Decompositions of Spatially Varying Quantile Distribution Estimates:

The Rise and Fall of Tokyo House Prices

Daniel McMillen

University of Illinois at Urbana-Champaign

Chihiro Shimizu

Nihon University

Outline

- Quantile procedures for estimating house price indices across the full distribution of prices:
  1. Hedonic
  2. Locally weighted versions to allow for smooth variation in appreciation rates over space.
  3. Tokyo & Chicago
  4. Decomposition of the change in distributions over time into the portions due to changes in the coefficients, explanatory variables, and location of the sales. Condo prices in Tokyo, 1986 – 1990 and 1991 – 1995.

Conventional House Price Indices

- Medians or Means
- “Quality – Controlled” Indices
  - Hedonic model with controls for structural characteristics
  - Repeat Sales
  - Hybrid

All focus on a central tendency – mean or median.

Quantile Price Indices

- $Q_{lnP}(q|X_{it}, D_{it}) = X_i\beta(q) + \sum_{t=2}^{T} D_{it}\delta_t(q)$
- q = .50 is comparable to hedonic estimation. Also directly comparable to repeat sales estimator if the sample is restricted to properties that have sold at least twice.
- Can trace out the full distribution by estimating across many quantiles.

Price Index Estimation

- Estimating Equation: $Q_{lnP}(q|X_{it}, D_{it}) = X_i\beta(q) + \sum_{t=2}^{T} D_{it}\delta_t(q)$
- Estimates in base period:
$$X\hat{\beta}(q)$$
- Estimates at time t:
$$X\hat{\beta}(q) + \hat{\delta}_t(q)$$
- n x B estimates in each period, where B is the number of quantiles.
- Kernel density function for the nB-vector of estimated prices shows the full distribution of predicted sales prices in time 0 and time t. Set all other time variables to 0 but keep X at actual values.

Other Counter-Factual Distributions

- Example: house age = 25, 75
- $Q_{lnP}(q|Age_i, X_i, D_{it}) = Age_i\lambda\,(q) + X_i\beta(q) + \sum_{t=2}^{T} D_{it}\delta_t(q)$
- Predicted lnP at time t for age = 25, 75:

$$25\hat{\lambda}(q) + X_i\hat{\beta}(q) + \hat{\delta}_t(q)$$

$$75\hat{\lambda}(q) + X_i\hat{\beta}(q) + \hat{\delta}_t(q)$$

- Kernel density function for the nB-vector of estimated prices shows the full distribution of predicted sales prices for age = 25 and age = 75.

Locally Weighted Estimation

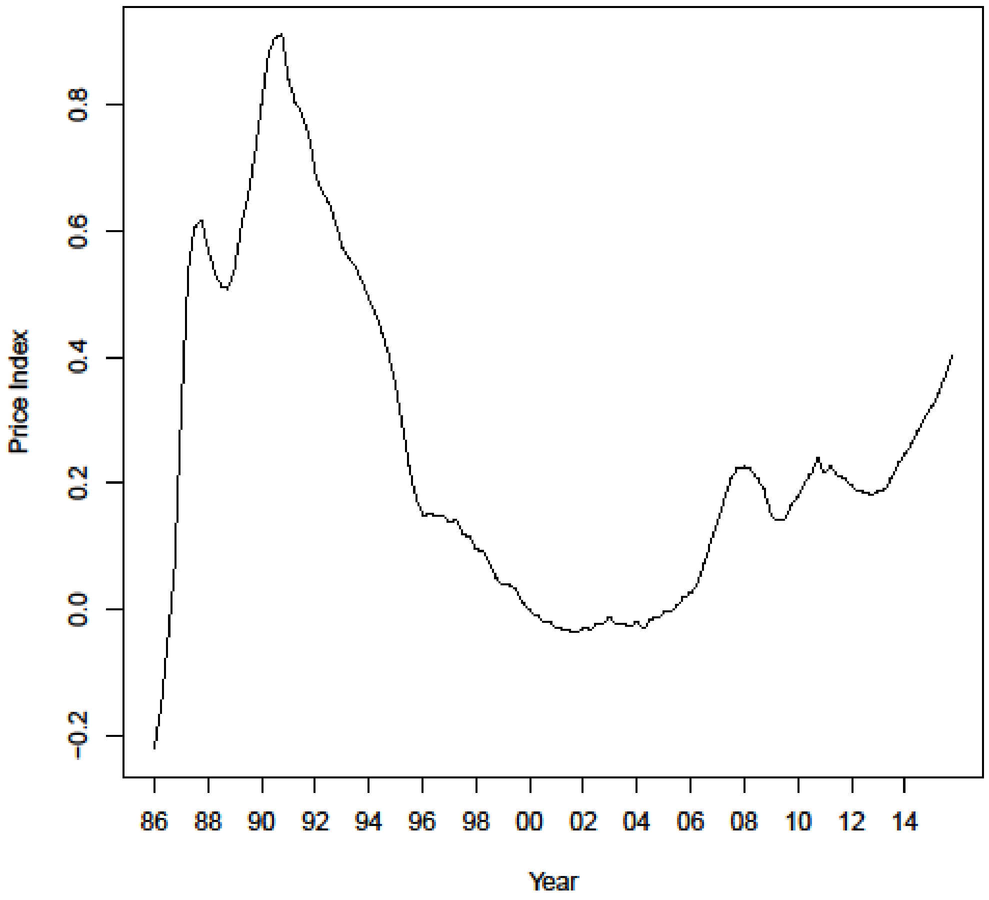
- Coefficients vary over space, e.g., z1 = longitude and z2 = latitude:

- $Q_{lnP}(q|X_i, D_{it}, z_{1i}, z_{2i}) = X_i\beta(q, z_{1i}, z_{2i}) + \sum_{t=2}^{T} D_{it}\delta_t(q, z_{1i}, z_{2i})$

- A weighted version of quantile regression. Weights decline with distance to target points. Can interpolate from target points to full data set.

- With k variables and B quantiles, have n x k x B estimated coefficients. But still have n x B predicted values. Again can use kernel density functions to summarize the results.

Tokyo: 226,983 condo sales, 1986 - 2014



Descriptive Statistics, 1986 – 1990, 32,029 sales

Variable	1986 – 1990	1991 – 1995
Price per Square Meter	104.721	88.079
Log Price per Square Meter	4.553	4.408
Building Area (square meters)	50.466	54.249
Log Building Area	3.854	3.928
Age	9.721	12.864
South View	0.209	0.383
1st Floor	0.098	0.097
2nd Floor	0.158	0.157
Floor (range 1 – 25)	4.753	4.766

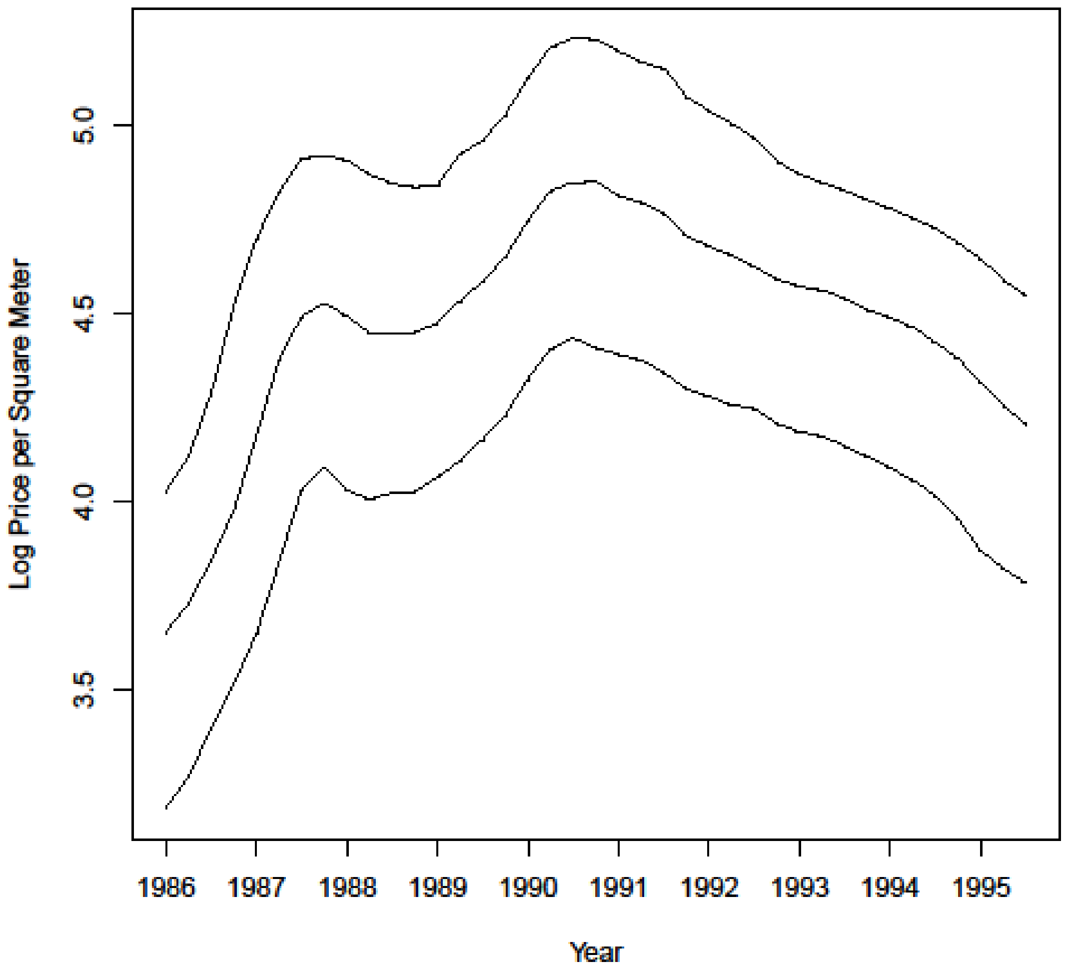
Estimates for 1986-1990 with controls for location and time

Variable	OLS	0.1	0.5	0.9	Quantile .9 - .1
Log Building Area	-0.0431 (0.0028)	-0.0258 (0.0038)	-0.0519 (0.0022)	-0.0584 (0.0041)	-0.0326 (0.0040)
Age	-0.0223 (0.0002)	-0.0246 (0.0003)	-0.0223 (0.0002)	-0.0218 (0.0003)	0.0028 (0.0004)
South View	0.0048 (0.0025)	0.0056 (0.0040)	0.0088 (0.0023)	-0.0009 (0.0044)	-0.0064 (0.0037)
1st Floor	-0.0140 (0.0035)	-0.0174 (0.0057)	-0.0168 (0.0033)	-0.0116 (0.0063)	0.0058 (0.0085)
2nd Floor	-0.0061 (0.0028)	-0.0040 (0.0046)	-0.0062 (0.0027)	-0.0079 (0.0051)	-0.0039 (0.0070)
Floor	0.0062 (0.0004)	0.0085 (0.0006)	0.0061 (0.0004)	0.0044 (0.0007)	-0.0041 (0.0008)

