2012 V40 4: pp. 745–773
DOI: 10.1111/j.1540-6229.2012.00343.x
REAL ESTATE
ECONOMICS

Repeat Sales as a Matching Estimator

Daniel P. McMillen*

The most common approaches for constructing house price indices—hedonic price functions and the repeat sales estimator—focus on changes over time in mean prices. Though the hedonic approach is less wasteful of data than the repeat sales estimator, it relies on an accurate specification of the underlying econometric model. I suggest using a matching estimator as an alternative to the hedonic and repeat sales approaches. Like the repeat sales approach, a matching estimator uses pairs of sales from different dates to estimate the mean difference in sales prices over time. The matching approach preserves much larger sample sizes than the repeat sales estimator while requiring less preimposed structure than the hedonic approach. The matching approach makes it easy to characterize changes in the full distribution of house prices.

The most common approaches for constructing house price indices—hedonic price functions and the repeat sales estimator—focus on changes over time in mean prices. Both approaches are designed to measure the expected sale price of a home after controlling for features of the structure and location. Hedonic price functions control directly for structural and locational characteristics by including them as explanatory variables in the estimated regression. The repeat sales approach controls for these variables more indirectly by restricting the analysis to properties that sold at least twice during the sample period. Time-invariant explanatory variables whose coefficients are constant over time drop out of the repeat sales regression because the estimating equation is based on changes in sales prices.

Both the hedonic approach and the repeat sales estimator have potential problems. Missing variables or functional form misspecification can produce biased estimates of hedonic price function coefficients. Repeat sales estimates may also be subject to bias if the explanatory variables for sales prices are not constant over time or if their coefficients change—a serious problem in places where homes are undergoing extensive renovations and some neighborhoods enjoy higher appreciation rates than others. Moreover, the repeat sales estimator suffers a serious loss in the number of observations available to estimate the index as all homes selling only once are dropped from the sample, and it

*Department of Economics, University of Illinois at Urbana-Champaign, Urbana, IL 61801 or mcmillen@illinois.edu.

is not clear that homes that have sold more than once are representative of the overall market. The potential sensitivity of the repeat sales approach to renovations, changing coefficients, sample size and sample selection is highlighted in such studies as Cannaday, Munneke and Yang (2005), Case *et al.* (2006), Case, Pollakowski and Wachter (1997), Case and Quigley (1991), Clapp, Giaccotto and Tirtiroglu (1991), Englund, Quigley and Redfearn (1999), Gatzlaff and Haurin (1997), McMillen (2003) and McMillen and Thorsnes (2006).

As mean-based estimators, both the hedonic and repeat sales approaches focus on a single measure of central tendency for the sale price distribution. The mean can be misleading when sales prices do not appreciate equally throughout the distribution. Using a quantile regression approach, McMillen (2008) finds that the distribution of sales prices in Chicago became less skewed between 1995 and 2005 because the high-price portion of the sale price distribution shifted further to the right than was the case for the lower end of the distribution. Since the sale price distribution was heavily skewed toward lower priced homes in 1995, the higher rate of appreciation among high-priced homes produced a more symmetric distribution of prices in 2005. By focusing exclusively on average prices, the hedonic and repeat sales approaches cannot account for such differences in appreciation rates over the range of the sale price distribution.

In this article I propose a matching estimator analog to the repeat sales approach that results in a particularly simple and useful method for characterizing changes across the entire distribution of sales prices. Beginning with Rosenbaum and Rubin (1983, 1984), the idea behind a matching estimator is to estimate average treatment effects by pairing each observation receiving treatment with a similar observation from the sample that did not receive the treatment. If the matching estimator succeeds in balancing the distributions of the covariates across the treated and untreated samples, a simple difference in means test provides an unbiased measure of the average treatment effect. The matching approach can also be used to calculate changes across other percentiles of the distribution. For example, it is just as straightforward to compare the 10th percentile of outcomes for the treated and untreated samples as to compare the averages.[1]

The repeat sales approach is a special case of a matching estimator in which houses are matched to themselves across time periods. For example, a home that sold in both 2000 and 2010 provides direct information on the appreciation rate across these two time periods. Treating 2010 as the treatment and 2000 as

[1]While this article was under review, I learned of a paper by Longford (2009) that also proposes uses a propensity score matching procedure to construct price indexes. He focuses on mean values rather than the full distribution of prices. The approach developed here was applied in Deng, McMillen, and Sing (2012) in an analysis of the housing market in Singapore. Though the current paper was written first, publication lags led the Singapore paper to appear first.

the base, a comparison of the sales prices for homes that sold in these two years provides a direct measure of the change in the house price distribution from the beginning to the end of the decade. The distinction between the treatment and control samples is less clear when sales are drawn from many years. However, the point of an index is to compare sales prices over time to a base period. The logical control date is thus the base time period, while each subsequent time becomes a separate form of treatment: how does the fact that a home sells at time $t + s$ rather than time t affect its sales price?

The matching estimator produces a very simple estimation procedure. It begins by matching each observation from the base period to a similar observation from each subsequent period. The match can be based on characteristics of the structure and location, which in the limit leads to the standard repeat sales estimator. Alternatively, matching can be based on a measure of closeness, with similar homes matched across similar neighborhoods. A particularly simple way of matching observations is to use the propensity score from probit or logit estimates of the probability that a home sold at time $t + s$ rather than in the base time t. As Wang and Zorn (1997) note, when the number of observations does not vary across time periods, the standard repeat sales estimator simplifies to a comparison of average sales prices across times for the repeat sales sample. Instead of restricting the sample to homes that sell at least twice, the matching estimator uses much larger samples of similar homes to compare prices across various points of the price distribution. After eliminating unusual homes—those without a close match—averages across many observations produce reliable estimates of constant-quality appreciation rates without resorting to the extreme of matching each home only to itself.

As compared with the repeat sales estimator, the more general matching approach has much larger sample sizes in each time period. Larger sample sizes are likely to lead to gains in precision while reducing the probability that unusual observations will unduly affect the results. The increase in sample size is a particular advantage when attempting to construct a price index for small geographic areas with relatively few repeat sales.

Although the matching estimator's closest analog is the repeat sales estimator, it also has advantages over the hedonic approach. It is much easier to estimate appreciation rates throughout the sale price distribution using the matching estimator than is the case with the mean-based hedonic approach. The matching approach is also less prone to functional form misspecification and to the effects of extreme observations. However, the hedonic approach has a significant potential advantage over the simple version of the matching approach in that it directly controls for structural and locational characteristics, which may lead to better measures of quality-adjusted price differences when the matching estimator does not succeed in perfectly balancing distributions across time. In this case, the best option may be to construct the index by estimating hedonic price functions using the matched sample: using the matched sample reduces the effect of outlier observations and functional form misspecification, while the hedonic regression helps to control for the effect of the observed covariates.

