

## Appendix

**Table A1.** Hedonic regression using full and repeat sales samples, 1998-2007

Variables	(1)		(2)	
	Estimate	SE	Estimate	SE
Distance				
Distance	0.028***	0.007	0.027*	0.010
Distance squared	-0.002**	0.001	-0.002*	0.001
Housing Characteristics	X		X	
Parking	X		X	
Fixed Effects				
ZIP code	X		X	
Time	X		X	
ZIP code linear time trend	X		X	
Observations	57,912		14,605	
Sample	Full		Repeat	

Note: The dependent variable is the natural logarithm of sales price, adjusted to 2000 dollars. Equation 1 is applied to the full and repeat sales sample. Housing characteristics include age, age squared, structural area, and dummy indicators for below average condition, having a basement, and split style. Parking is defined by the number of all-day spaces in parking lots/garages within one mile of the station. Time fixed effects are represented by sales year and quarter. Robust standard errors clustered at the ZIP code level.

\* Statistical significance level at \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

**Table A2.** Robustness check: Using percent Black instead of ZIP code

Variables	Hedonic			Repeat	
	1	2	3	4	5
Distance	0.028*** (0.007)	0.024** (0.007)	0.018* (0.007)	0.050*** (0.011)	0.040*** (0.009)
Distance squared	-0.002** (0.001)	-0.001** (0.000)	-0.002*** (0.000)	-0.002* (0.001)	-0.001+ (0.001)
Controls					
Housing characteristics	X	X	X		
Parking	X	X	X	X	X
Fixed Effects					
Housing				X	X
Time	X	X	X		
ZIP code	X	X			
ZIP code linear time trend	X			X	
Percent Black			X		
Percent Black linear time trend		X	X		X
Observations	16,511	11,243	8,956	12,132	15,980
Sample	Full	Full	Full	Repeat	Repeat
Model	Hedonic	Hedonic	Hedonic	Repeat	Repeat

Note: The dependent variable is the natural logarithm of sales price, adjusted to 2000 dollars. Each column is a separate regression. Models 1 and 4 are the main results. Percent Black is constructed by assigning the property census tract to a group based on the share of its black population with categories defined at 10 percent intervals. Housing controls include age, age squared, structural area, and dummy indicators for below average condition, having a basement, and split style. Parking is defined by the number of all-day spaces in parking lots/garages within one mile of the station. Time fixed effects are represented by sales year and quarter.

\* Statistical significance level at \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

**Table A3.** Alternative to main results, using sales price as dependent variable

Variables	(1)		(2)	
	Estimate	SE	Estimate	SE
Distance				
Distance	5,762*	2,230	8,101***	2,421
Distance squared	-336+	167	-343+	184
Housing Characteristics	X			
Parking	X		X	
Fixed Effects				
Housing			X	
Time	X		X	
ZIP code	X			
ZIP code linear time trend	X		X	
Observations	57,914		16,511	
Sample	Full		Repeat	
Model	Hedonic		Repeat	

Note: The dependent variable is sales price, adjusted to 2000 dollars. Robust standard errors clustered at the ZIP code level. Housing controls include age, age squared, structural area, and dummy indicators for below average condition, having a basement, and split style. Parking is defined by the number of all-day spaces in parking lots/garages within one mile of the station. Time fixed effects are represented by sales year and quarter.

\* Statistical significance level at \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

**Table A4.** Effect of proximity to rail on housing prices using categorical distance

Variables	(1)		(2)	
	Estimate	SE	Estimate	SE
Distance categories				
1-1.99 mi	0.000	0.018	0.109*	0.046
2-2.99 mi	0.026	0.019	0.136***	0.041
3-3.99 mi	0.056*	0.022	0.165***	0.044
4-4.99 mi	0.093**	0.028	0.252***	0.046
5-5.99 mi	0.070**	0.023	0.220***	0.047
6-6.99 mi	0.079**	0.027	0.262***	0.049
7-7.99 mi	0.091***	0.024	0.295***	0.05
8-8.99 mi	0.097***	0.025	0.303***	0.051
9-9.99 mi	0.043	0.037	0.287***	0.058
10 mi or greater	0.070	0.043	0.355***	0.058
<hr/>				
Housing Characteristics	X			
Parking	X		X	
Fixed Effects				
Housing			X	
Time	X		X	
ZIP code	X			
ZIP code linear time trend	X		X	
<hr/>				
Observations	57,912		16,511	
Sample	Full		Repeat	
Model	Hedonic		Repeat	

