of China People's Bank, low income households that purchase low-end housing (with its size and price below a threshold) are eligible for subsidized mortgage loans for 60 months. Estimating the impact of policies is not feasible without more information on the income of the households involved in real estate transactions.

We have discussed a number of possible explanations for the high estimated price/rent ratio of low-end estates. It would be useful to conduct statistical tests on the economic explanations. However, the data needed for such tests are not available at this point. We will leave the empirical estimation of the factors discussed above for future research.

4.4 Policy Implications of High Price of Low-End Housing

An adverse welfare implication of high price-to-rent ratio of low-end housing is that low income households pay a relatively high price for the service of the housing that they own. The policy remedy to the resultant inequity is far from being obvious. Numerous studies (e.g., Susin (2002) and Vigdor (2006) discussed above) show that housing subsidies to low-income families in the form of a mortgage or rent assistance likely make low-end housing more expensive, which at least partially offsets the intended policy objective of housing affordability. Owning low-end housing can be rationalized by a higher expected price appreciation or higher risk premium for hedging rent growth. However, the ownership of low-end housing makes low- and middle-income families more vulnerable to a downturn in the housing market when they invest a large portion of household wealth in housing. Because at the low-end it is more expensive to own than to rent, government policies that promote ownership of low-end housing may reduce the welfare of its owner ex-post.

The large difference in the price-to-rent ratio between high-end and low-end housing found in this study also implies that different segments of housing should be treated as different asset classes from an investment perspective. Owner-occupied housing is the most valuable investment for many households. Higher price-to-rent ratio of low-end housing means that its owner holds an asset that resembles a stock with a high price-to-dividend ratio. Because owner-occupied housing is sorted by income, unlike investment in financial assets that are available to all investors, low-income households unwittingly hold a class of housing investment with different risk characteristics from those held by high-income households. It is useful to analyze investment strategies for households with different income levels while taking into account the difference in risk of owner-occupied housing. For the purpose of diversification, it is useful to construct housing indices of low-end housing as a separate asset class from high-end housing.

5. Conclusion

In this paper, we have investigated how prices of housing units relate to their latent rents. Our cross-sectional analysis of the housing price-to-rent ratio adds to the literature of real-estate pricing based on the present value of rent, hedonic features and repeated sales. The model which consists of price and rent equations contains several thousands of parameters. A Bayesian model selection indicates strong heterogeneity in the parameters across estates. The estimated price-to-rent ratios substantially differ across estates (e.g., the cross-estate standard deviation of the logarithm of price-to-rent ratio is more than forty percent in Shanghai.) We have analyzed the factors that influence the price-to-rent ratio. We find that the estate fixed-effect of the price-rent ratio is negatively correlated with the estate fixed-effect of rent. This finding agrees with several economic theories and affords us new perspectives in investment and public policy in housing.

Acknowledgement

We thank the editor and two referees for their constructive comments. We also thank Centaline (China) Property Consultants Ltd. for providing the Shanghai and Shenzhen real estate data.

References

Ayuso, J. and Restoy, F., (2006), House Prices and Rents: An Equilibrium Asset Pricing Approach, *Journal of Empirical Finance,* **13**, 371–388.

Bao, H. X. and Wan, A. T., (2004), On the Use of Spline Smoothing in Estimating Hedonic Housing Price Models: Empirical Evidence Using Hong Kong Data, *Real Estate Economics,* **32**, 3, 1–25.

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

Can, A., (1990), The Measurement of Neighborhood Dynamics in Urban House Prices, *Economic Geography,* **66**, 3, 254–272.

Capozza, D. R. and Seguin, P. J., (1996), Expectations, Efficiency, and Euphoria in the Housing Market, *Regional Science and Urban Economics,* **26**, 369–386.

Case, B., Pollakowski, H. O., and Wachter, S. M., (1997), Frequency of Transaction and House Price Modeling, *The Journal of Real Estate Finance and Economics,* **14**, 173–187.