The rest of the article is organized as follows. I first provide a review of the standard procedures for constructing house price indices. The matching approach is introduced in the third section. The fourth section summarizes the hedonic, repeat sales and matched sample data sets. I then compare the approaches using sales of homes in Chicago for 1993–2008. The sixth section presents results for using some alternative matching procedures. The final section concludes the article.

Standard Price Indices

Both the hedonic and the repeat sale price index can be derived from the following estimating equation:

$$y_{it} = \sum_{t=1}^{T} \delta_t D_{it} + \beta_t' X_{it} + \lambda_t' Z_{it} + u_{it}. \tag{1}$$

In Equation (1), y_{it} represents the natural logarithm of the sale price of house i at time t ($t = 1, \ldots, T$), D_{it} is a variable indicating the house sold at time t, X_{it} is a set of observed characteristics of the structure and location, Z_{it} is a set of unobserved characteristics that also influence the sales price, and u_{it} is a random error term. The parameters to be estimated comprise the vectors $\delta = (\delta_1 \cdots \delta_T)'$, β and λ. Equation (1) is a general formulation that encompasses all commonly used procedures for constructing sale price indices. Special cases are produced by altering the assumptions regarding which variables and coefficients vary over time.

Hedonic price indices are generally based on one of two specifications. The more common specification is a special case of Equation (1) in which the coefficients for X and Z are constrained not to vary over time:

$$y_{it} = \sum_{t=1}^{T} \delta_t D_{it} + \beta' X_{it} + \lambda' Z_{it} + u_{it}. \tag{2}$$

In this specification, the price index is simply the series of estimated values $\hat{\delta}_1 \cdots \hat{\delta}_T$ from a regression of y on X and, if it were observable, Z. Less commonly, Equation (1) serves as the basis for a series of estimating equations, with a separate regression estimated for each time period. In this case, the price index is either the series of average predicted values from each regression or the set of predictions for target values of X and, when observable, Z:

$$\hat{y}_t = \hat{\delta}_t + \beta_t' \bar{X} + \lambda_t' \bar{Z}, \tag{3}$$

where $\bar{X}$ and $\bar{Z}$ represent either the average values for X and Z or a set of arbitrary target values. However, if $\bar{X}$ and $\bar{Z}$ represent the average values of X and Z at time t, it is important to bear in mind an elementary feature of linear regression: since the regression line goes through the set of means, $\hat{\delta}_t + \beta_t' \bar{X}_t + \lambda_t' \bar{Z}_t$ simply equals the average value of y at time t, $\bar{y}_t$. Thus, there is no need to estimate a regression if the price index is based on period-by-period predicted values at the sample average values of the explanatory variables at time t. Good examples of the hedonic approach include Kiel and Zabel (1997), Mark and Goldberg (1984), Palmquist (1980) and Thibodeau (1989).

The advantages of the hedonic approach are that (1) it accomplishes the objective of producing a quality-adjusted index by including controls for observable characteristics of the house and location and (2) it takes advantage of the large sample sizes available for home sales over time. The primary problem with the approach is the need to specify a functional form and the possibility that omitted variables may affect the results. Functional form misspecification can also be viewed as a form of omitted variables. For example, if the correct functional form is quadratic in a single variable x but the equation is specified to be linear, then x^2 enters the vector of missing variables, Z.

Since the price index is generally constructed as the series of estimates $\hat{\delta}_1 \cdots \hat{\delta}_T$ from Equation (2), the key question is whether Z is correlated with the set of time of sale indicator variables, D. Thus, if the sample of sales in a time period tends to be composed of unusually high- or low-quality homes—*i.e.*, homes with high or low average values for Z—then the effect of the missing variables will be to bias the estimated value of δ_t upward or downward. Since the estimated values of δ are biased if the missing variables Z are correlated with X, the predicted values will also be biased when the index is constructed for an arbitrary set of target values for x. Simple sample averages will also provide a misleading measure of quality-controlled price changes because unusually high or low values of the explanatory variables will lead to unusually high or low values for the sample mean of sales price. This issue is generally referred to as a problem of "sample selection" in the literature, although the real problem is missing variables that are correlated with the time of sale.

The repeat sales approach attempts to control for the effects of missing variables by restricting the analysis to the set of homes that sold at least twice during the sample period (Bailey, Muth and Nourse 1963, Case and Shiller 1987, 1989).

Using Equation (1), the change in the log of sale price for a home that sold at time t and an earlier time s is:

$$y_{it} - y_{is} = (\delta_t D_{it} - \delta_s D_{is}) + (\beta_t' X_{it} - \beta_s' X_{is}) + (\lambda_t' Z_{it} - \lambda_s' Z_{is}) + (u_{it} - u_{it}). \tag{4}$$

If X, Z, β and λ are all constant over time, then Equation (4) reduces to a very simple estimating equation:

$$y_{it} - y_{is} = \delta_t D_{it} - \delta_s D_{is} + u_{it} - u_{it}. \tag{5}$$

The problem is that housing characteristics and their coefficients may change over time. If so, Equation (4) shows that the standard repeat sales estimator, like the hedonic approach, is subject to a missing variable problem. The repeat sales approach will produce biased estimates of the price index if $(\beta_t' X_{it} - \beta_s' X_{is})$ or $(\lambda_t' Z_{it} - \lambda_s' Z_{is})$ is correlated with $\delta_t D_{it} - \delta_s D_{is}$. Unobserved renovations—changes in Z—will cause an upward bias in the rate of appreciation, as will a concentration of sales from a location whose rate of appreciation is higher than is typical for the sample area (*i.e.*, changes in λ). Again, this problem of a concentration of sales with characteristics or locations that tend to have unusual rates of appreciation tends to be referred to in the literature as a "sample selection" issue, though Equation (4) shows that it can just as easily be viewed as an omitted variables problem.

Matching

Although the hedonic and repeat sales estimators are derived from equations that include large sets of controls for characteristics of the home and location, it is the coefficients for the time dummy variables that are of interest when the objective is to construct a quality-adjusted price index. Indeed, it is because β and λ are "nuisance" terms that Equation (2) can be differenced to form the repeat sales estimator. What is less commonly recognized is that the coefficients of interest, $\delta_1 \ldots \delta_T$, are forms of "treatment effects." To see this point clearly consider a simple two-period model:

$$\begin{aligned} y_{it} = \delta_1 D_{i1} + \delta_2 D_{i2} + \beta' X_{it} + \lambda' Z_{it} + u_{it} = \delta_1 + (\delta_2 - \delta_1) D_{i2} \\ + \beta' X_{it} + \lambda' Z_{it} + u. \end{aligned} \tag{6}$$

Equation (6) could be estimated directly to construct an index with 0 as the base, time 1, value and the estimated coefficient for D_2 as the value in the second period. Alternatively, a repeat sales model could be estimated by regressing $y_2 - y_1$ on D_2 while omitting the intercept. The repeat sales version of the model is equivalent to a standard difference-in-means test: do the average log sales prices differ between period 1 and 2 for the sample of homes that sold in both periods?