Note: The dependent variable is the natural logarithm of sales price, adjusted to 2000 dollars. Dropped distance category is less than one mile. Model 1 presents results for the standard hedonic model using the full sample with robust standard errors clustered at the ZIP code level. Model 2 presents results for the repeat sales model using the repeat sales sample with robust clustered standard errors.

\* Statistical significance level at \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

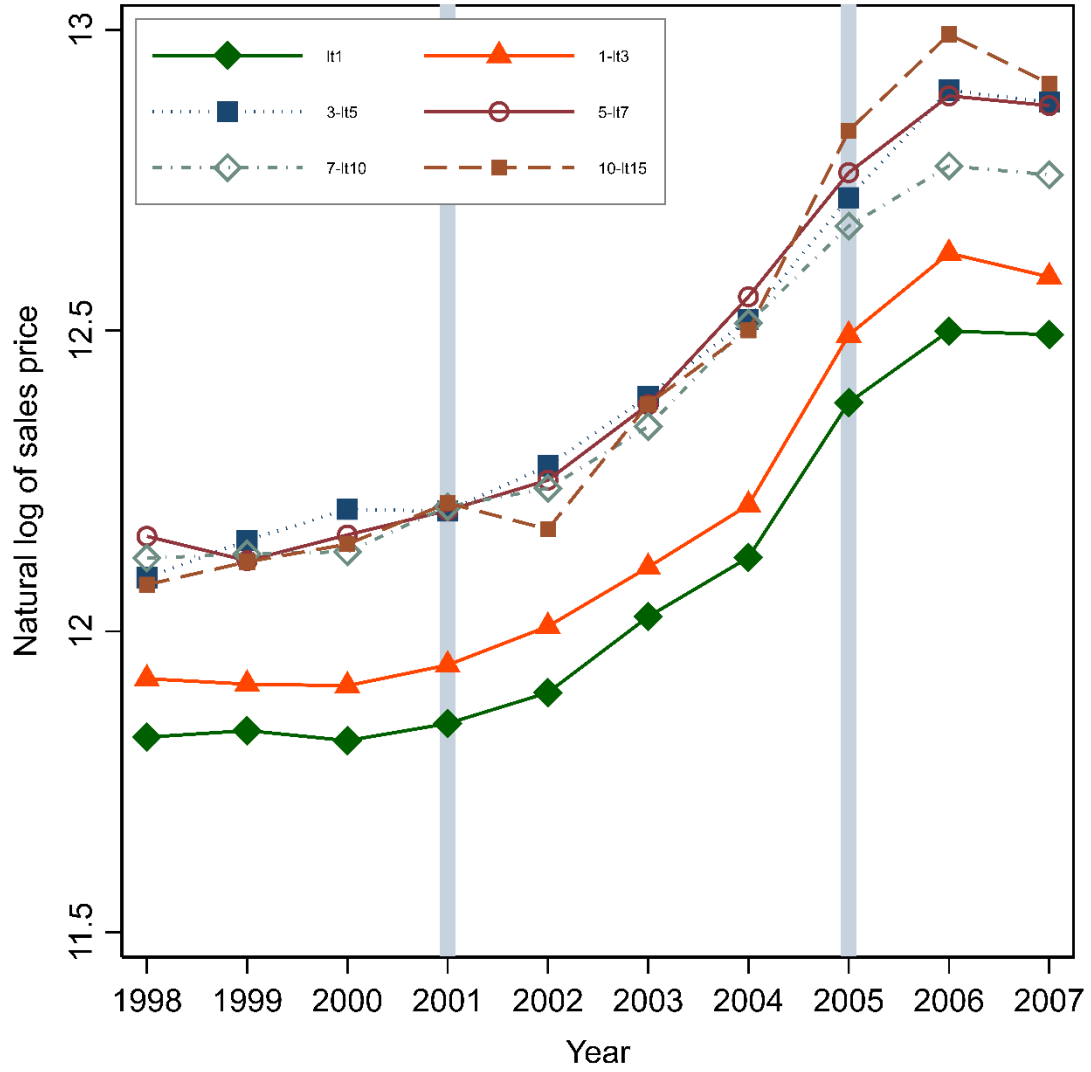


Figure A1. Average sales price by distance over time, full sample

Note: Sales price in 2000 dollars. Blue vertical lines represent opening of new stations.