Case, K. E. and Shiller, R. J., (1989), The Efficiency of the Market for Single Family Homes, *American Economic Review,* **79**, 1, 125–137.

Clark, T. E., (1995), Rents and Prices of Housing across Areas of the United States: A Cross-section Examination of the Present Value Model, *Regional Science and Urban Economics,* **25**, 237–247.

Clayton, J., (1996), Rational Exceptions, Market Fundamentals and Housing Price Volatility, *Real Estate Economics,* **24**, 441–470.

Coulson, N. E. and Bond, E. W., (1990), A Hedonic Approach to Residential Succession, *The Review of Economics and Statistics,* **72**, 3, 433–444.

Dey, D. K. and Rao, C. R., (2005), *Handbook of Statistics, Volume 25: Bayesian Thinking, Modeling and Computation*, New York: North Holland.

Dokko, Y., Edelstein, R. H., Pomer, M., and Urdang, E. S., (1991), Determinants of the Rate of Return for Nonresidential Real Estate: Inflation Expectations and Market Adjustment Lags, *Real Estate Economics,* **19**, 371–388.

Downes, T. A. and Zabel, J. E., (2002), The Impact of School Characteristics on House Prices: Chicago 1987-1991, *Journal of Urban Economics,* **52**, 1–25.

Dufour, J. M. and Jasiak, J., (2001), Finite Sample Limited Information Inference Methods for Structural Equations and Models with Generated Regressors, *International Economic Review,* **42**, 815–843.

Genesove, D. and Mayer, C. J., (1997), Equity and Time to Sale in the Real Estate Market, *American Economic Review,* **87**, 255–269.

Geweke, J., (2005), *Contemporary Bayesian Econometrics and Statistics*, New York: Wiley.

Goodman, A. C., (1988), An Econometric Model of Housing Price, Permanent Income, Tenure Choice, and Housing Demand', *Journal of Urban Economics,* **23**, 327–353.

Himmelberg, Charles, C. M., and Sinai, T., (2005), Assessing High House Prices: Bubbles, Fundamentals and Misperceptions, *Journal of Economic Perspectives*, Fall, 67–92.

Hughes, Jr., W. T., and Sirmans, C. F., (1992), Traffic Externalities and Single-Family House Prices, *Journal of Regional Science*, **32**, 487–500.

Hui, S. K., Cheung, A., and Pang, J., (2010), A Hierarchical Bayesian Approach for Residential Property Valuation: Application to Hong Kong Housing Market, *International Real Estate Review*, **13**, 1–29.

Kass, R. E., and Raftery, A. E., (1995), Bayes Factors and Model Uncertainty, *Journal of the American Statistical Association,* **90**, 773–795.

Kim, S., (1992), Search, Hedonic Prices and Housing Demand, *The Review of Economics and Statistics*, **74**, 3, 503–508.

Mankiw, N. G., and Weil, D. N., (1989), The Baby Boom, the Baby Bust and the Housing Market, *Regional Science and Urban Economics*, **19**, 235–238.

Mats, W., (2002), Household Expenditure Pattern of Housing Attributes: A Linear Expenditure System with Hedonic Prices, *Journal of Housing Economics,* **1**, 1, 75–93.

McCue, T., and Kling, J. L., (2006), Real Estate Returns and the Macroeconomy: Some Empirical Evidence from Real Estate Investment Trust Data, 1972-1991, *Journal of Real Estate Research,* **9**, 277–287.

Meese, R., and Wallace, N., (1994), Testing the Present Value Relation for Housing Prices: Should I Leave My House in San Francisco?, *Journal of Urban Economics*, **35**, 245–266.

O 'Hagan, A., and West, M., (2010), *The Oxford Handbook of Applied Bayesian Analysis*, Oxford: Oxford University Press.

Ridker, R. G., and Henning, J. A., (1967), The Determinants of Residential Property Values with Special Reference to Air Pollution, *The Review of Economics and Statistics,* **49**,2, 246–257.