Either approach provides an estimate of the expected difference in the price of a home selling in period 2 rather than period 1. This difference is written as an average treatment effect (ATE) as follows: $ATE = \frac{1}{n_2}\sum_{i=1}^{n_2} D_{i2}E[y_i(t_2) - y_i(t_1)]$, which represents the expected difference in sale price between the first and second period for the sample of homes that actually sold in the second period. ATE is the "average treatment effect on the treated," where the "treated" are the n_2 properties that sold in the second time period.[2] Since period 1 serves as the base for the price index, this expression applies directly to subsequent time periods also:

$$ATE(t_j) = \frac{1}{n_j}\sum_{i=1}^{n_j} D_{ij}E[y_i(t_j) - y_i(t_1)]. \tag{7}$$

Equation (7) is a version of a Laspeyres price index: what is the expected change in sale price relative to the base period for those properties that actually sold in time t_j?

An important insight from the literature on treatment effects is that the requirements for estimating an average effect are much less stringent than the requirements for accurately estimating an entire set of parameters. There would be no need to estimate a regression if sales were randomly distributed across time periods; a comparison of average sales prices across time periods is all that would be required to construct the price index. The hedonic approach attempts to control directly for nonrandom sample selection by including observable characteristics as control variables, while the repeat sales estimator controls for the effect of such characteristics indirectly by limiting the sample to homes selling twice. Neither approach would be necessary if sample selection were purely random.

Matching estimators attempt to reduce the effects of nonrandom sample selection through various algorithms designed to match treatment observations with similar observations from the control sample. Within the context of price index estimation, a "treatment" observation can naturally be considered as a sale from time t while a "control" observation is a sale from the base time period. Many methods have been proposed for constructing the matches. The repeat sales method is an extreme form in which each observation is matched only to an

[2]The literature on treatment effects is large and growing. Excellent overviews are presented in Ho *et al.* (2007) and Imbens and Wooldridge (2009).

earlier sale of the same home. A natural extension of the approach would be to follow the sort of "comparable sales" approach used by appraisers, whereby each home is matched to houses with similar characteristics within the same neighborhood. If sales prices are appreciating at the same rate across house types and neighborhoods—an implicit assumption behind the repeat sales and hedonic approaches—then it is not necessary to have perfect matches to estimate the average treatment effect. Differences between matched observations will average out over many observations, leading to an accurate estimate of the rate of appreciation across time periods.

Although it is straightforward to match sales directly across many housing characteristics, it is simpler to use the propensity score approach. Since "treatment" is a sale at time t and the "control" is a sale at time 1, the predicted values from a probit or logit model of sales time can be used to construct matches. Defining $I_t = 1$ if a sale occurs at time t and $I_t = 0$ if it sold during the base period, the propensity score is simply the set of predicted values from a probit or logit regression of I_t on the observed set of housing and location characteristics. Alternatively, each observation might be matched to a weighted average of a set of nearby observations, with the measure based on either the propensity score or a direct distance metric.[3]

In the empirical section of the article, the data set covers the time from the first quarter of 1993 to the last quarter of 2008. I use 1993:1 as the base time. Since the initial periods have fewer sales than subsequent periods, I use an algorithm that matches each observation in 1993:1 to similar observations in subsequent quarters. The advantage of this approach is that it assures that matched samples for later quarters are restricted to observations with similar sales in the smaller base period, which assures that the matched samples remain similar across all time periods. A potential disadvantage is that a series of logit models does not account for potential dependence over time in the errors of the underlying latent variables. Moreover, the results may be sensitive to the choice of time unit. The simplicity of the approach at least partly offsets these potential problems.

The following algorithm is used to construct the matched samples:

1. For each quarter q from 1993:2 to 2008:4, estimate a logit model using all sales taking place in 1993:1 and q. The dependent variable equals

[3]Ho *et al.* (2007) present a particularly useful discussion of the practical issues involved in matching estimators. Their "MatchIt" routine in R constructs matches very rapidly for a wide variety of algorithms. Good examples of matching estimator application in urban economics include Bondonio and Engberg (2000), Bondonio and Greenbaum (2007), Cho (2009), List, McHone and Millimet (2004), McMillen and McDonald (2002), O'Keefe (2004), Reed and Rogers (2003) and Romero (2009).

one if the sale took place in quarter q and zero if the sale is from 1993:1. The explanatory variables for the logit regressions are the same as those used for the hedonic price function estimates.

2. Use the estimated propensity score from each logit regression to match n_1 observations from quarter q to sales from 1993:1, where n_1 is the number of sales in 1993:1. Based on a random ordering of the 1993:1 observations, each observation is matched without replacement to its closest counterpart in quarter q.

At the end of this matching process, the matched-sample data set comprises approximately $60n_1$ observations—n_1 matched sales for each quarter from 1993:2 to 2008:4, plus the initial n_1 sales from 1993:1 ("approximately" because some observations may not have a close match).

Wang and Zorn (1997) show that the standard repeat sales estimator reduces to period-by-period averages of sales prices in the repeat-sale sample when the number of sales does not vary across time periods. Thus, the obvious counterpart to the repeat sales estimator for the matched-sample data set is to construct quarter-by-quarter averages of the natural logarithm of sales price. Unlike the repeat sales estimator, which controls for housing characteristics by matching each home to itself across time, the matching estimator controls for these characteristics through the matching algorithm. The goal of the matching process is to balance the distributions across time, *i.e.*, to ensure that observations from each quarter from 1993:2 to 2008:4 are similar to those from 1993:1. By eliminating extreme observations, the differences in the housing characteristics average out, leaving the period-by-period sample mean as a relatively efficient estimate of the price of housing in each quarter. Similar in spirit to the repeat sales estimator, period-by-period means from matched sample estimates are likely to be much more efficient because sample sizes are much larger when single-sale properties are included.

As Ho *et al.* (2007) emphasize, the matching process is not itself an estimation procedure. Instead, it is a method of preprocessing the data to ensure that estimates are less model dependent. A hedonic price function estimated using the matched-sample data is likely to be more accurate and less sensitive to the model specification than when the estimates are based on the full data set. Any model that can be estimated using the full data set can also be estimated using the matched samples.