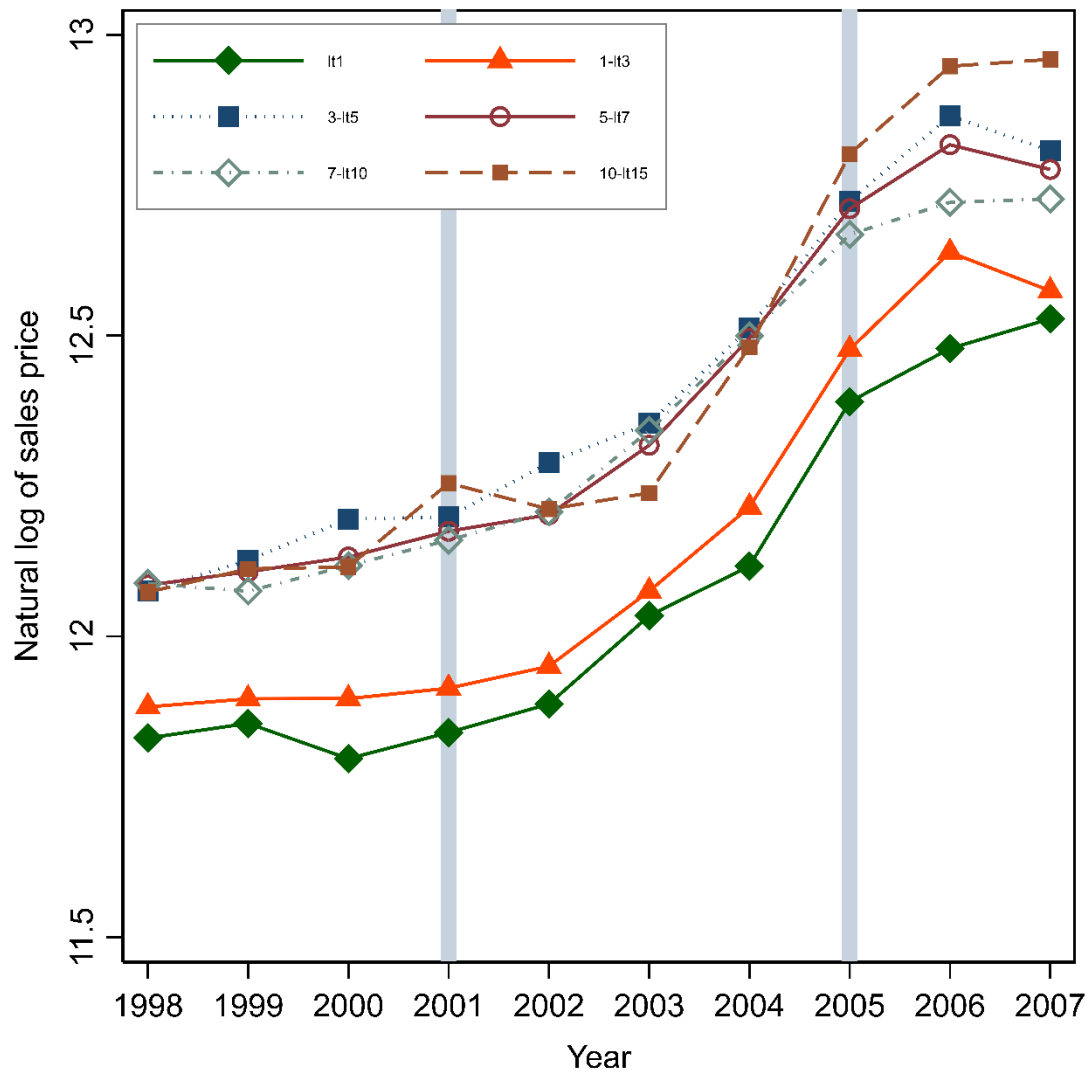


Figure A2. Average sales price by distance over time, repeat sales sample

Note: Sales price in 2000 dollars. Blue vertical lines represent opening of new stations. lt=less than.