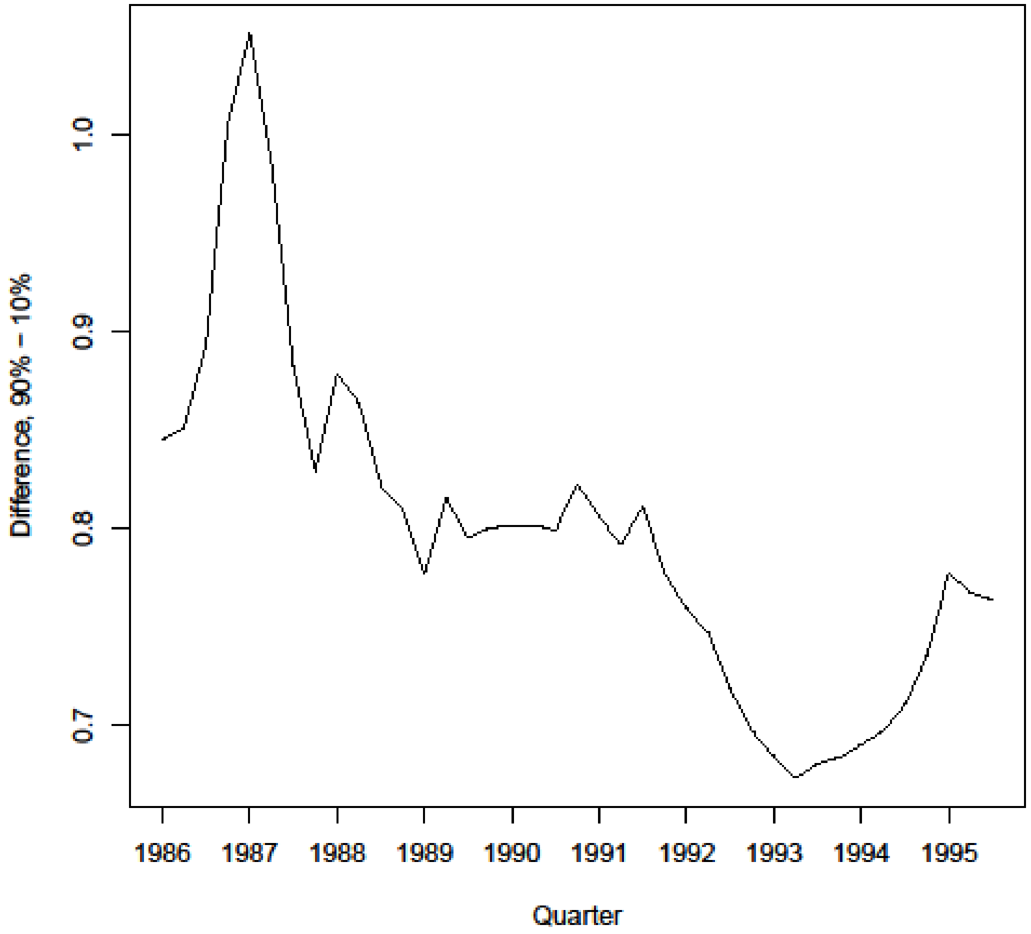
Estimates for 1991 - 1995 with controls for location and time

Variable	OLS	0.1	0.5	0.9	.9 - .1
Log Building Area	-0.0440 (0.0020)	0.0104 (0.0026)	-0.0472 (0.0017)	-0.0686 (0.0030)	-0.0790 (0.0048)
Age	-0.0221 (0.0001)	-0.0233 (0.0002)	-0.0221 (0.0001)	-0.0213 (0.0002)	0.0020 (0.0002)
South View	0.0135 (0.0014)	0.0132 (0.0020)	0.0158 (0.0013)	0.0104 (0.0023)	-0.0028 (0.0026)
1st Floor	-0.0117 (0.0025)	-0.0128 (0.0038)	-0.0129 (0.0025)	-0.0109 (0.0045)	0.0020 (0.0052)
2nd Floor	-0.0063 (0.0020)	-0.0043 (0.0031)	-0.0072 (0.0020)	-0.0082 (0.0036)	-0.0039 (0.0040)
Floor	0.0057 (0.0003)	0.0065 (0.0004)	0.0055 (0.0003)	0.0049 (0.0005)	-0.0016 (0.0006)

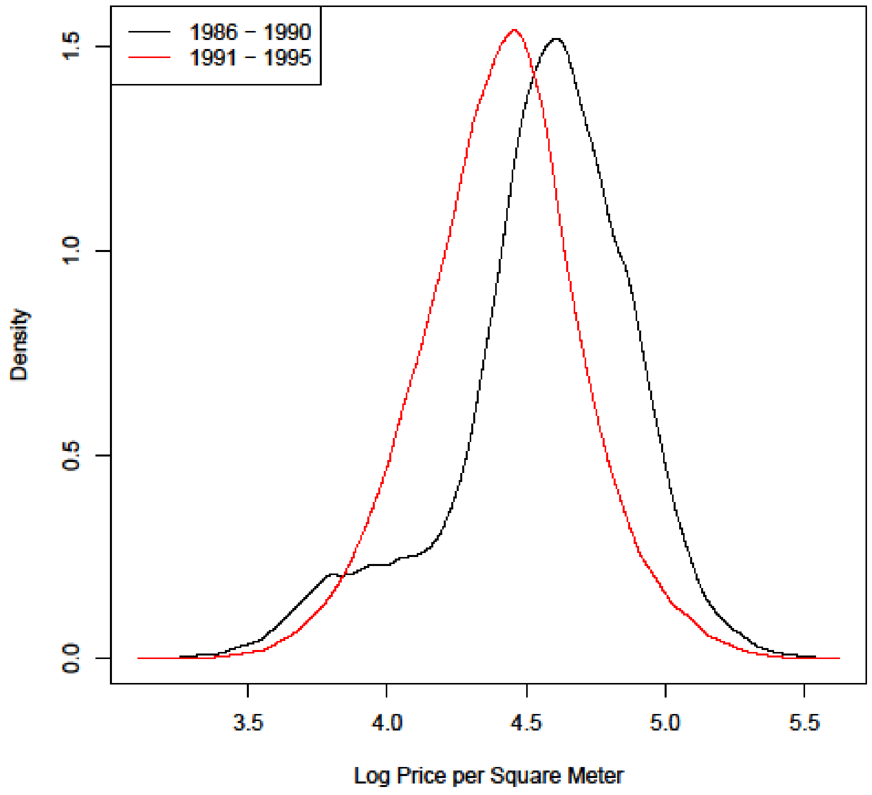
10%, 50%, and 90% Quantile Estimates

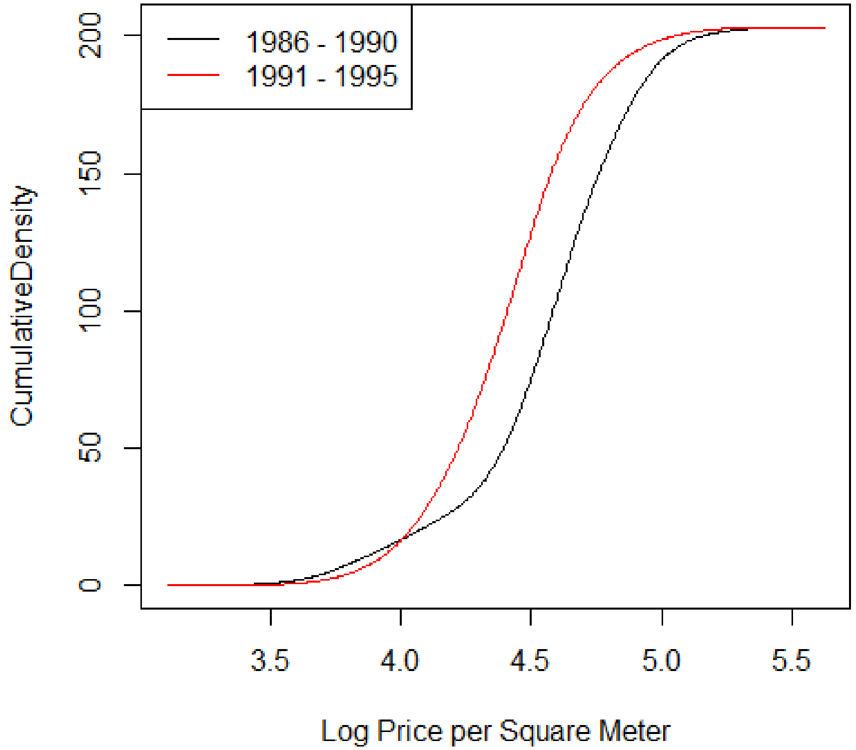


Difference between 90% and 10% Quantile Estimates, 1986 – 1995

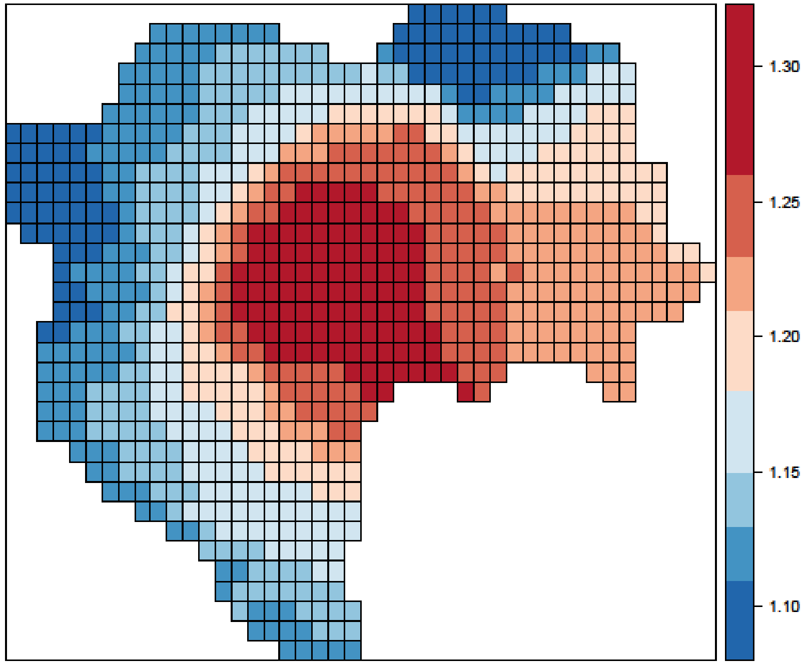


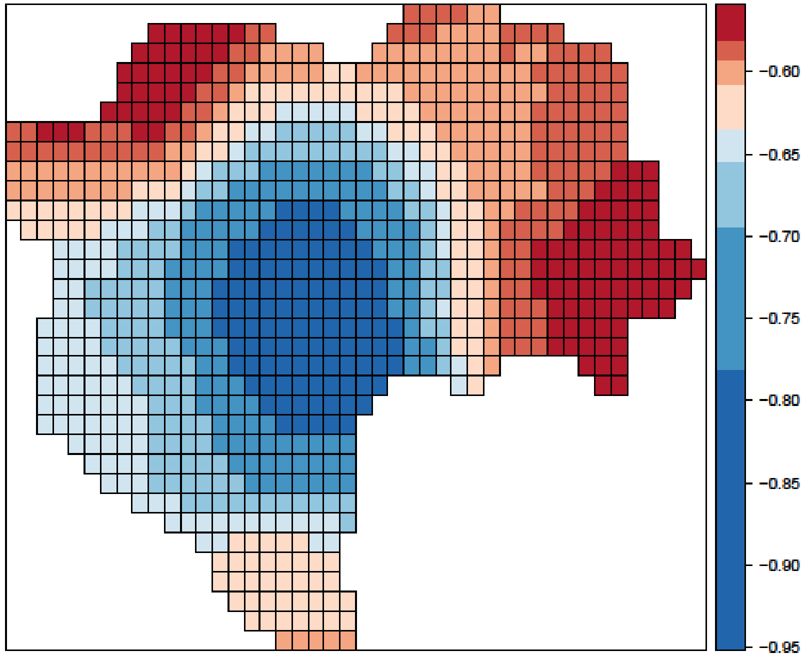
Kernel Density Estimates for Log Price per Square Meter