The matching approach makes it very easy to construct price indices across the full range of the sale price distribution. It is just as easy to construct period-by-period estimates of the median or the 10th or 90th percentile of the sample of sales prices as it is to construct the series of averages. The series of medians, 10th percentiles and 90th percentiles can then be plotted on a graph to show the evolution of these points in the sale price distribution over time. Alternatively, kernel density estimates could be used to show the entire distribution of sales prices at each time.[4]

Data

The data set includes all sales of single-family homes in Chicago for 1993–2008. In addition to sale price and the date of sale, the data set includes standard housing characteristics, including lot size, building area and the year of construction; the number of rooms, bedrooms and bathrooms; and variables indicating that the house has central air conditioning, a fireplace, brick construction and a one-car or a two-plus-car garage.[5] The properties have been geo-coded by the Cook County Assessor's Office. I then used a GIS program to determine the distance of each home from the traditional city center of Chicago at the intersection of State and Madison streets, and whether each home is within $\frac{1}{4}$ mile of a rail line, within $\frac{1}{2}$ mile of Lake Michigan, and within $\frac{1}{4}$ mile of a stop on one of Chicago's rapid transit rail lines (the "EL").

The full data set includes 168,642 sales. Of these, 51,658 sold at least twice during the 1993–2008 period. Table 1 presents summary statistics for the full sample and the subsample of 51,658 repeat sales pairs.

The base for the matched sample is the 1,651 sales taking place in the first quarter of 1993 as the base. Other times might be used as the base for the matched sample. Figure 1 shows the number of observations by quarter for each of the data sets. There is marked seasonality in the data, with far more observations in the second and third quarters of each year than in the first and fourth quarters. Matching on the first quarter has two advantages. First, it fits in the spirit of a price index by comparing sales in subsequent periods to sales in

[4]It would also be possible to use a quantile estimator to show how the sales price distribution evolves over time. McMillen (2008) uses this approach to compare the distribution of sales prices in 1995 to 2005 and finds that the distribution shifted further to the right at higher quantiles. Note that the average predicted value at each quantile is simply the target percentile of the distribution of the dependent variable. Thus, a period-by-period series of quantiles is equivalent to period-by-period quantile regression estimates. In principle, it is possible to estimate a pooled quantile regression across all observations, including quarter dummy variables and controls for characteristics of the house and location. In practice, I could not estimate these models because of the large number of observations and explanatory variables.

[5]The housing characteristics are measured with some error because the variables are drawn from the 2003 Cook County assessment roll. Some bias will be introduced by observations representing homes that were renovated between the time of sale and the time of assessment.

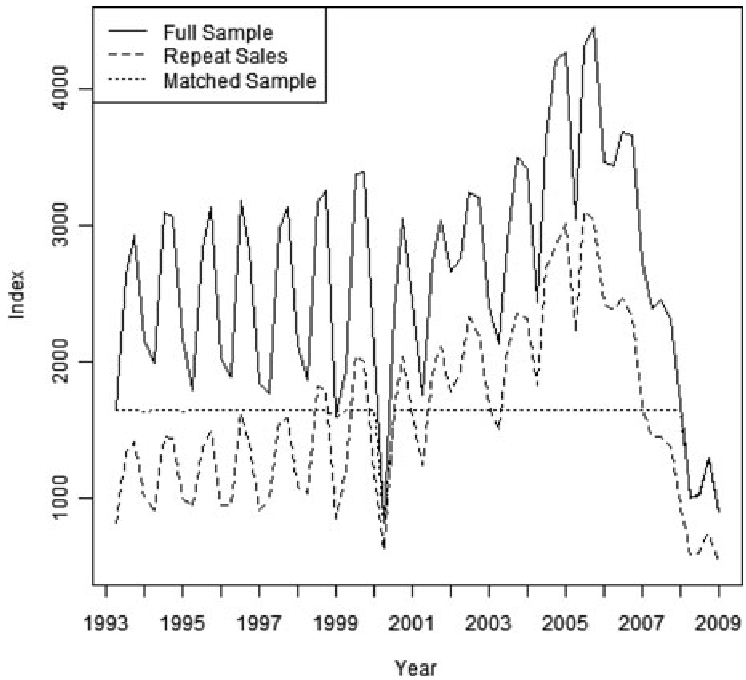
Table 1 ■ Descriptive statistics.

	Full Sample		Repeat Sales		Matched Sample	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Log of sale price	11.934	0.603	12.211	0.566	11.940	0.623
Log of lot size	8.255	0.326	8.223	0.337	8.268	0.316
Log of building area	7.063	0.299	7.057	0.301	7.076	0.301
Number of rooms	5.477	1.358	5.476	1.369	5.524	1.358
Number of bedrooms	2.853	0.777	2.833	0.783	2.879	0.778
Number of bathrooms	1.312	0.491	1.313	0.494	1.328	0.501
Central air conditioning	0.202	0.401	0.196	0.397	0.218	0.413
Fireplace	0.092	0.289	0.092	0.289	0.103	0.304
Brick construction	0.619	0.486	0.594	0.491	0.633	0.482
Garage, 1 car	0.298	0.457	0.302	0.459	0.286	0.452
Garage, 2+ car	0.464	0.499	0.452	0.498	0.490	0.500
Distance from city center	9.014	2.700	8.715	2.699	9.046	2.629
Within 1/4 mile of EL stop	0.048	0.213	0.054	0.227	0.043	0.203
Within 1/2 mile of Lake Michigan	0.012	0.108	0.010	0.101	0.012	0.110
With 1/4 mile of rail line	0.086	0.280	0.085	0.279	0.082	0.275
Age of house at time of sale	69.994	26.020	74.807	26.246	70.165	25.241
Number of observations	168,642		51,658		102,064	

the base time period. Thus, it is analogous to a Laspeyres price index, showing how prices vary for homes that are directly comparable to those that sold during the first quarter of the sample. The second advantage is that using the relatively small number of observations from the first quarter as the base is a conservative matching approach that helps deter relatively poor matches from entering the matched sample. If 2005 were used as the base, for example, the matching procedure would attempt to match each observation from this very large group of sales to observations from times with fewer sales. A later section of this article compares some alternative matching procedures.

The repeat sales estimator is a matching approach based on *exact* matches: each home is matched only to previous sales of the same property. Matches can also be based on various distance metrics and subclassification schemes. The excellent MatchIt program in *R* includes a wide variety of alternative matching schemes (Ho *et al.* 2011). I use a propensity score approach to construct the matches. Using 1993:1 as the base, I estimate 63 logit models of the probability that a home sold in quarter q (where $q = 1993{:}2, \ldots, 2008{:}4$) using the subset of observations of sales from 1993:1 and quarter q. The matching algorithm begins by matching a randomly drawn observation from

Figure 1 ■ Number of observations by quarter.