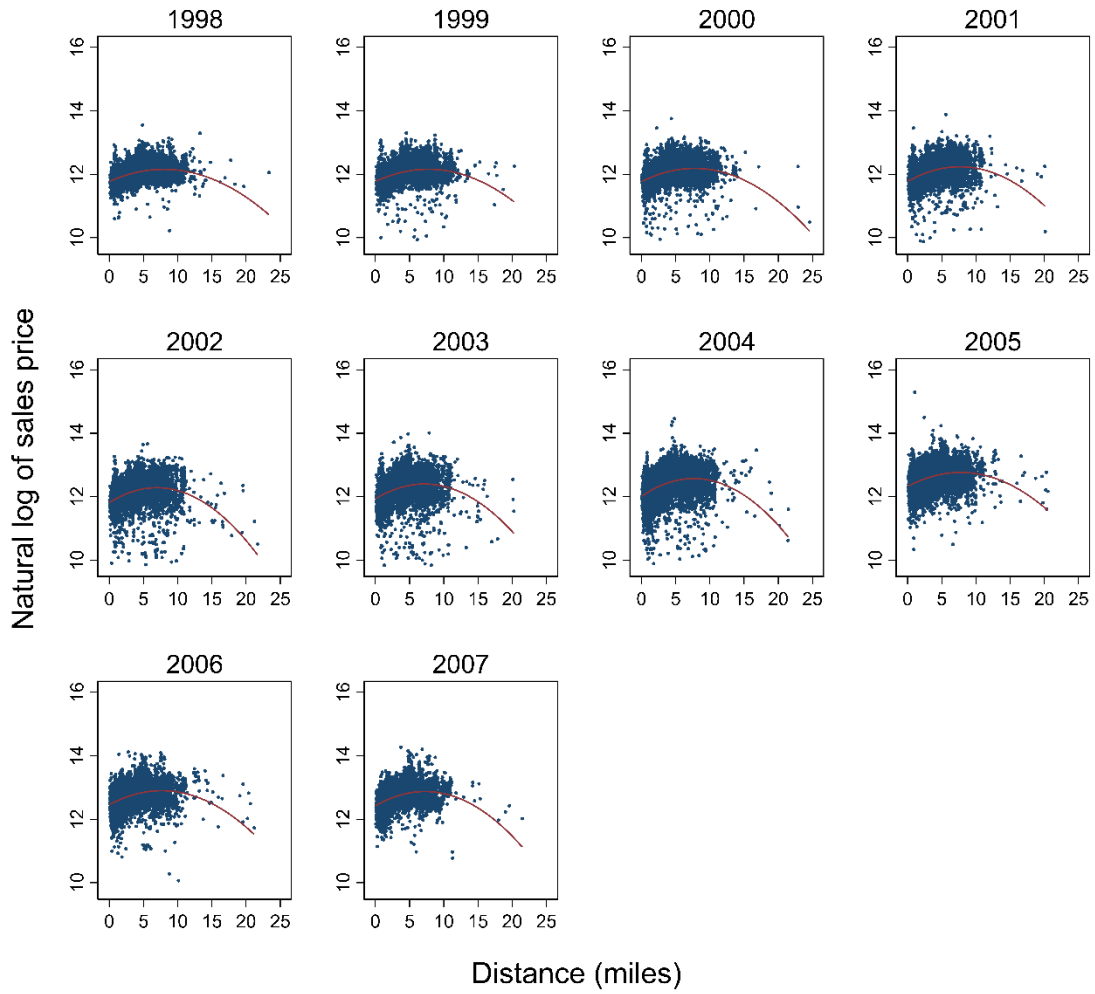
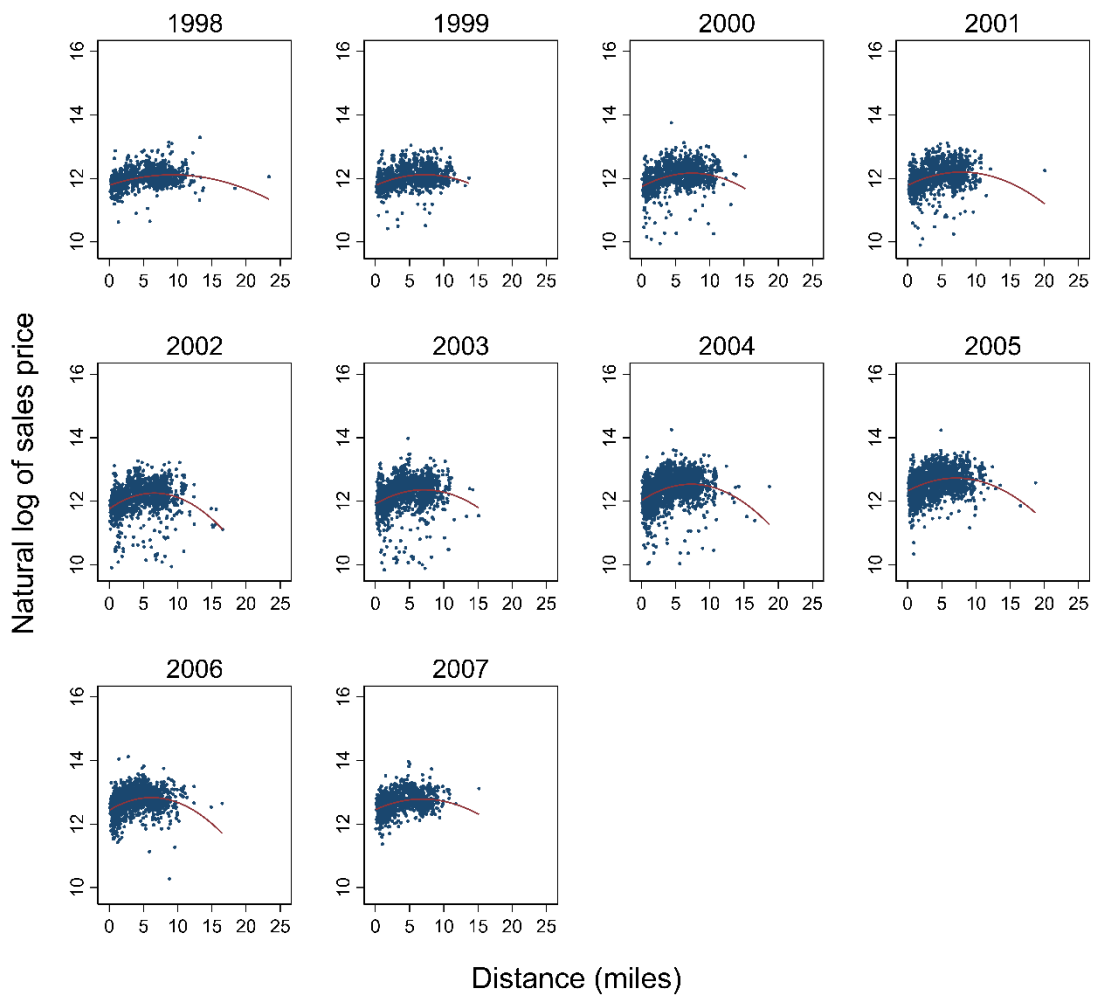


Figure A3. Sales transaction by distance by year, full sample

Note: Sales price in 2000 dollars.