Rossi, P. E., and Allenby, G. M., (1993), A Bayesian Approach to Estimating Household Parameters, *Journal of Marketing Research,* **30**, 171–182.

Schwartz, A. E., Susin, S., and Voicu, I., (2003), Has Falling Crime Rate Driven New York City's Real Estate Boom?, *Journal of Housing Research,* **14**, 101–135.

Sinai, T. M., and Souleles, N. S., (2005), Owner Occupied Housing as a Hedge Against Rent Risk, *Quarterly Journal of Economics,* **120**, 2, 763–789.

Stein, J. C., (1995), Prices and Trading Volume in the Housing Market: A Model with Downpayment Effects, *The Quarterly Journal of Economics,* **110**, 379–406.

Susin, S., (2002), Rent Vouchers and the Price of Low-Income Housing, *Journal of Public Economics,* **83**, 109–152.

Sweeney, J. L., (1974), Quality, Commodity Hierarchies, and Housing Markets, *Econometrica,* **42**, 147–167.

Vanderford, S. E., Mimura, Y., and Sweaney, A. L., (2005), A Hedonic Price Comparison of Manufactured and Site-Built Homes in the Non-MSA United States, *Journal of Real Estate Research,* **27**, 1, 83–104.

Vigdor, J. L., (2006), Liquidity Constraints and Housing Prices: Theory and Evidence from the VA Mortgage Program, *Journal of Public Economics,* **90**, 1579–1600.

Witte, A. D., Sumka, H. J., and Erekson, H., (1979), An Estimation of a Structural Hedonic Price Model of the Housing Market: An Application of Rosen's Theory of Implicit Market, *Econometrica,* **47**, 5, 1151–1174.

Zellner, A., (1970), Estimation of Regression Relationships Containing Unobservable Independent Variables, *International Economic Review,* **11**, 441–454.

Appendix

The joint posterior is:

$$\pi\left(\alpha,\beta,\mu,\theta,\sigma,\tau,c\middle|D\right)$$

$$\propto L\left(P\middle|\hat{R},\alpha,c,\beta,\sigma\right)\pi\left(\hat{R}\middle|\mu,\theta,\tau,\hat{x}\right)L(\left(R\middle|\theta,\mu,\tau\right)\pi(\beta)\pi(\theta)\pi(\sigma)\pi(\tau)\pi(\alpha)\pi(\mu)\pi(c)$$

$$\propto\prod_{i=1}^{I}\exp\left\{-\frac{\sum_{j=1}^{J_i}\left[P_{ij}-\left(c_i+\alpha_i\hat{R}_{ij}+y_{ij}'\beta\right)\right]^2}{2\sigma_i^2}-\frac{\sum_{j=1}^{J_i}\left[\hat{R}_{ij}-\left(\mu_i+\hat{x}_{ij}'\theta\right)\right]^2}{2\tau_i^2}\right\}$$

$$\times\prod_{i=1}^{I}\exp\left\{-\frac{\sum_{k=1}^{K_i}\left[R_{ik}-\left(\mu_i+x_{ik}'\theta\right)\right]^2}{2\tau_i^2}\right\}\exp\left\{-\frac{1}{2}\left(\beta-\overline{\beta}\right)'\mathrm{B}^{-1}\left(\beta-\overline{\beta}\right)\right\}$$

$$\times\exp\left\{-\frac{1}{2}\left(\theta-\overline{\theta}\right)'\Omega^{-1}\left(\theta-\overline{\theta}\right)\right\}\prod_{i=1}^{I}\left\{\sigma_i^{-2(s_\sigma+1)}\exp\left(-\frac{v_\sigma}{\sigma_i^2}\right)\sigma_i^{-J_I}\tau_i^{-2(s_\tau+1)}\exp\left(-\frac{v_\tau}{\tau_i^2}\right)\tau_i^{-K_I}\right\}$$

$$\times\prod_{i=1}^{I}\left\{\exp\left[-\frac{(\alpha_i-\overline{\alpha})^2}{2v_\alpha^2}\right]\exp\left[-\frac{(\mu_i-\overline{\mu})^2}{2v_\mu^2}\right]\exp\left[-\frac{(c_i-\overline{c})^2}{2v_c^2}\right]\right\}.$$