1993:1 to its closest counterpart from quarter q, where distance is simply the difference in the estimated propensity scores from the logit model for quarter q. The algorithm then moves to the next randomly drawn (without replacement) observation from 1993:1, matching it with its closest counterpart from the unmatched observations from quarter q. Most of the matched samples will have the same number of observations as the 1993:1 subsample, 1,651. The primary exceptions are quarters with fewer sales than 1993:1. In addition, I discard any observations that fall outside the support for the propensity scores for the 1993:1 observations in the logit model for quarter q. In the end, the combined matched samples have 102,064 observations, of which 1,651 are the base 1993:1 sales.

Summary statistics for the matched sample are shown in Table 1. The number of observations by quarter is shown in Figure 1. By construction, the matched sample has approximately the same number of observations in each quarter, and the full, hedonic sample has more observations than the matched sample. Figure 1 also shows the number of observations represented in the repeat sales

Table 2 ■ Hedonic and matched-sample regressions.

	Full Sample		Matched Sample	
	Mean	Standard Error	Mean	Standard Error
Constant	7.083	0.032	7.053	0.041
Log of lot size	0.260	0.003	0.270	0.004
Log of building area	0.285	0.004	0.281	0.005
Number of rooms	−0.004	0.001	−0.005	0.001
Number of bedrooms	0.008	0.002	0.008	0.002
Number of bathrooms	0.048	0.002	0.046	0.002
Central air conditioning	0.016	0.002	0.013	0.002
Fireplace	0.079	0.003	0.075	0.003
Brick construction	0.047	0.002	0.054	0.002
Garage, 1 car	0.043	0.002	0.043	0.003
Garage, 2+ car	0.059	0.002	0.058	0.002
Distance from city center	0.019	0.001	0.020	0.002
Within 1/4 mile of EL stop	0.019	0.004	0.029	0.005
Within 1/2 mile of Lake Michigan	0.025	0.008	0.014	0.009
With 1/4 mile of rail line	−0.013	0.003	−0.016	0.003
Age of house at time of sale	−0.002	0.000	−0.002	0.000
R^2	0.772		0.791	
Number of observations	168,642		102,064	

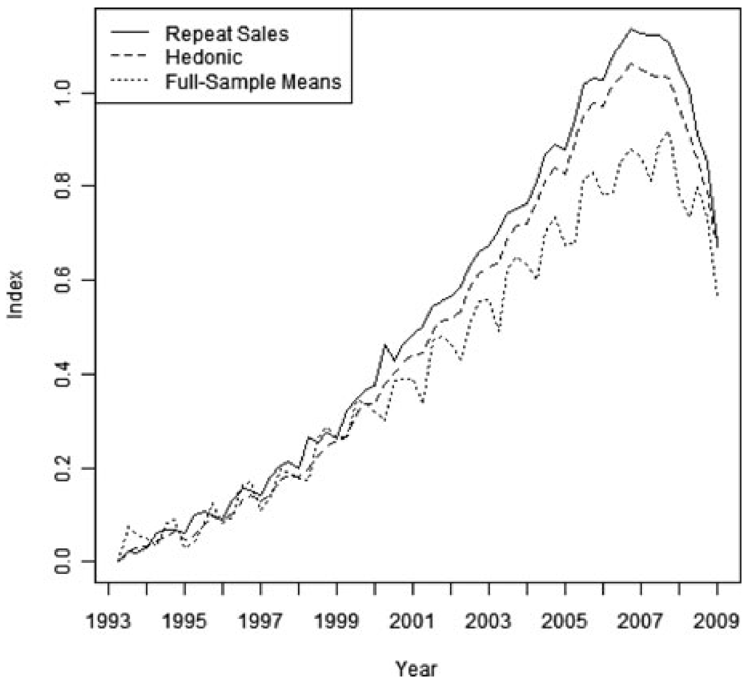
pairs.[6] The matched sample has more sales than the pooled repeat sales sample in the later periods, primarily because homes that sold in the later years are more likely to have sold at least twice since there are simply more sales in later years and those sales have a long history from which an earlier sale may have occurred. The repeat sales estimator is often found to be very sensitive to the small number of sales from early time periods. Figure 1 shows that an advantage of the matching approach is that it greatly increases the number of observations from these early quarters.

Results for Chicago, 1993–2008

Table 2 compares the estimated hedonic price functions for the full sample and the combined matched samples. The natural logarithm of the sale price is the dependent variable for these regressions. In addition to the variables shown in Table 2, the regressions also include controls for the quarter of sale and the community area in which the home is located.[7] Although these coefficients

[6]The number of observations for the repeat sales pairs is actually overstated somewhat. For example, if a home sold in 1995:1, 2000:1 and 2005:1, it will be represented by two repeat sales pairs, with the first pair showing the change in price from 1995:1 to 2000:1 and the second pair showing the change from 1995:1 to 2005:1. As a result of this implicit matching procedure, the sale from 2000:1 does double duty.

Figure 2 ■ Standard price indices.

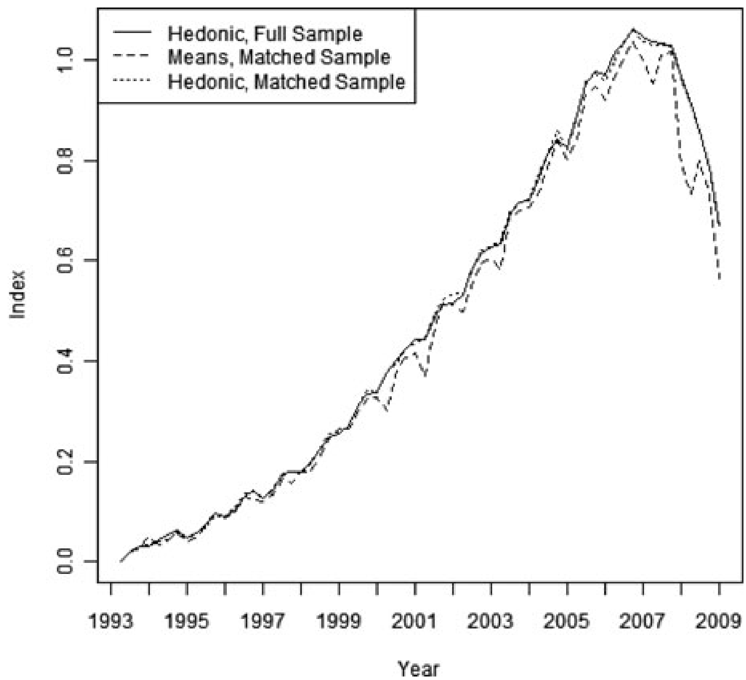


are not the focus of the study, it is clear that most of the estimates are highly significant with the expected signs, and the overall explanatory power of the regressions is very good. The coefficients for the quarterly dummy variables form the price indices for the hedonic and matched samples, with the value for 1993:1 normalized to zero. Using the repeat sales sample, the estimated price index is constructed by regressing the change in the natural log of sale price on a series of 63 indicator variables that equal one if the first sale in a repeat sales pair takes place at the time indicated by the variable, -1 if the second sale in the pair occurs at that time, and zero otherwise.