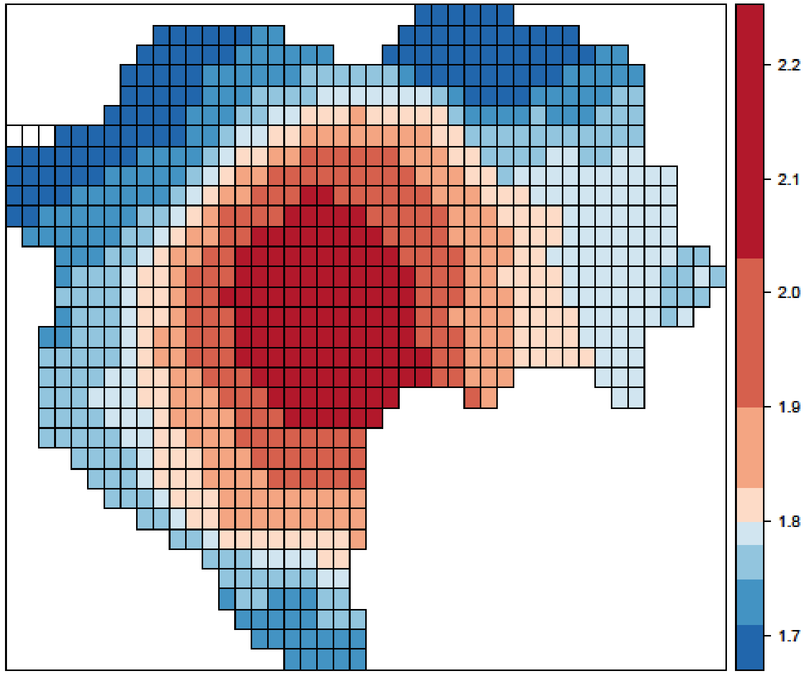




Spatially Varying Median Appreciation Rates





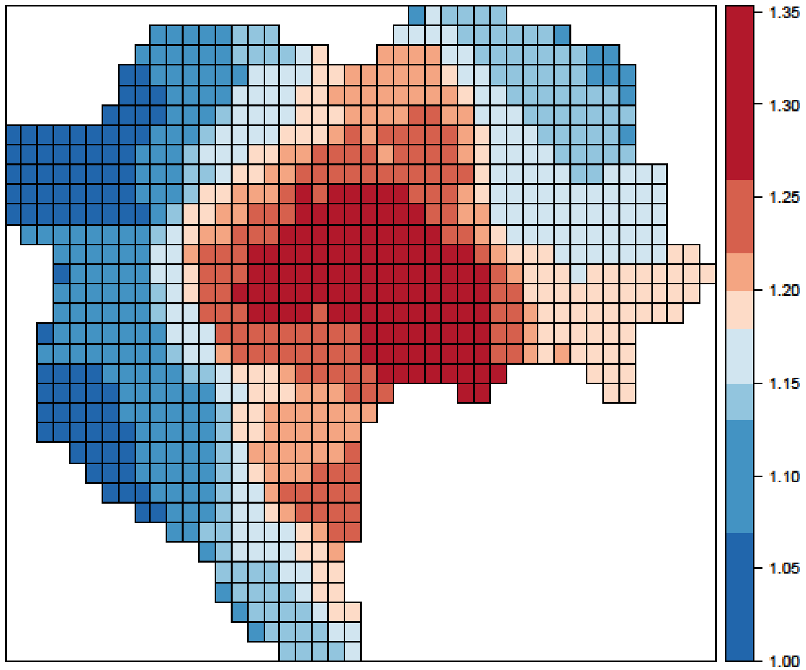


1986 – 1990

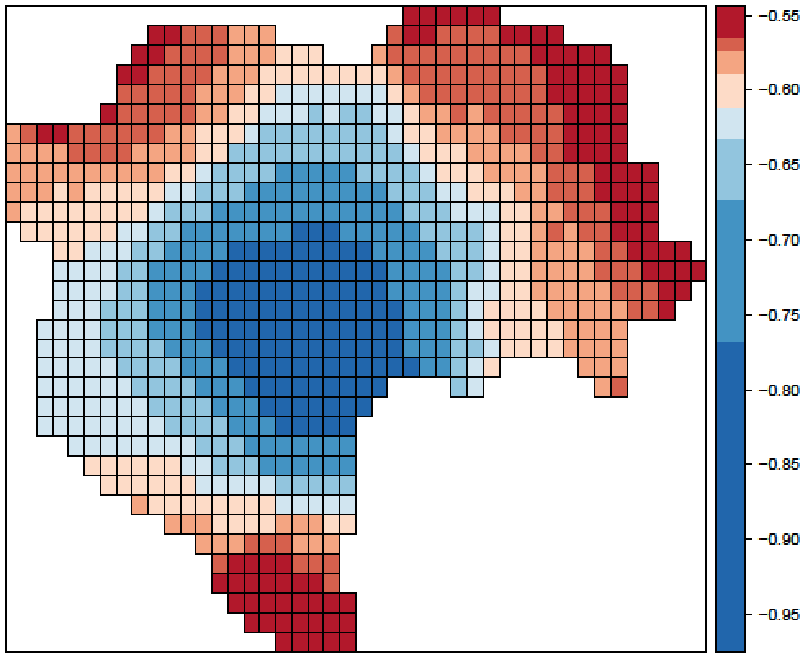
1991 – 1995

Difference, 1986-90 – 1991-95

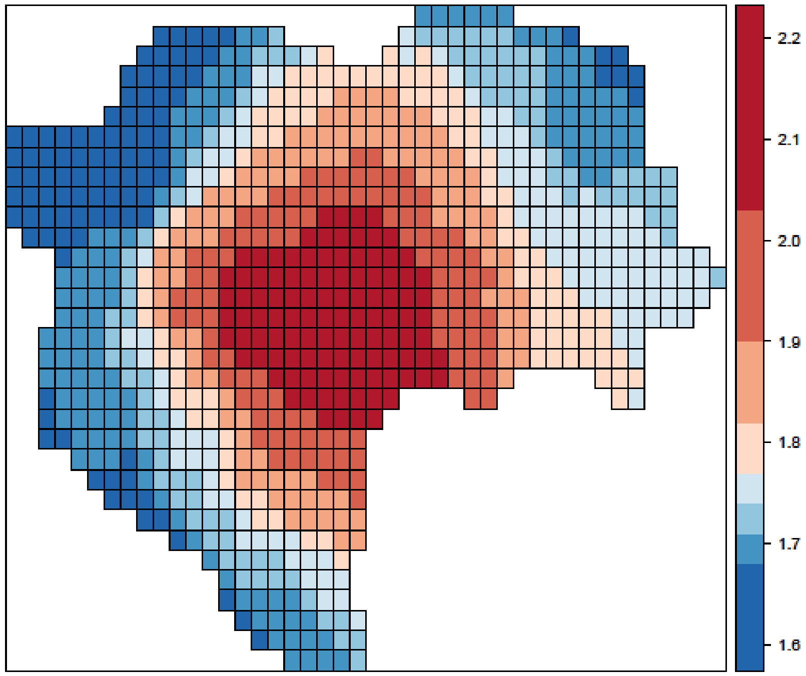
Spatially Varying 10% Appreciation Rates



1986 – 1990

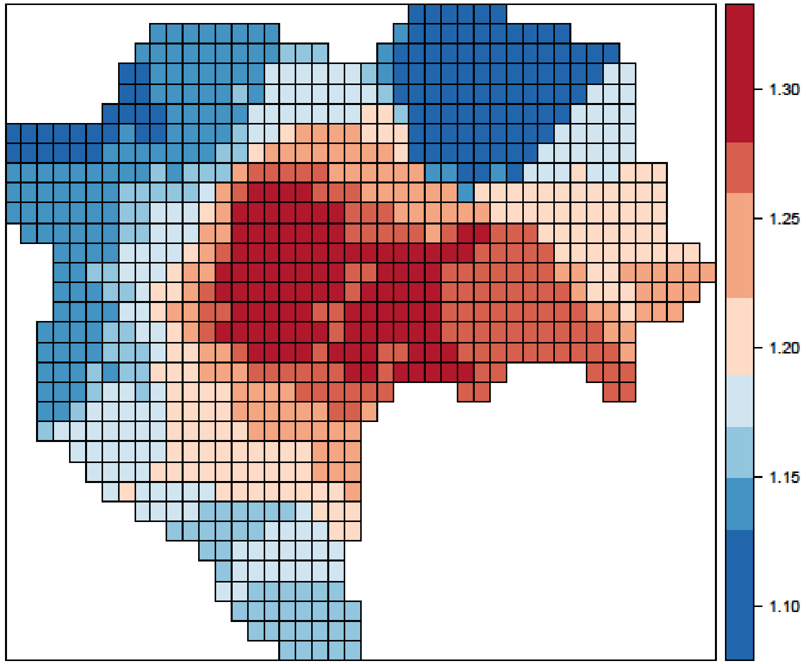


1991 – 1995

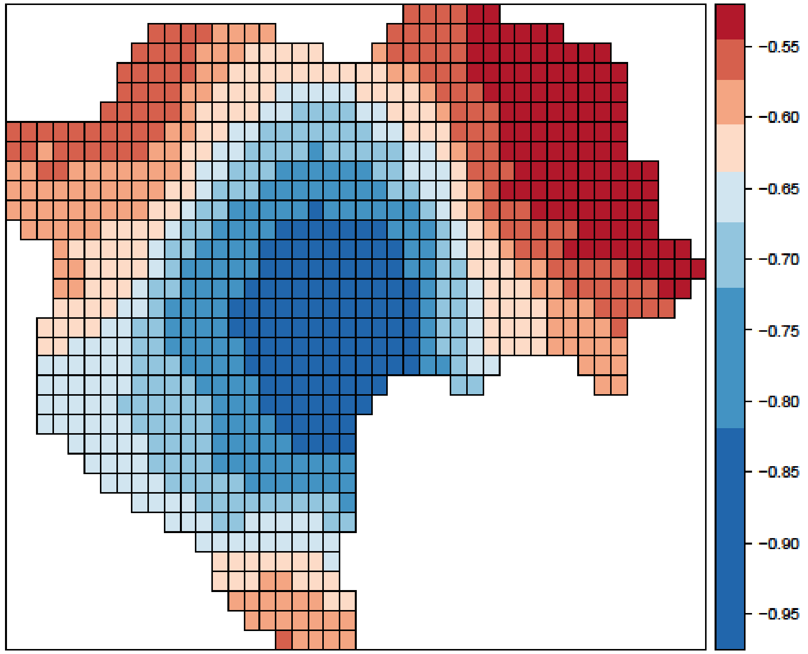


Difference, 1986-90 – 1991-95

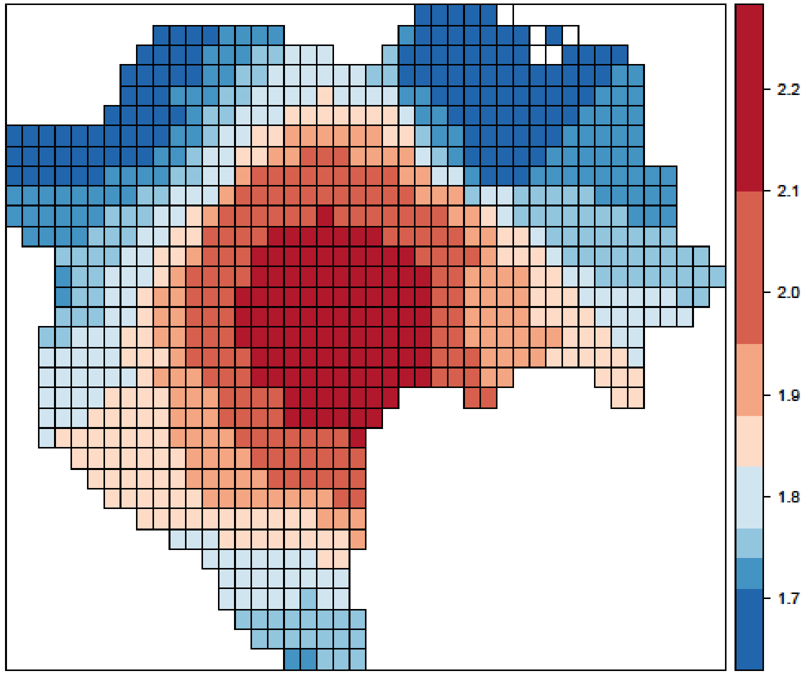
Spatially Varying 90% Appreciation Rates



1986 – 1990



1991 – 1995



Difference, 1986-90 – 1991-95

Decompositions: Non-Spatial Version

- Machado & Mata, Journal of Applied Econometrics (2005)

$$y_1(\tau) = x_1\beta_1(\tau) + \delta_1(\tau)$$

$$y_2(\tau) = x_2\beta_2(\tau) + \delta_2(\tau)$$

$$y_1(\tau) - y_2(\tau)$$

$$= \underbrace{\left(x_1\beta_1(\tau) - x_2\beta_1(\tau)\right)}_{\text{Variables}} + \underbrace{\left(x_2\beta_1(\tau) - x_2\beta_2(\tau)\right)}_{\text{Coefficients}}$$

$$+\underbrace{\left(\delta_1(\tau) - \delta_2(\tau)\right)}_{\text{Time}}$$

Computations

- n observations, k explanatory variables, B quantiles
- $x_1\beta_1(\tau)$: x_1 is $n_1$xk, $\beta_1(\tau)$ is kxB
- $x_2\beta_2(\tau)$: x_2 is $n_2$xk, $\beta_2(\tau)$ is kxB
- $x_2\beta_1(\tau)$: x_2 is $n_2$xk, $\beta_1(\tau)$ is kxB

Machado & Mata: Draw M draws from the rows of x_1 and the columns of $\beta_1(\tau)$. Also make M draws from the rows of x_2 and the columns of and $\beta_2(\tau)$.