The following conditional posteriors are used for Gibbs sampling:

(i) $\pi\left(\mu_i\middle|\ \tau_i,\theta,\hat{R}_{i,}D\right)$

$$\propto\exp\left\{-\frac{(\mu_i-\overline{\mu})^2}{2\overline{v}_\mu^2}\right\}\exp\left\{-\frac{\sum_{k=1}^{K_i}\left[\mu_i-\tilde{\mu}_{ik}\right]^2}{2\tau_i^2}\right\}\exp\left\{-\frac{\sum_{j=1}^{J_i}\left[\mu_i-\hat{\mu}_{ij}\right]^2}{2\tau_i^2}\right\}$$

$$\propto N\left(\frac{\frac{\overline{\mu}}{\overline{v}_\mu^2}+\frac{\sum_{k=1}^{K_i}\tilde{\mu}_{ik}+\sum_{j=1}^{J_i}\hat{\mu}_{ij}}{\tau_i^2}}{\frac{1}{\overline{v}_\mu^2}+\frac{K_i+J_i}{\tau_i^2}},\frac{1}{\frac{1}{\overline{v}_\mu^2}+\frac{K_i+J_i}{\tau_i^2}}\right)$$

where $\tilde{\mu}_{ik}=R_{ik}-x_{ik}'\theta$, $\hat{\mu}_{ij}=\hat{R}_{ij}-x_{ij}'\theta$

(ii) $\pi\left(c_i \middle| \alpha_i, \beta, \theta, \sigma_i, \hat{R}_{i,} D\right)$

$$\propto \exp\left\{-\frac{(c_i-\bar{c})^2}{2\bar{v}_c^2}\right\}\exp\left\{-\frac{\sum_{j=1}^{J_i}\left[c_i-\hat{c}_{ij}\right]^2}{2\sigma_i^2}\right\} \propto N\left(\frac{\frac{\bar{c}}{\bar{v}_c^2}+\frac{\sum_{j=1}^{J_i}\hat{c}_{ij}}{\sigma_i^2}}{\frac{1}{\bar{v}_c^2}+\frac{J_i}{\sigma_i^2}}, \frac{1}{\frac{1}{\bar{v}_c^2}+\frac{J_i}{\sigma_i^2}}\right),$$

where $\hat{c}_{ij} = P_{ij} - \alpha_i \hat{R}_{ij} - y_{ij}^{'}\beta$

(iii) $\pi\left(\alpha_i \middle| c_i, \beta, \sigma_i, \hat{R}_{i,} D\right) \propto \exp\left\{-\frac{(\alpha_i-\bar{\alpha})^2}{2\bar{v}_\alpha^2}\right\}\exp\left\{-\frac{\sum_{j=1}^{J_i}\left[\alpha_i\hat{R}_{ij}-w_{ij}\right]^2}{2\sigma_i^2}\right\}$

$$\propto N\left(\frac{\frac{\bar{\alpha}}{\bar{v}_\alpha^2}+\frac{\sum_{j=1}^{J_i}\hat{R}_{ij}w_{ij}}{\sigma_i^2}}{\frac{1}{\bar{v}_\alpha^2}+\frac{\sum_{j=1}^{J_i}\hat{R}_{ij}^2}{\sigma_i^2}}, \frac{1}{\frac{1}{\bar{v}_\alpha^2}+\frac{\sum_{j=1}^{J_i}\hat{R}_{ij}^2}{\sigma_i^2}}\right),$$

where $w_{ij} = P_{ij} - c_i - y_{ij}^{'}\beta$

(iv) $\pi\left(\beta \middle| c, \alpha, \sigma, \hat{R}_{i,} D\right)$

$$\propto \exp\left\{-\frac{1}{2}\left(\beta-\bar{\beta}\right)^{'}\mathrm{B}^{-1}\left(\beta-\bar{\beta}\right)\right\}\exp\left\{-\sum_{i=1}^{I}\frac{\left(g_i-y_i\beta\right)^{'}\left(g_i-y_i\beta\right)}{2\sigma_i^2}\right\}$$