The standard hedonic and repeat sale price index estimates are shown in Figure 2. Both indices indicate much higher rates of appreciation over 1993–2008 than is indicated by simple quarterly averages of the log of sales

[7] "Community areas" were defined by a group of University of Chicago sociologists in the 1930s, and are still widely used today to refer to Chicago neighborhoods. There are 77 community areas in Chicago, compared with 865 census tracts. The average area covered by a community area is approximately three square miles.

Figure 3 ■ Hedonic price index and matched sample means.



prices over this time. The repeat sales estimator indicates a higher rate of appreciation than the hedonic estimator.

Figure 3 compares the hedonic estimates to two approaches for estimating price indices with the combined matched sample. The first approach follows directly from Equation (7): once the distributions of covariates are balanced by matching sales across the control and treatment samples—1993:1 versus subsequent quarters—the difference in means provides a direct estimate of the rate of appreciation from 1993:1 to quarter q. Figure 3 shows that these differences in matched-sample means are remarkably close to the hedonic price index estimates.

If the matching procedure is not fully successful in balancing the distributions over time, the estimates can potentially be improved by including the explanatory variables from the matching procedure as controls when calculating the difference in means. The resulting estimator is simply the standard hedonic model applied to the matched sample. Figure 3 shows the result, which is nearly identical to the base hedonic price index.

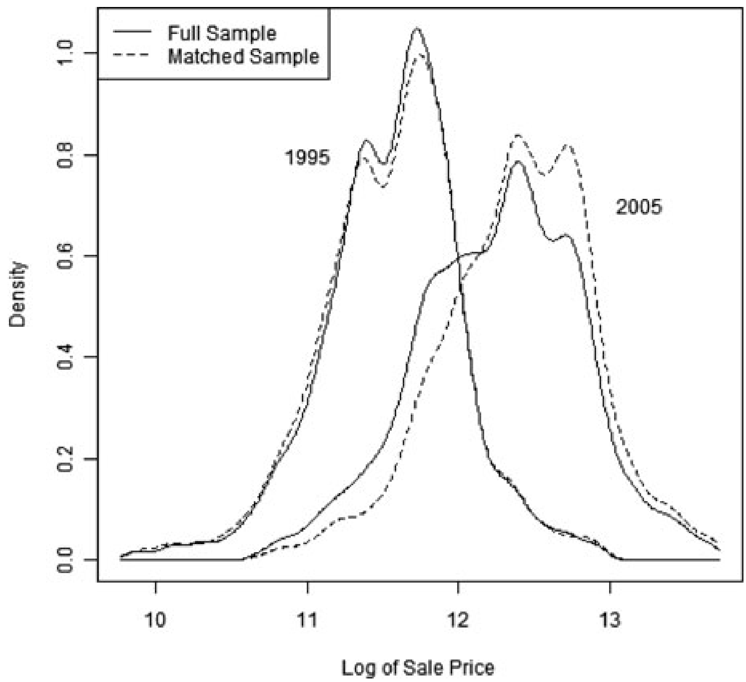
Since all of these approaches involve estimating unknown parameters, there is no way to know with certainty which price index is closest to the truth. The matching estimator shows clearly that the hedonic estimator is quite robust in this application. Whereas the full sample means are much different from the hedonic price index, the matched sample means are quite similar to the price index produced by estimating a hedonic regression with the matched sample. This result suggests that the matching approach successfully removes unrepresentative, low-priced observations that are pulling down average sales prices in later quarters. The hedonic regression controls for these effects directly by including controls for location and housing characteristics; the matching approach removes them through pre-estimation screening of the sample.

Once the observations without matches are removed from the sample, a simple comparison of means appears to serve as a reliable estimator of the price index. This result is important because it suggests that other points of the sale price distribution may be estimated reliably using simple percentiles of the log of sales prices for each quarter's matched sample. In fact, the full distribution of sales prices can be estimated using a kernel density approach. Kernel density results are shown in Figure 4 for 1995 and 2005 for the full sample and the combined matched samples. Since 1995 is close to the base year, it is perhaps not surprising that the estimated density of the log of sales prices is very similar for the full sample and the smaller, matched sample. By 2005, however, there are some important differences in the estimated price densities. In particular, the left side of the 2005 sale price density has shifted farther to the right for the matched sample than is the case for the full sample.

Another way to see this result is to compare percentiles for the log of sales prices across time. These results are shown for the full sample and the combined matched samples in Figure 5. The lines represent the 10%, 50% and 90% percentiles of the distribution of the natural log of sale price in the matched samples. The difference between the 10% and 90% percentiles is shown in Figure 6. Both samples suggest that the spread in the distribution is widening over time due to higher appreciation rates at higher quantiles. However, the widening of the distribution is somewhat less marked in the matched sample.

In McMillen (2008), I used a quantile estimator to show that the distribution of sales prices in Chicago became less skewed between 1995 and 2005, with the distribution shifting farther to the right for higher sales prices than is the case on the left side of the distribution. The kernel density estimates shown in Figure 4 for the full sample are nearly identical to those shown in McMillen (2008). The matching estimator produces a similar pattern of results, but the matched sample indicates a higher rate of appreciation in the low-price portion of the distribution than is indicated for the full sample. The reason for this

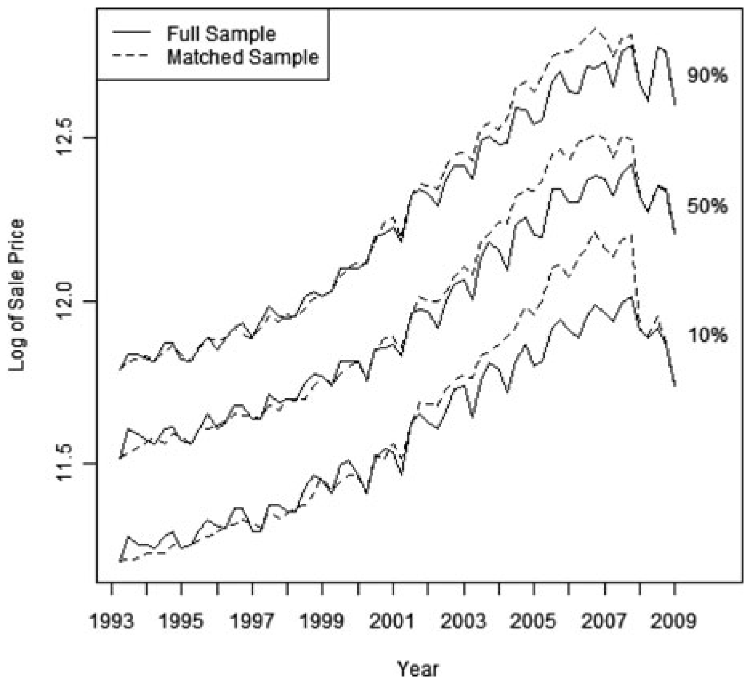
Figure 4 ■ Sale price—densities.