Construct $x_1\beta_1(\tau)$, $x_2\beta_2(\tau)$, and $x_2\beta_1(\tau)$ from the new Mxk and kxM matrices.

Spatially varying coefficients

$$y_1(z_1,\tau) = x_1\beta_1(z_1,\tau) + \delta_1(z_1,\tau)$$

$$y_2(z_2,\tau) = x_2\beta_2(z_2,\tau) + \delta_2(z_2,\tau)$$

$$y_1(z_1,\tau) - y_2(z_2,\tau)$$

$$= \underbrace{\big(x_1\beta_1(z_1,\tau) - x_2\beta_1(z_1,\tau)\big)}_{\text{Variables}} + \underbrace{\big(x_2\beta_1(z_1,\tau) - x_2\beta_2(z_1,\tau)\big)}_{\text{X Coefficients}} + \underbrace{\big(\delta_1(z_1,\tau) - \delta_2(z_1,\tau)\big)}_{\text{Time Coefficients}}$$

$$+ \underbrace{\big(x_2\beta_2(z_1,\tau) - x_2\beta_2(z_2,\tau)\big)}_{\text{Location, X}} + \underbrace{\big(\delta_2(z_1,\tau) - \delta_2(z_2,\tau)\big)}_{\text{Location, Time}}$$

Need to estimate $\beta_2(z_1,\tau)$ and $\delta_2(z_1,\tau)$: the coefficients in time 2 at the time 1 sale locations.

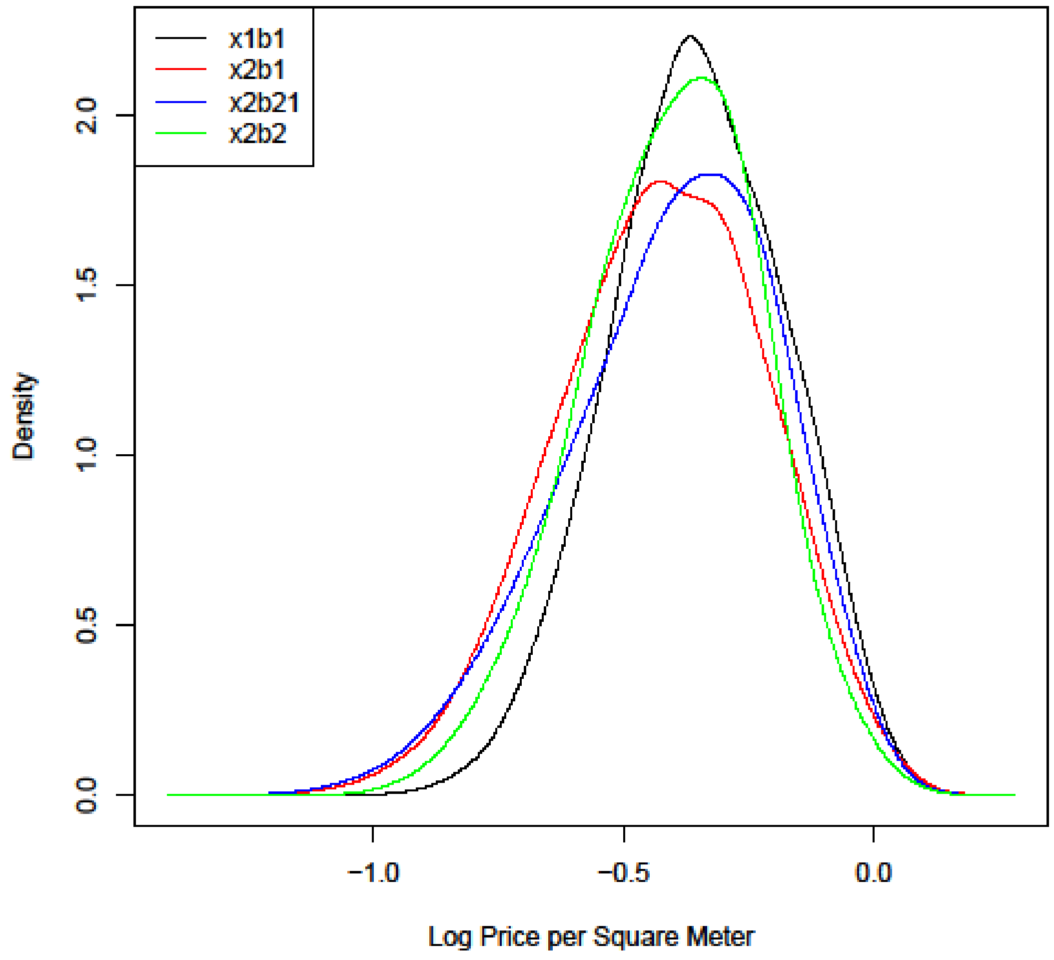
Spatially varying coefficients

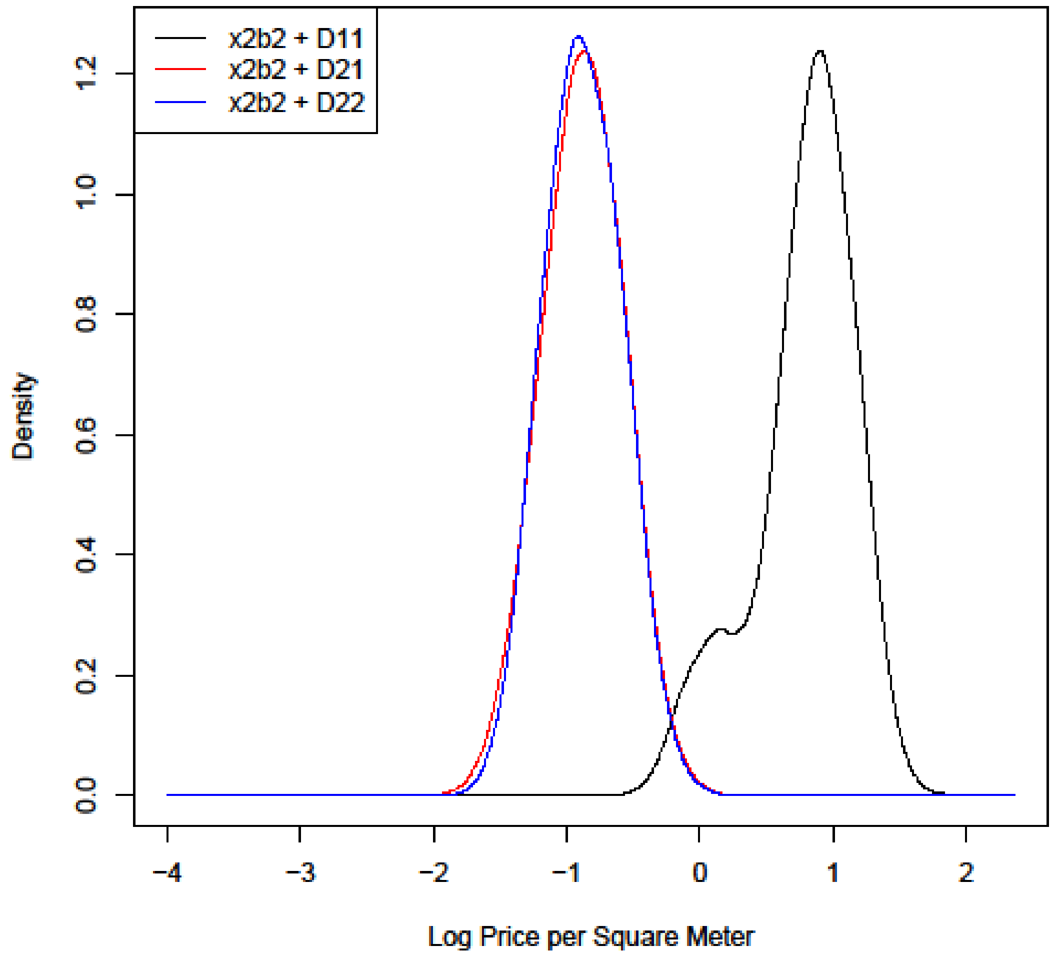
Estimate 3 locally weighted models:

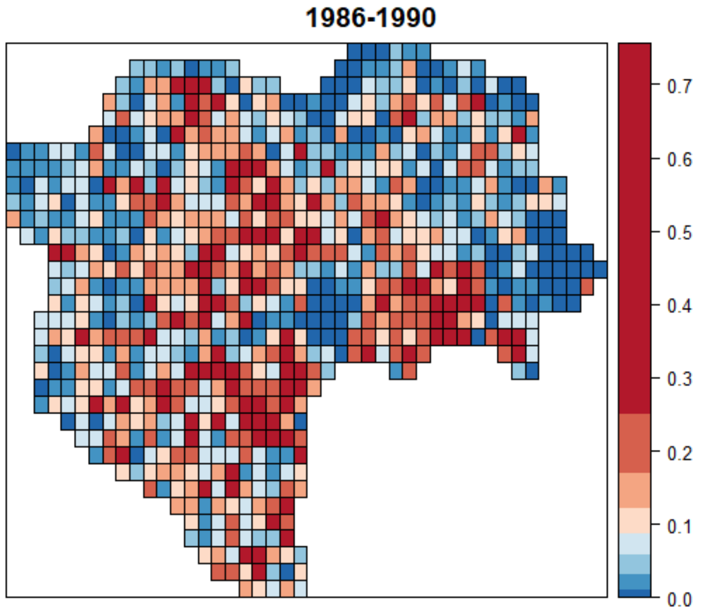
- $y_1(z_1, \tau) = x_1 \beta_1(z_1, \tau) + \delta_1(z_1, \tau)$, time 1 data and locations, B quantiles
- $y_2(z_2, \tau) = x_2 \beta_2(z_2, \tau) + \delta_2(z_2, \tau)$, time 2 locations, time 2 data, B quantiles
- $y_2(z_1, \tau) = x_2 \beta_2(z_1, \tau) + \delta_2(z_1, \tau)$, time 1 locations, time 2 data, B quantiles