$$\propto N\left(\left[\mathrm{B}^{-1}+\sum_{i=1}^{I}\sigma_i^{-2}y_i^{'}y_i\right]^{-1}\left[\mathrm{B}^{-1}\bar{\beta}+\sum_{i=1}^{I}\sigma_i^{-2}y_i^{'}g_i\right], \left[\mathrm{B}^{-1}+\sum_{i=1}^{I}\sigma_i^{-2}y_i^{'}y_i\right]^{-1}\right),$$

where $g_i = P_i - c_i I - \alpha_i \hat{R}_i$.

(v) $\pi\left(\theta \middle| \mu, \tau, \hat{R}_{i,} D\right)$

$$\propto \exp\left\{-\frac{1}{2}\left(\theta-\bar{\theta}\right)^{'}\Omega^{-1}\left(\theta-\bar{\theta}\right)\right\}\exp\left\{-\sum_{i=1}^{I}\frac{\left(q_i-x_i\theta\right)^{'}\left(q_i-x_i\theta\right)+\left(\hat{q}_i-\hat{x}_i\theta\right)^{'}\left(\hat{q}_i-\hat{x}_i\theta\right)}{2\tau_i^2}\right\}$$

$$\propto N\left(\frac{\Theta^{-1}\bar{\theta}+\sum_{i=1}^{I}\tau_i^{-2}\left(x_i^{'}q_i+\hat{x}_i^{'}\hat{q}_i\right)}{\Theta^{-1}+\sum_{i=1}^{I}\tau_i^{-2}\left(x_i^{'}x_i+\hat{x}_i^{'}\hat{x}_i\right)}, \left[\Theta^{-1}+\sum_{i=1}^{I}\tau_i^{-2}\left(x_i^{'}x_i+\hat{x}_i^{'}\hat{x}_i\right)\right]^{-1}\right),$$

where $q_i = R_i - \mu_i \mathrm{I},\ \hat{q}_i = \hat{R}_i - \mu_i \mathrm{I}$

(vi) $\pi\left(\sigma_i^2 \middle| c_i, \alpha_i, \beta, \hat{R}_{i,} D\right) \propto \left(\sigma_i^2\right)^{-(s_\sigma+1)} \exp\left(-\frac{v_\sigma}{\sigma_i^2}\right)\left(\sigma_i^2\right)^{-\frac{J_i}{2}} \exp\left(-\frac{\hat{v}_i}{2\sigma_i^2}\right)$

$$\propto IG\left(s_\sigma + \frac{J_i}{2}, v_\sigma + \frac{\hat{v}_i}{2}\right),$$

where $\hat{v}_i = \sum_{j=1}^{J_i}\left[P_{ij} - \left(c_i + \alpha_i \hat{R}_{ij} + y_{ij}^{'}\beta\right)\right]^2$.

(vii) $\pi\left(\tau_I^2 \middle| \mu_i, \theta, \hat{R}_{i,} D\right)$

$$\propto \left(\tau_i^2\right)^{-(s_\tau+1)} \exp\left(-\frac{v_\tau}{\tau_i^2}\right)\left(\tau_i^2\right)^{-\frac{K_i}{2}}\left(\tau_i^2\right)^{-\frac{J_i}{2}} \exp\left(-\frac{\tilde{h}_i}{2\tau_i^2}\right)\exp\left(-\frac{\hat{h}_i}{2\tau_i^2}\right)$$

$$\propto IG\left(s_\tau + \frac{K_I + J_i}{2}, v_\tau + \frac{\tilde{h}_i + \hat{h}_i}{2}\right),$$

where $\tilde{h}_i = \sum_{k=1}^{K_i}\left[R_{ik} - \left(\mu_i + x_{ik}^{'}\theta\right)\right]^2$, $\hat{h}_i = \sum_{j=1}^{J_i}\left[\hat{R}_{ij} - \left(\mu_i + \hat{x}_{ij}^{'}\theta\right)\right]^2$.