difference in results is the same as the reason that the matched sample means show a higher rate of appreciation than the full sample means: the pre-estimation matching procedure eliminates many of the low-priced observations that are in the full sample in later time periods. Excluding these observations is important if the objective is to measure a quality-controlled rate of appreciation. The matching estimator suggests part of the reason that the distribution shifted farther to the right for high-priced properties is that there were many more sales of low-priced properties with characteristics that differed markedly from the homes that sold in earlier time periods.[8]

[8]Guerrieri, Hartley and Hurst (2010) present a model in which endogenous gentrification causes the sale price distribution to shift to the right. In their model, higher income households want to live near other high-income households, which causes high-priced homes to appreciate in high-income neighborhoods. Note that this sort of externality is not accounted for by the explanatory variables in a standard hedonic price function, and the repeat sales estimator fails to account for the externality because the characteristics of the neighborhood are changing over time. The matching estimator is also affected by omitted variables.

Figure 5 ■ Price quantiles.

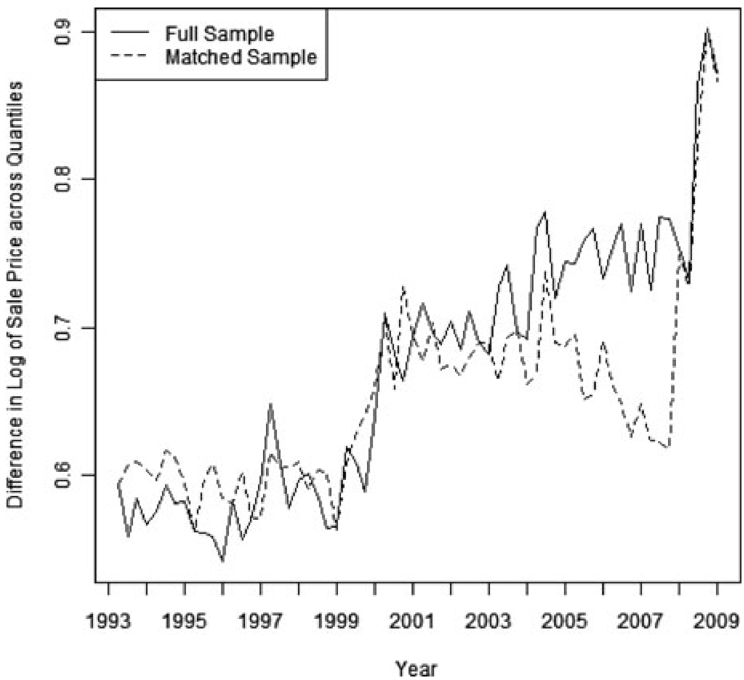


Figures 4 and 5 illustrate the advantages of the matching approach to price index estimation. The initial matching process produces samples with similar distributions of housing characteristics. Because the distribution of characteristics is similar over time, it is appropriate to simply compare means, medians or other points in the sale price distribution over time. Thus, the density estimates shown in Figure 4 for the matched sample show changes in the full distribution of quality-controlled sales prices over time. Figure 5 shows price indices that are analogous to standard mean-based estimates, but they provide much more information by showing how prices evolve over different parts of the distribution. Thus, the matching estimator produces a much richer set of results than is available using either the repeat sales or standard hedonic approach.

Alternative Matching Approaches

The standard hedonic data set includes many time periods and many explanatory variables. Although I have followed a common practice in basing matches on estimated propensity scores, many alternative matching procedures are available. Most prominent among the alternatives is directly matching observations

Figure 6 ■ Difference between 10% and 90% quantiles.



based on a measure of distance between the explanatory variables for the hedonic regression. In addition, there is the question of *time*: although I have chosen to use the first quarter as the base time period for the matching procedure, any time can potentially serve as the base. In this section, I compare propensity score matching with matching based on Mahalanobis measures of distance. I also vary the time used as the base.

Figure 1 helps illustrate the issues involved in the choice of the base time period. Only 1,651 sales took place during the first quarter, 1993:1. By using it as the base, each subsequent period's matches are limited to no more than 1,651 comparable properties. In contrast, the peak number of sales, 4,442, took place during the third quarter of 2005. If 2005:3 were used as the base, nearly every observation would be included in the matched data sets—the exceptions being those that are excluded for being outside the common support of the estimated propensity scores. Thus, it seems reasonable to use a time with relatively few sales as the base. In this example, using the first quarter as the base has two advantages. First, it is a time with relatively few sales, which leads to a conservative matching procedure in which a relatively large number of observations is discarded. Second, the interpretation of the results is the natural one for a price index: what is the expected sale price for future dates if homes have characteristics that are similar to those selling in the initial time period?

A primary goal of matching is to assure the distributions of the explanatory variables are similar across the base and matched samples. For a simple two-period case, Ho *et al.* (2011) suggest comparing quantiles of the distributions across samples for all explanatory variables. Such a comparison becomes very cumbersome for a data set with 64 time periods. To save space, I focus on mean values.

Table 3 compares mean values for the base time period to the mean values across all of the other 63 matched samples. The matched samples are those that have been used so far—a series of matched samples based on period-by-period logit regression models using 1993:1 as the base. Although a simple comparison of 1993:1 to 1993:2–2008:4 neglects the variation across quarters, it does provide a measure of the extent to which matching reduces the difference between the means for each explanatory variable across all 63 quarters.

The first row of results in Table 3 is illustrative. The mean of the log of lot size is 8.270 for 1993:1, which compares with a mean of 8.254 across the remaining 63 quarters. Thus, the raw data reveal that lot sizes were larger on average in the base period than subsequently. After constructing matches based on estimated propensity scores from a series of logit regressions, the mean value of the log of lot size is 8.267 across the 63 matched samples. After matching, the difference in means between the 1993:1 and 1993:2–2008:4 samples falls from 0.016 to 0.003, which on a percentage basis is an improvement of 77.963%. Comparable improvements are found for all of the other explanatory variables. In general, Table 3 suggests that homes that sold in 1993:1 were larger and perhaps of higher quality than the homes selling in later time periods—a result that explains the large difference between the full-sample means and the standard hedonic and repeat sales indices shown in Figure 2.

Figures 7 and 8 show mean values across quarters for two of the most important explanatory variables, lot size and building area. Both graphs include horizontal lines that show the variable's mean value in 1993:1. The dashed lines represent the quarterly means for the full, unmatched samples. The solid lines show the means across quarters for the matched samples. The matching procedure significantly decreases the variability in the means across quarters. These graphs are representative of those for the other variables. Using a series of logit models to construct matches produces a series of matched samples whose explanatory variables have mean values that are close to the 1993:1 means.