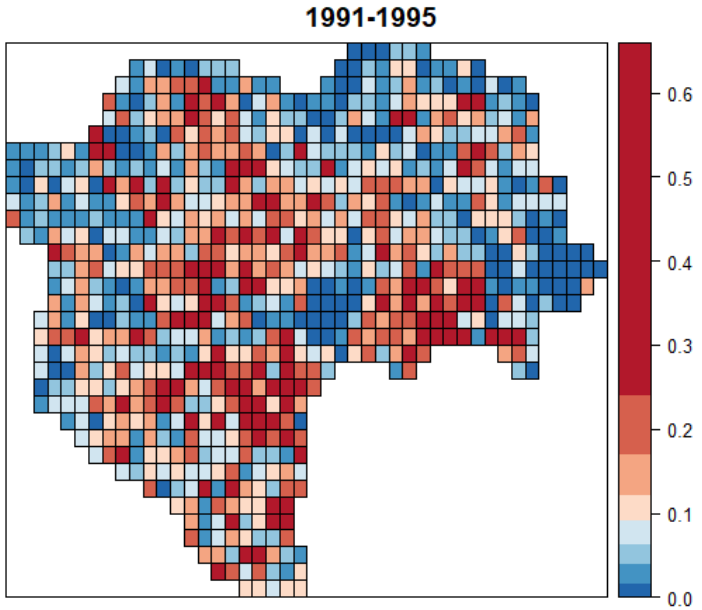
Results: $\hat{\beta}_1$, $\hat{\delta}_1$, $\hat{\beta}_2$, $\hat{\delta}_2$, $\hat{\beta}_{21}$, $\hat{\delta}_{21}$

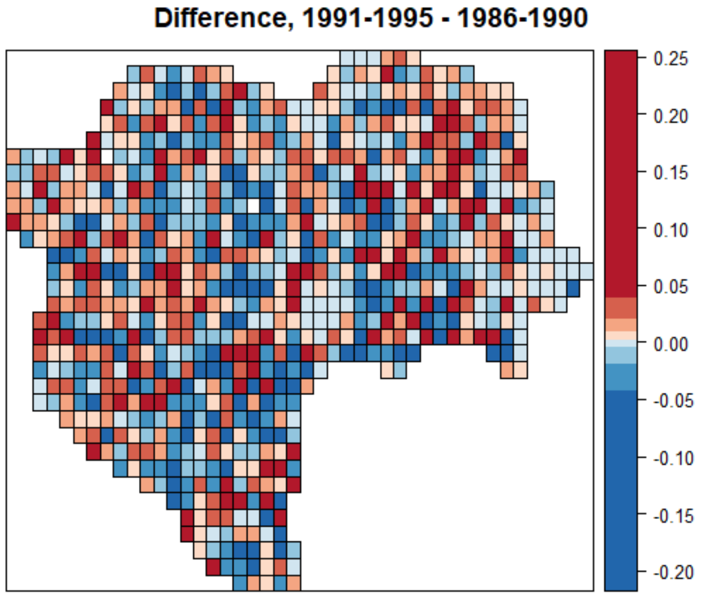
Decomposition, Tokyo: In building area, age, south view, floor variables