**Figure A4.** Sales transaction by distance by year, repeat sales sample

Note: Sales price in 2000 dollars.



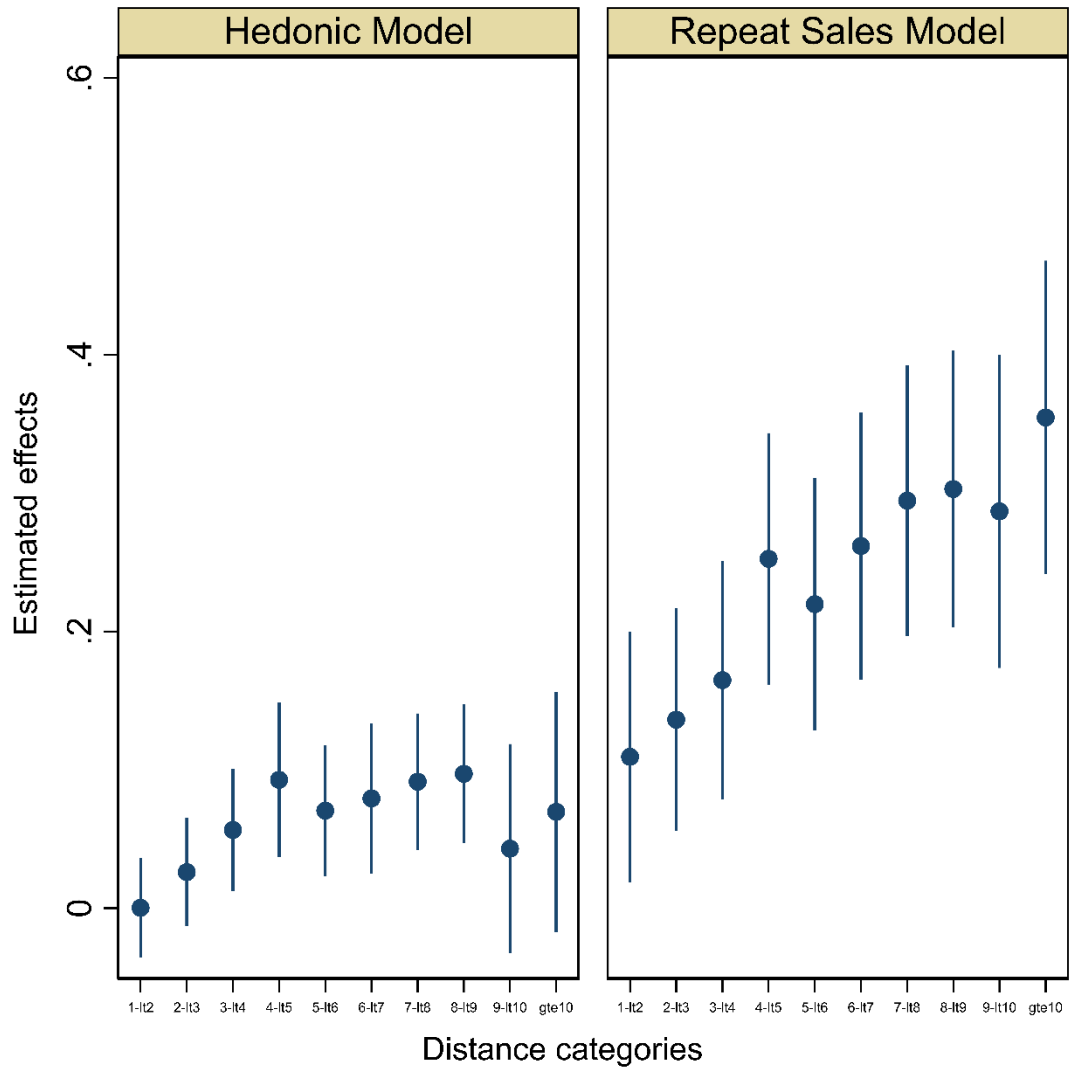


Figure A5. Regression coefficients of hedonic and repeat sales models using distance categories

Note: gte = 10 miles or greater. lt=less than.