(viii) $\pi\left(\hat{R}_{ij} \middle| \mu_i, \alpha_i, c_i, \beta, \theta, \sigma_i, \tau_i, D\right)$

$$\propto \exp\left\{-\frac{\left[P_{ij} - \left(c_i + \alpha_i \hat{R}_{ij} + y_{ij}^{'}\beta\right)\right]^2}{2\sigma_i^2} - \frac{\left[\hat{R}_{ij} - \left(\mu_i + \hat{x}_{ij}^{'}\theta\right)\right]^2}{2\tau_i^2}\right\}$$

$$\propto N\left(\frac{\frac{\alpha_i w_i}{\sigma_i^2} + \frac{\mu_i + x_{ij}^{'}\theta}{\tau_i^2}}{\frac{1}{\tau_i^2} + \frac{\alpha_i^2}{\sigma_i^2}}, \frac{1}{\frac{1}{\tau_i^2} + \frac{\alpha_i^2}{\sigma_i^2}}\right),$$

where $w_{ij} = P_{ij} - c_i - y_{ij}^{'}\beta$.

The conditional posteriors suggest a Gibbs sampling MCMC algorithm: In cycle *k*, with $\left(c_i^{k-1}, \mu_i^{k-1}, \alpha_i^{k-1}, \sigma_i^{k-1}, \tau_i^{k-1}, \theta^{k-1}, \beta^{k-1}\right)$ drawn already,

(1) Draw $\left(\mu_i^k \middle| \tau_i^{k-1}, \theta^{k-1}, \hat{R}_i^{k-1}, D\right)$ (for all $i = 1, \cdots, I$) from the normal distribution in (i).

(2) Draw $\left(c_i^k \middle| \alpha_i^{k-1}, \sigma_i^{k-1}, \beta^{k-1}, \hat{R}_i^{k-1}, D\right)$ (for all $i = 1, \cdots, I$) from the normal distribution in (ii).

(3) Draw $\left(\alpha_i^k \middle| c_i^k, \sigma_i^{k-1}, \beta^{k-1}, \hat{R}_i^{k-1}, D\right)$ (for all $i = 1, \cdots, I$) from the normal distribution in (iii).

(4) Draw $\left(\beta^k \middle| c_i^k, \alpha_i^k, \sigma_i^{k-1}, \hat{R}_i^{k-1}, D\right)$ from the normal distribution in (iv).

(5) Draw $\left(\theta^k \middle| \mu_i^k, \tau_i^{k-1}, \hat{R}_i^{k-1}, D\right)$ from the normal distribution in (v).

(6) Draw $\left(\sigma_i^k \middle| c_i^k, \alpha_i^k, \beta^k, \hat{R}_i^{k-1}, D\right)$ (for all $i = 1, \cdots, I$) from the IG distribution in (vi) (for all $i = 1, \cdots, I$).

(7) Draw $\left(\tau_i^k \middle| \mu_i^k, \theta^k, \hat{R}_i^{k-1}, D\right)$ (for all $i = 1, \cdots, I$) from the IG distribution in (vii) (for all $i = 1, \cdots, I$).

(8) Draw $\left(\hat{R}_{ij}^k \middle| c_i^k, \alpha_i^k, \mu_i^k, \tau_i^k, \sigma_i^k, \beta^k, \theta^k, D\right)$ (for all $i = 1, \cdots, I$) from the normal distribution in (viii).

Some of the Gibbs sampling steps can be combined to simulate a larger block of parameters, for example, Steps (2) and (3) can be combined to simulate (c, α). We present the algorithm based on the conditional posterior for each parameter vector because it is most transparent and applicable to different models (for example, the four combinations that c and/or α may be estate-specific or constant across estates).