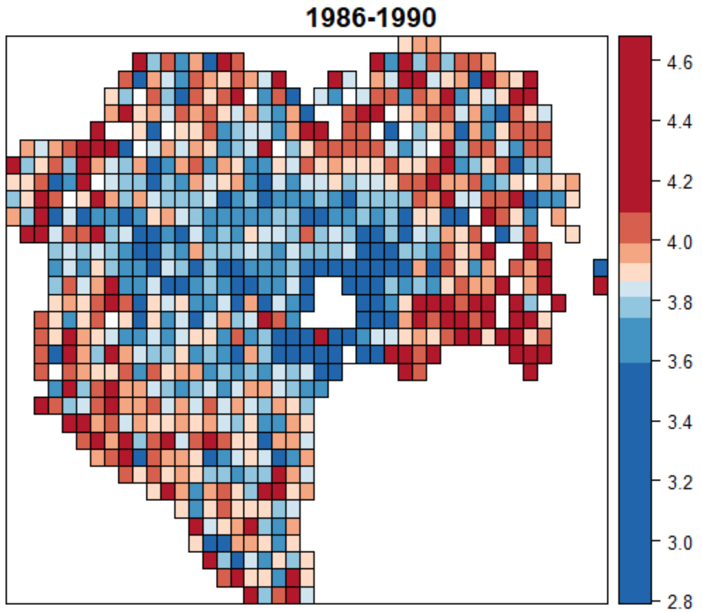


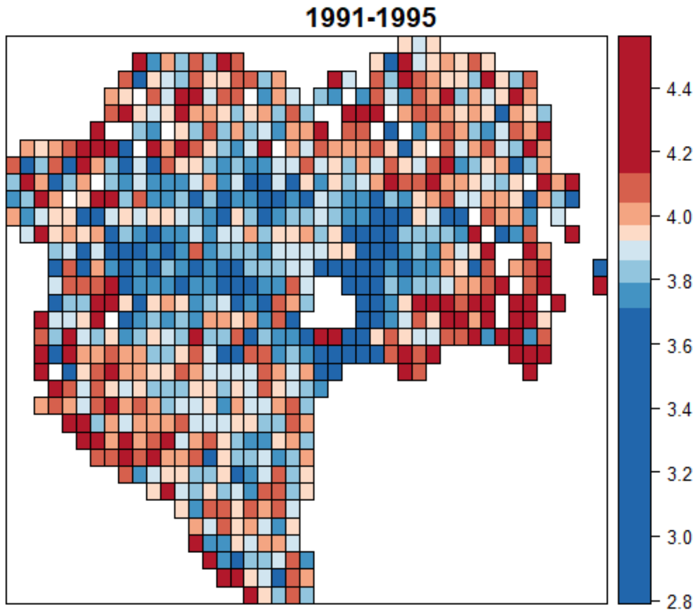


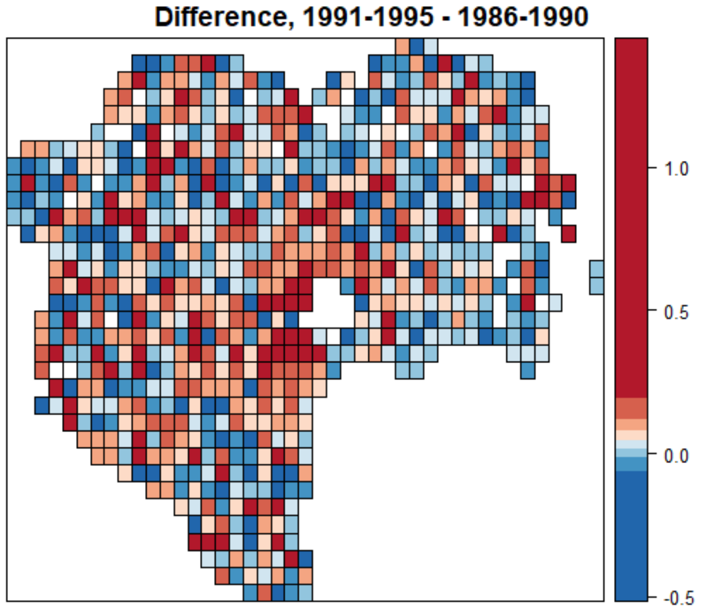




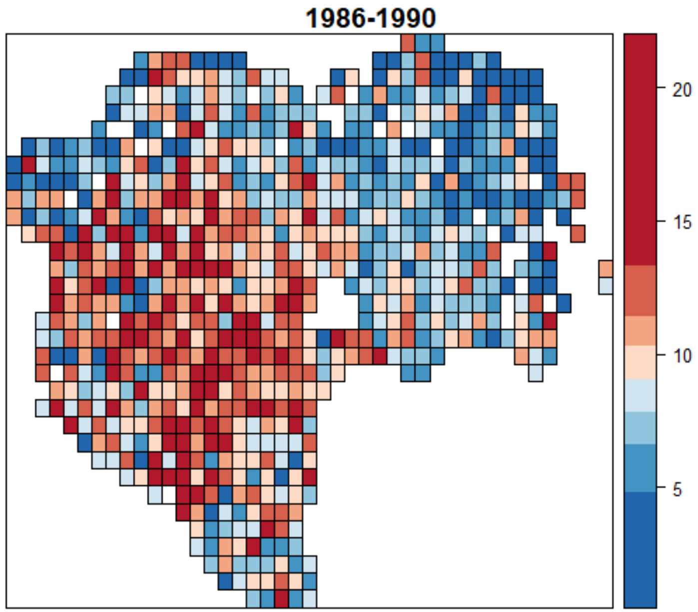
Percent of Observations by Tract

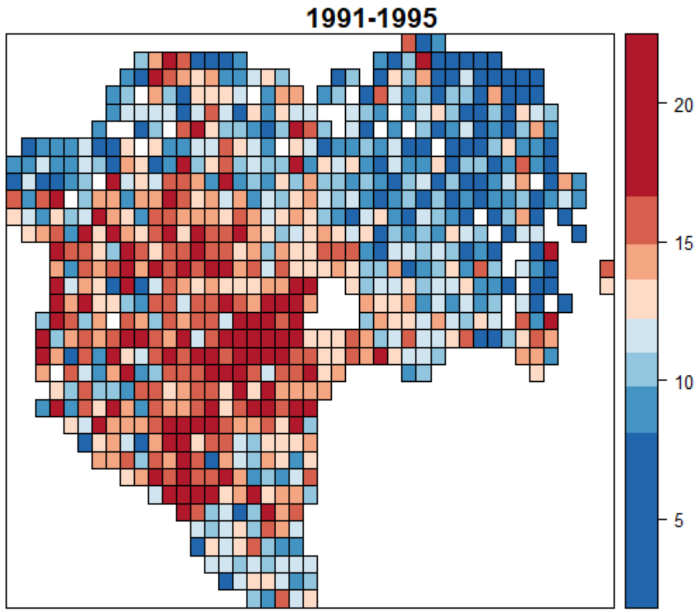


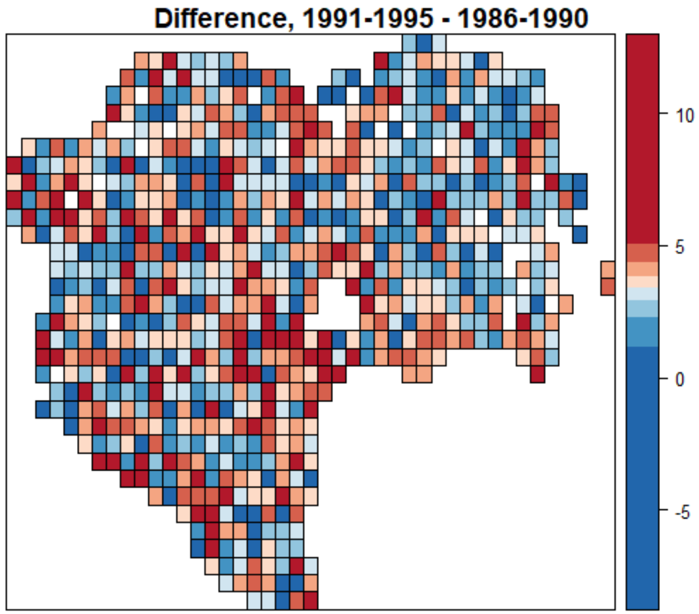




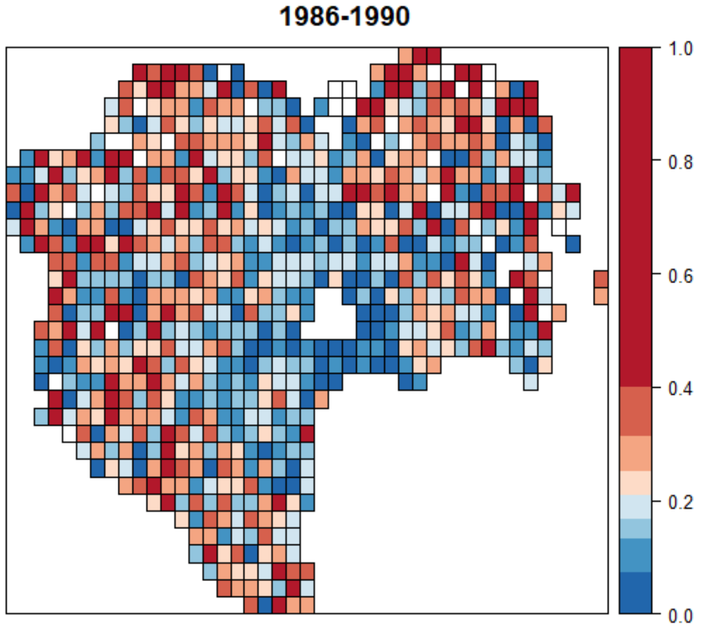
Mean Log Building Area

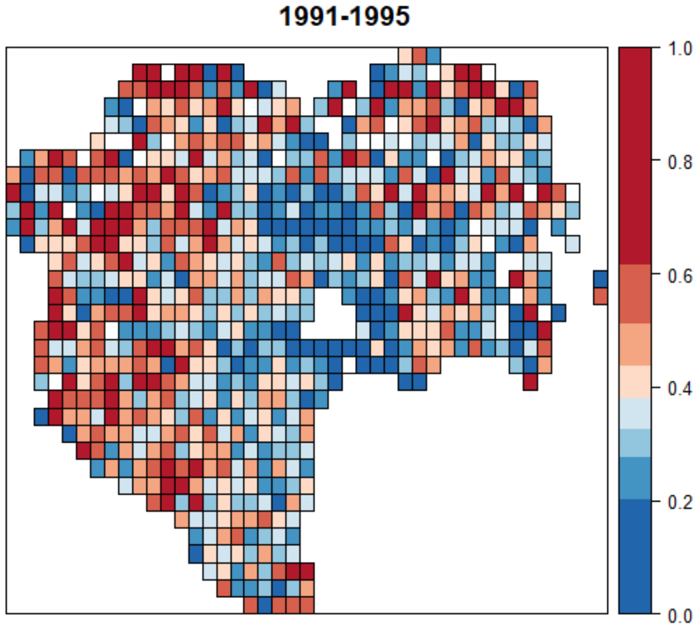

1986-1990
20
15
10
5

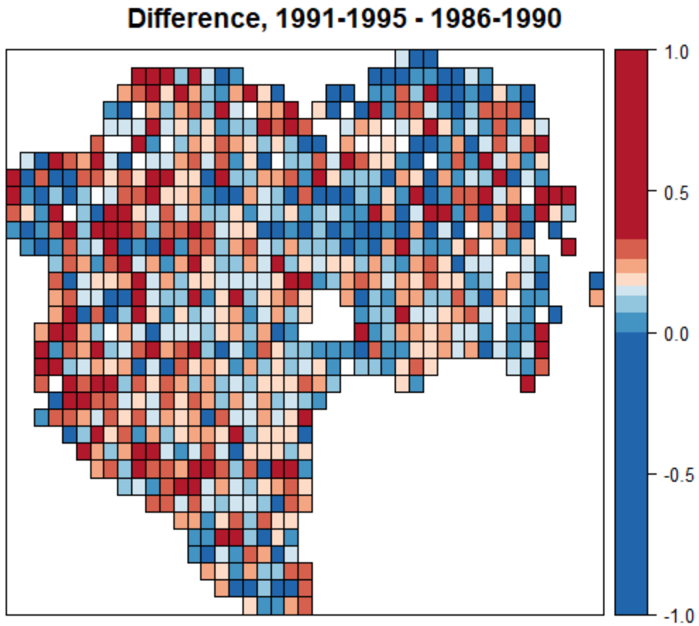

1991-1995
20
15
10
5


Difference, 1991-1995 - 1986-1990
10
5
0
-5

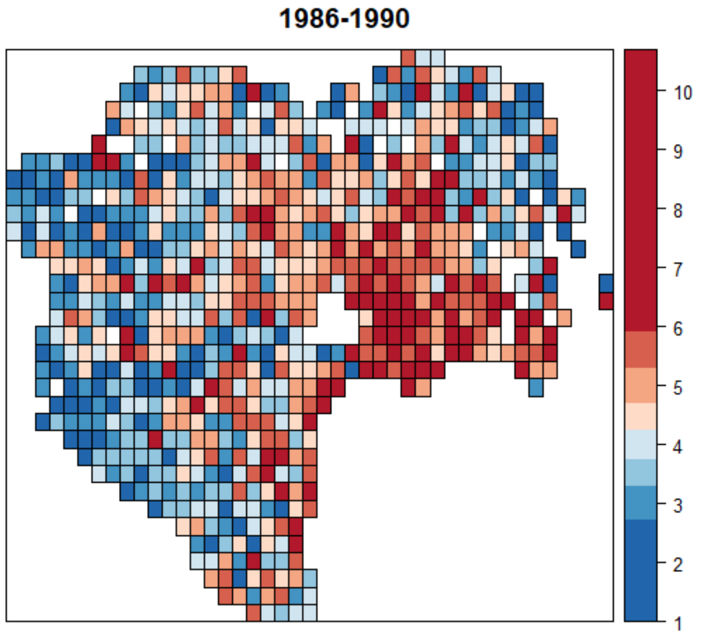
Mean Age

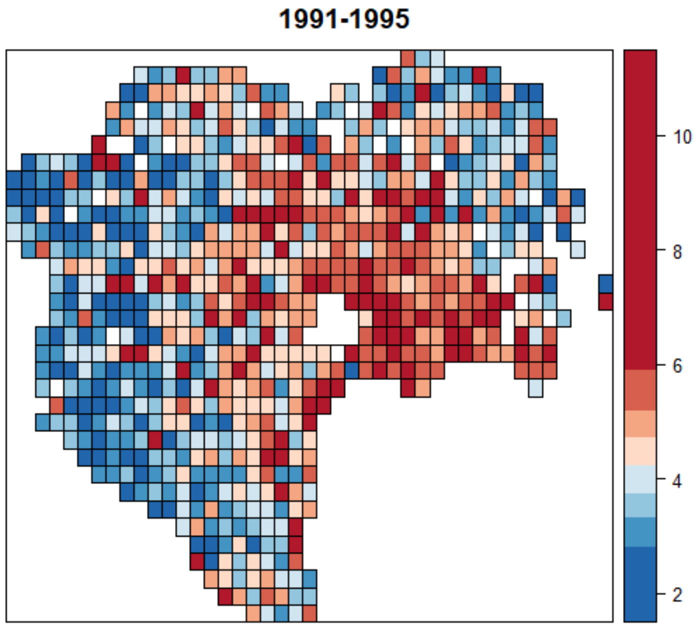

1986-1990
1.0
0.8
0.6
0.4
0.2
0.0

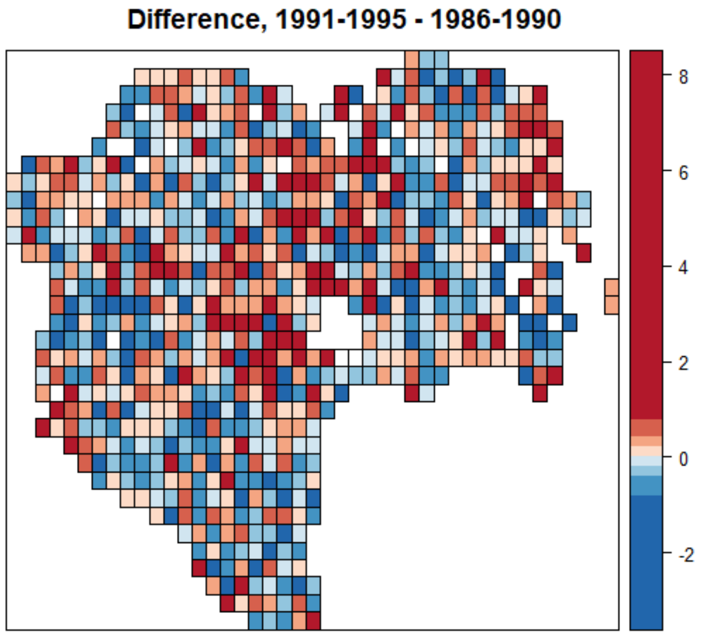

1991-1995
1.0
0.8
0.6
0.4
0.2
0.0


Difference, 1991-1995 - 1986-1990
1.0
0.5
0.0
-0.5
-1.0

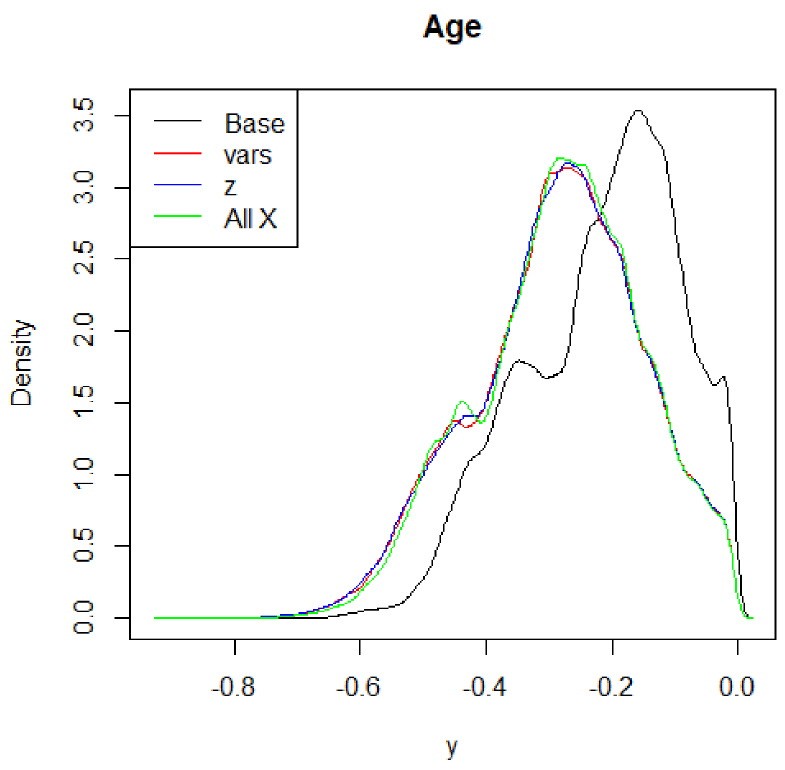
South View

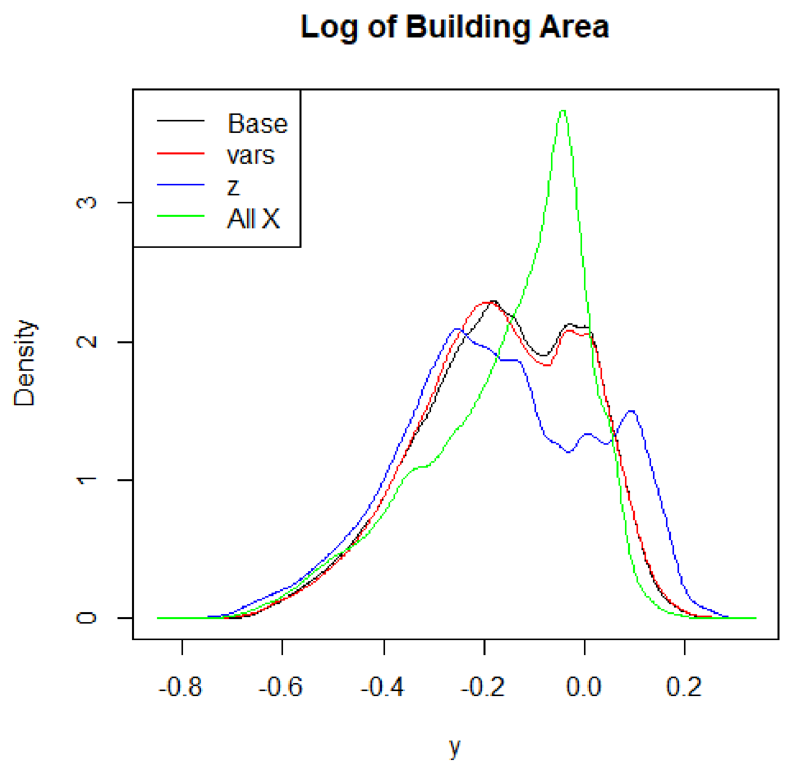

1986-1990
10
9
8
7
6
5
4
3
2
1


1991-1995
10
8
6
4
2

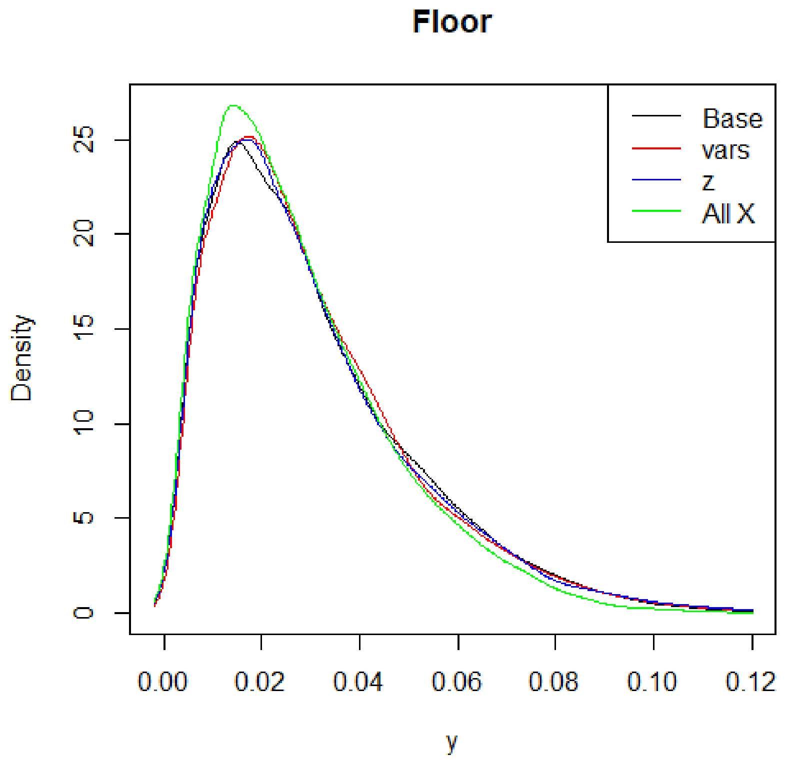

Difference, 1991-1995 - 1986-1990
8
6
4
2
0
-2

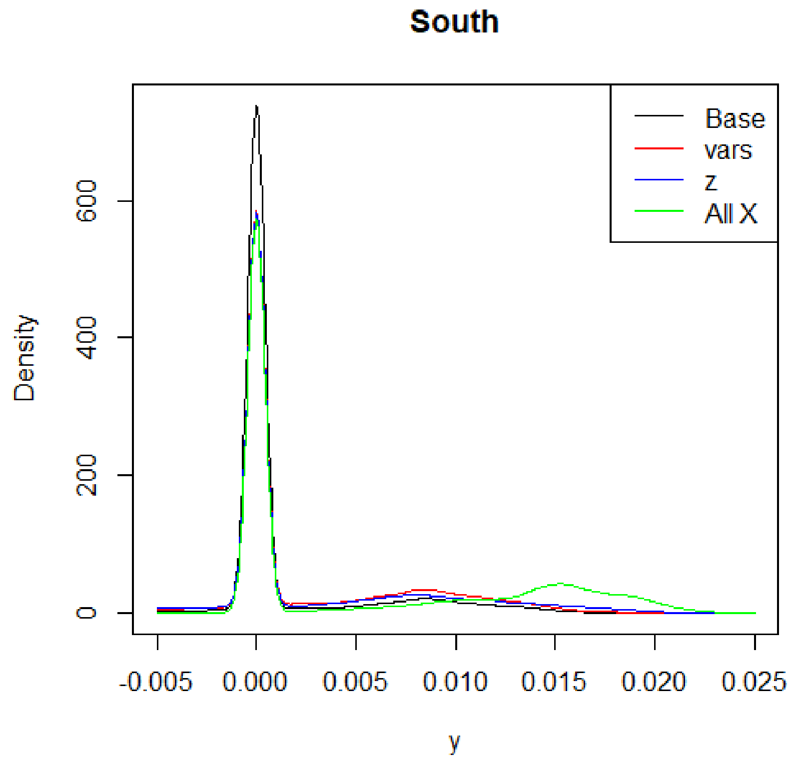
Floor



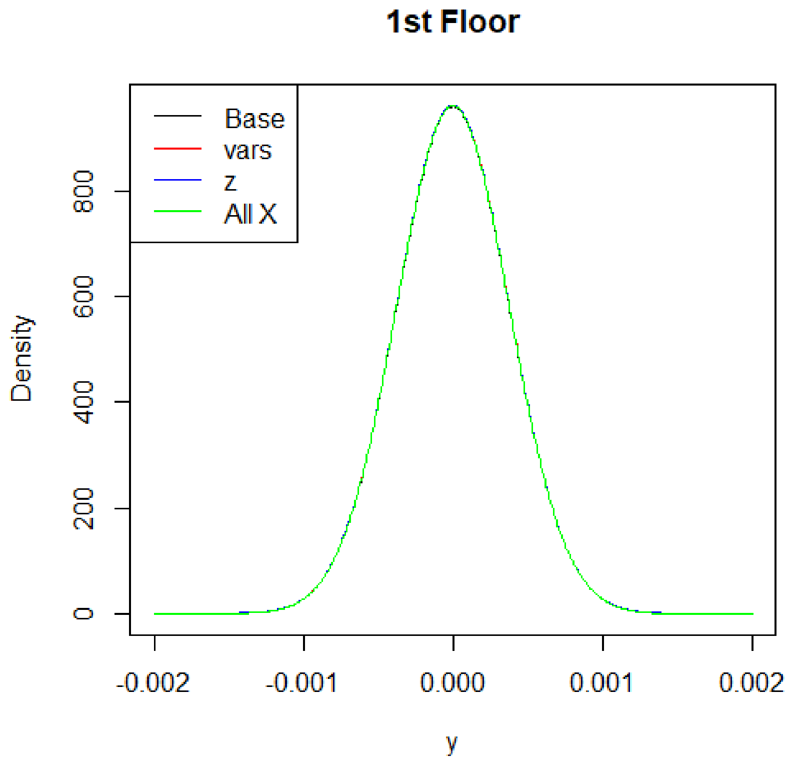


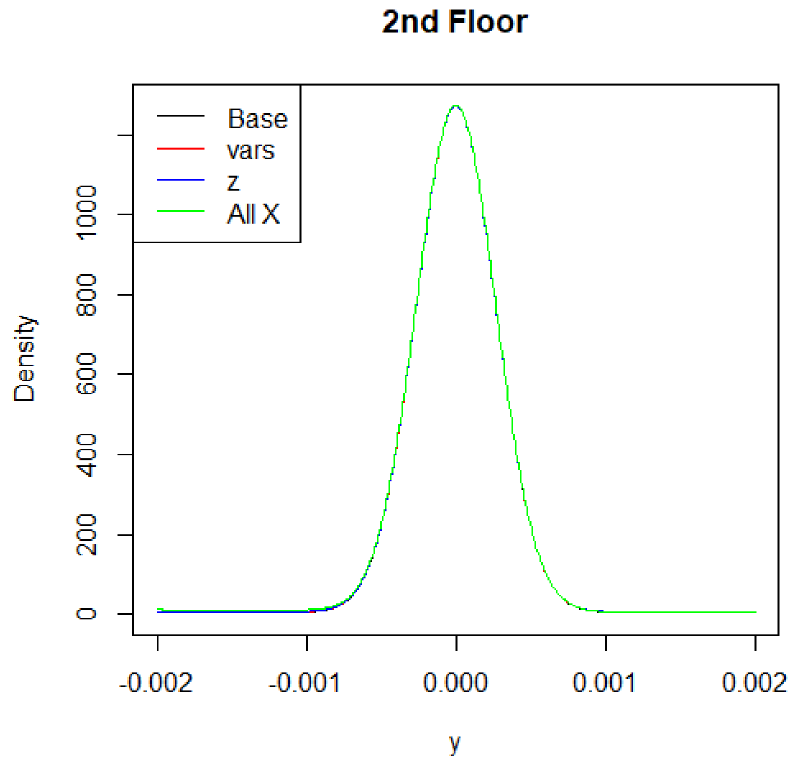
Decomposition for Individual Variables, Tokyo



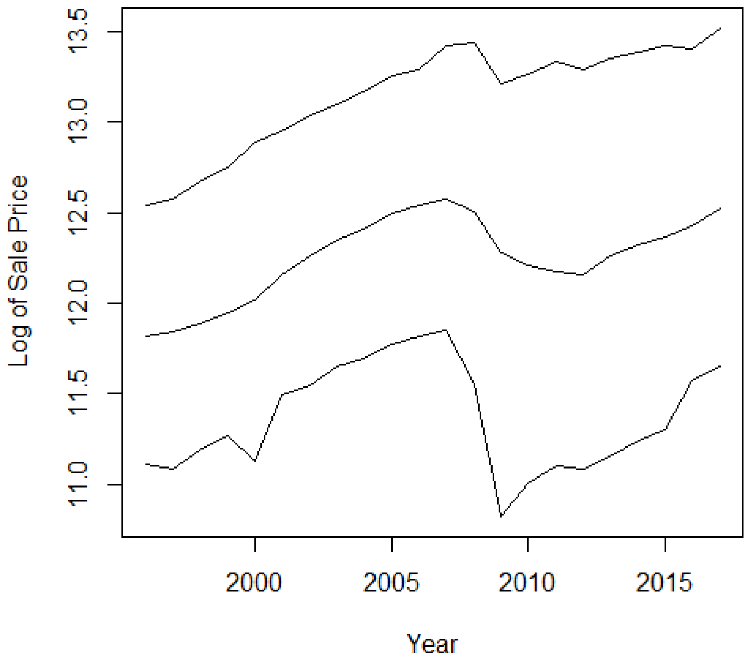


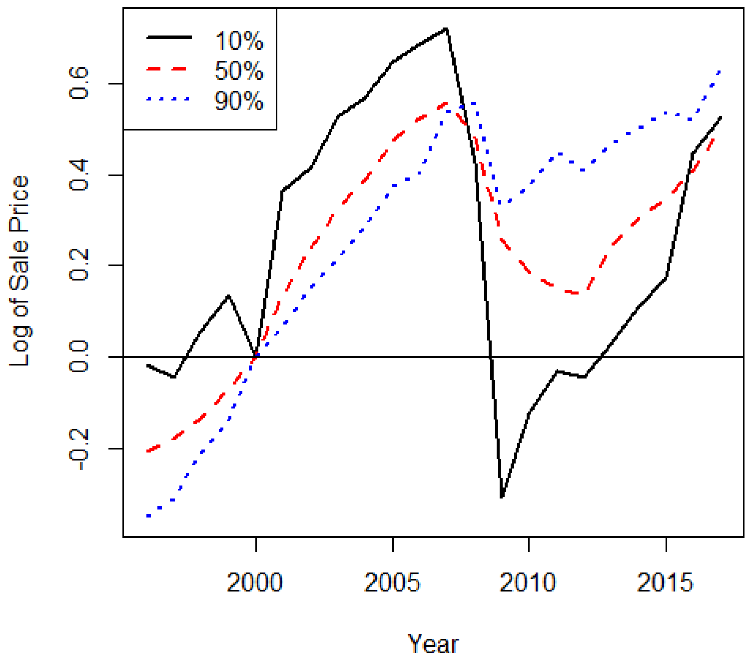
Decomposition for Individual Variables, Tokyo



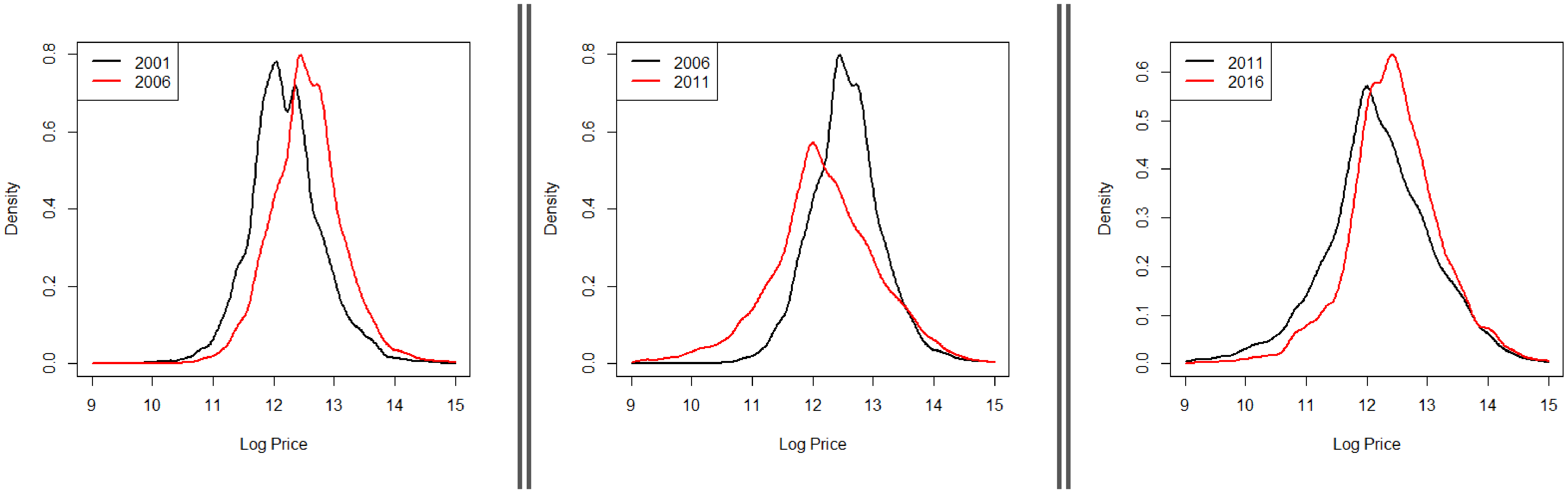


Decomposition for Individual Variables, Tokyo

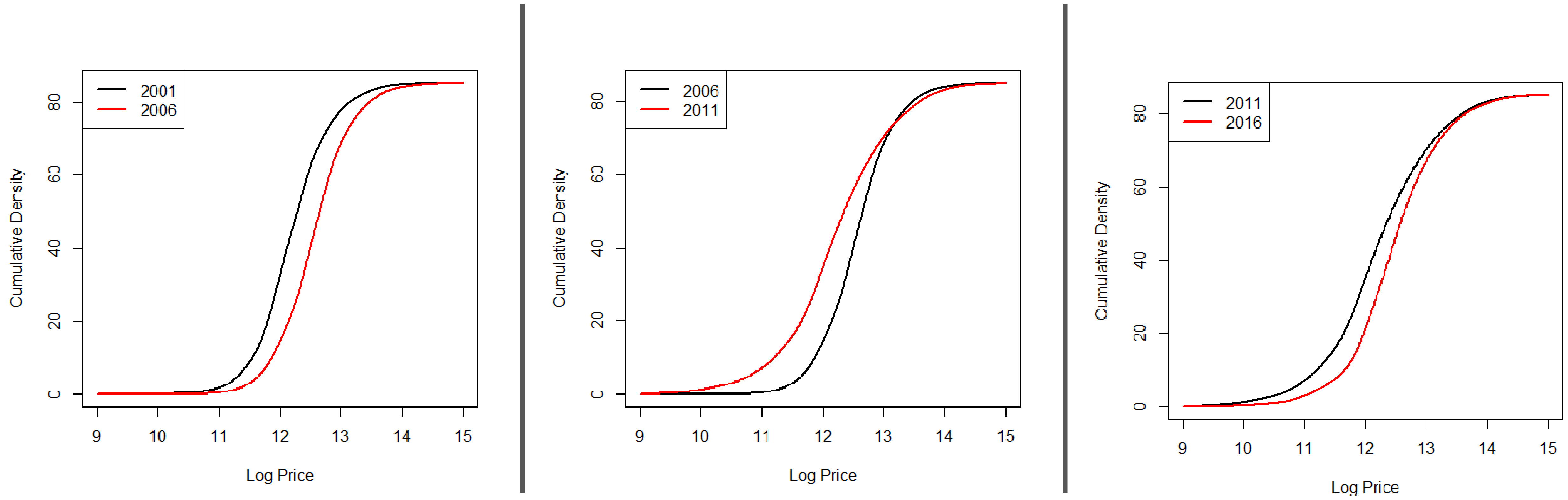




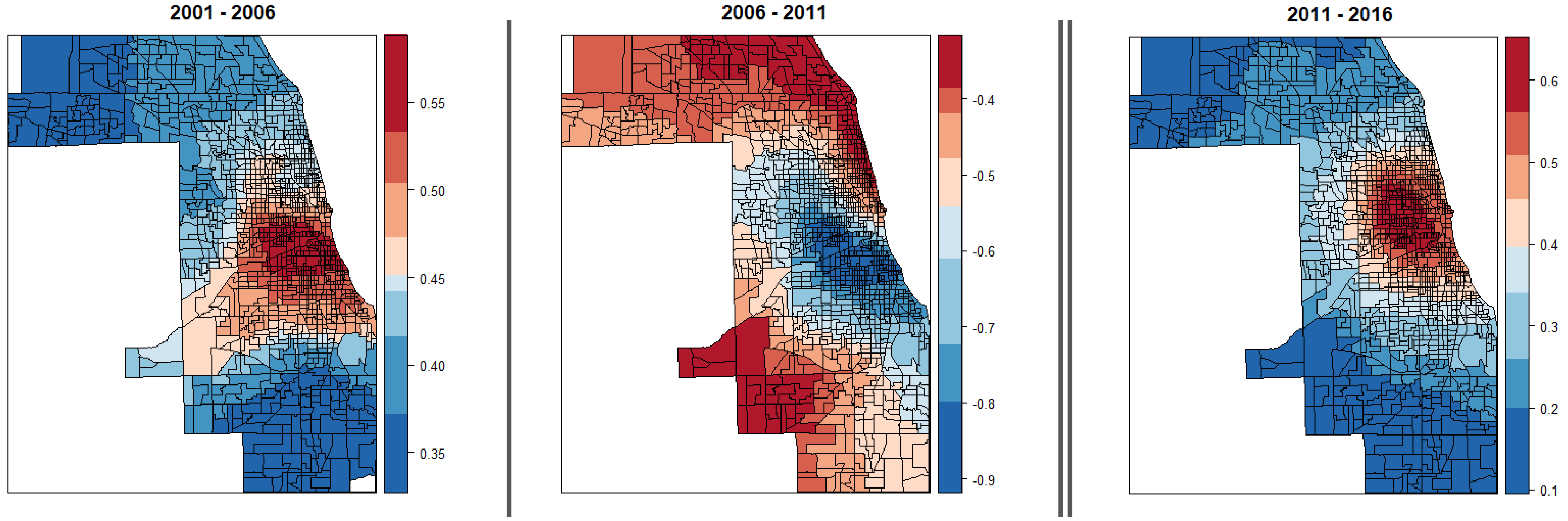
Chicago, 1996-2017



Density of Sales Prices in Chicago

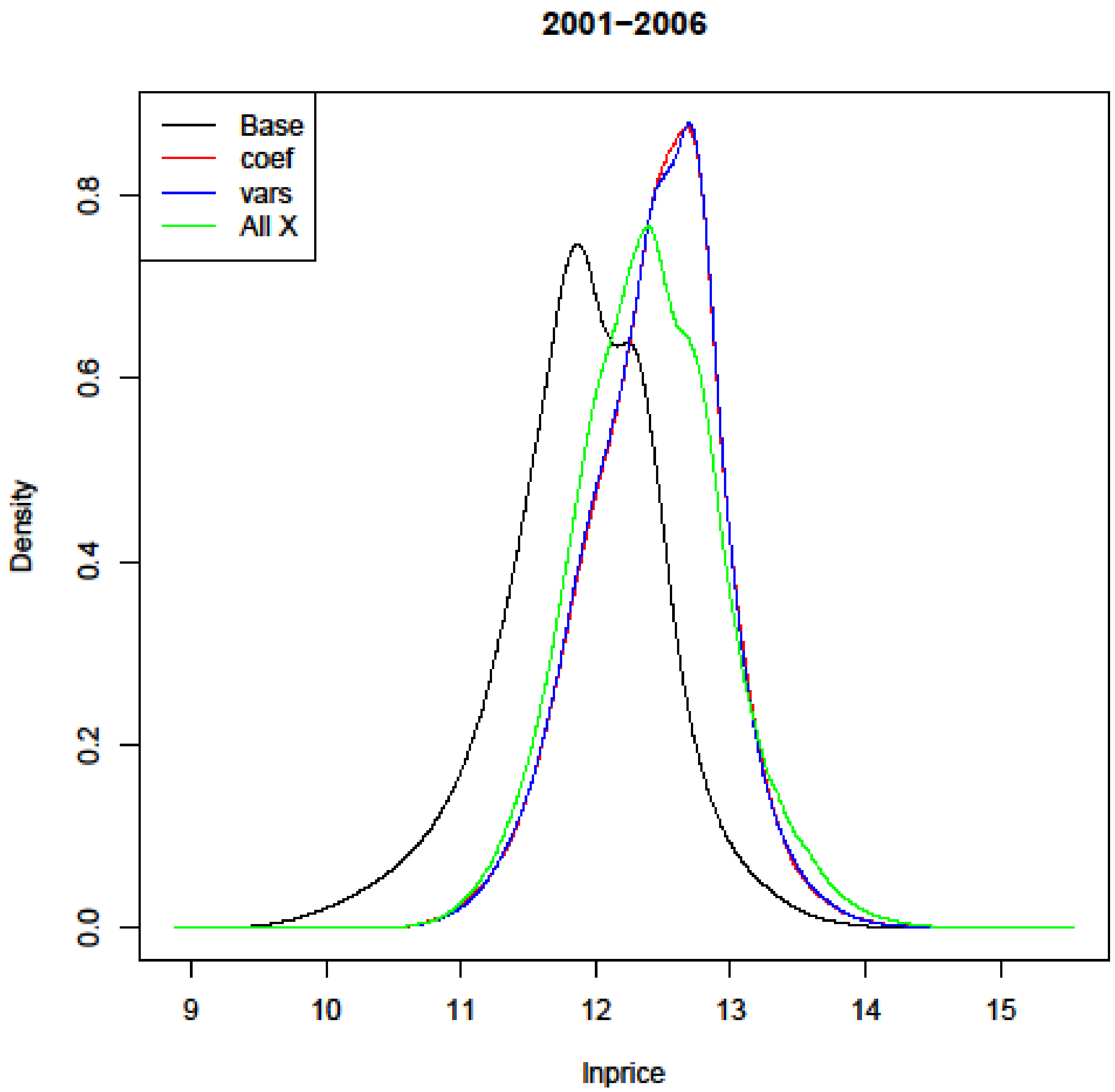


Cumulative Densities, Chicago



Appreciation Rates, Chicago

Chicago


2001−2006
Base
coef
vars
All X
Density
0.0
0.2
0.4
0.6
0.8
9
10
11
12
13
14
15
lnprice