Table 3 ■ Matching summary statistics.

Sample	Means, 1993:1 Full = Matched	Means, 1993:2–2008:4 Full	Means, 1993:2–2008:4 Matched	Difference in Means Full	Difference in Means Matched	Percentage Improvement
Log of lot size	8.270	8.254	8.267	0.016	0.003	77.963
Log of building area	7.083	7.063	7.079	0.020	0.004	81.183
Number of rooms	5.555	5.476	5.544	0.079	0.011	85.993
Number of bedrooms	2.891	2.852	2.888	0.038	0.003	91.981
Number of bathrooms	1.338	1.312	1.333	0.026	0.005	80.479
Central air conditioning	0.223	0.202	0.219	0.022	0.004	79.596
Fireplace	0.108	0.092	0.104	0.016	0.004	76.298
Brick construction	0.632	0.619	0.632	0.013	0.001	95.858
Garage, 1 car	0.283	0.298	0.284	−0.015	−0.001	92.709
Garage, 2+ car	0.495	0.464	0.493	0.032	0.002	93.400
Distance from city center	8.996	9.014	8.998	−0.018	−0.002	91.527
Within 1/4 mile of EL stop	0.044	0.048	0.044	−0.004	−0.001	83.443
Within 1/2 mile of Lake Michigan	0.014	0.012	0.013	0.002	0.001	57.071
With 1/4 mile of rail line	0.083	0.086	0.083	−0.003	0.001	81.728
Year built	1929.936	1930.633	1929.929	−0.697	0.007	99.016
Latitude	41.848	41.839	41.846	0.009	0.002	81.881
Longitude	−87.708	−87.704	−87.707	−0.004	−0.001	77.131

Figure 7 ■ Means for matched and unmatched observations: Log of building area.

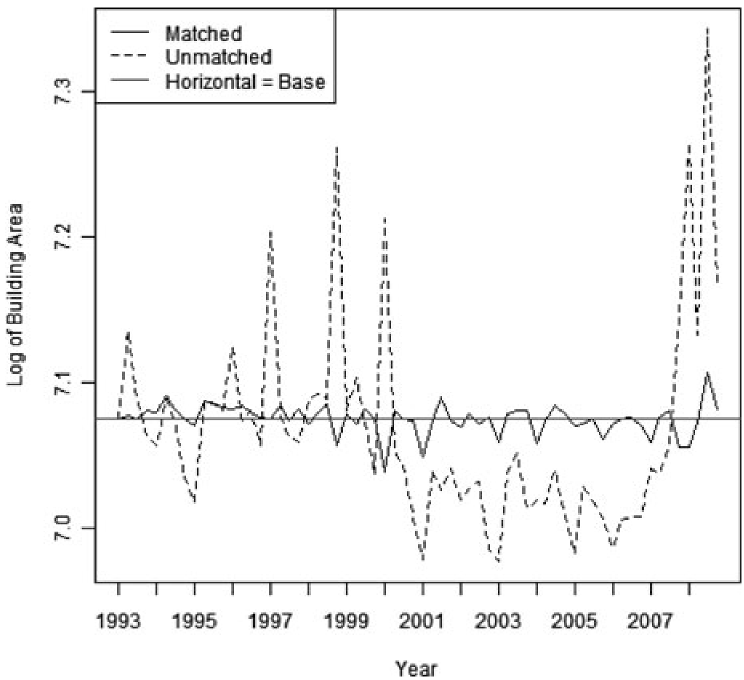


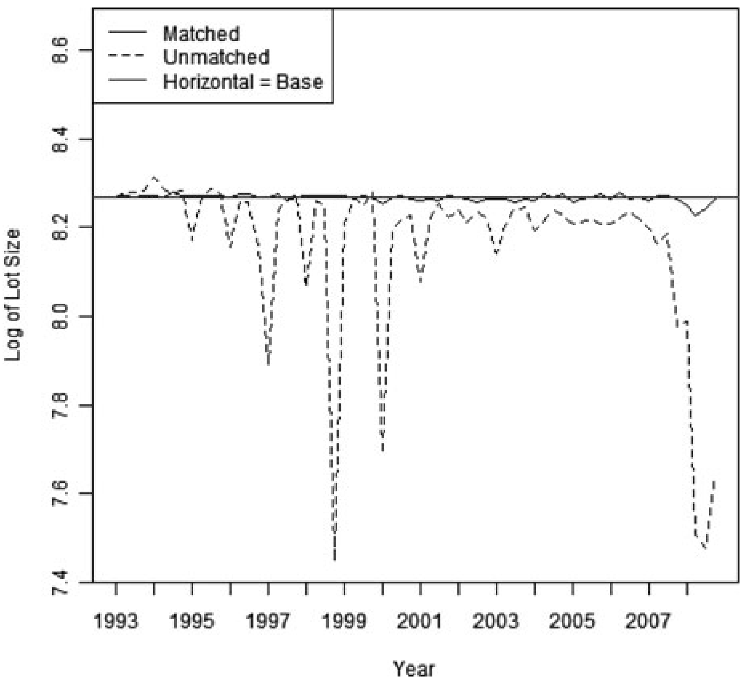
Table 4 compares results for alternative definitions of the base time period. It also presents the results when a Mahalanobis measure of distance is used in place of logit propensity scores. The numbers in Table 4 are comparable to those in the last column of Table 3, and the results for Model 1 in Table 4 are identical to those in Table 3.

Models 1–3 differ by base period. Whereas 1993:1 has 1,651 observations, 2000:3 has 3,049 and 2008:1 has only 1,019. Since Model 2 has a large number of observations, the matched samples are large when 2000:3 is the base. With fewer observations eliminated, the matched samples are not as homogeneous as the case where the base period has a small number of observations. Thus, Models 1 and 3 tend to have larger percentage improvements in the difference in means across samples (*i.e.*, in the "balance" of the means across samples). There is not a large difference in the percentage improvements in the mean balance between the 1993:1 and 2008:1 base periods, even though 2008:1 has significantly fewer observations. The results for propensity scores matching suggest that it is a good idea to choose a base time period with a relatively small

Table 4 ■ Percentage improvement in mean balance for matched samples.

Matching Method	Propensity Score				Mahalanobis			
Model Number	1	2	3	4	5	6	7	8
Base time	1993:1	2000:3	2008:1	1993:1, closest	1993:1	2000:3	2008:1	1993:1, closest
Log of lot size	77.963	3.856	96.659	80.757	36.527	−18.460	−64.749	25.416
Log of building area	81.183	−96.675	90.247	84.610	−22.193	63.898	−24.502	−20.359
Number of rooms	85.993	17.144	83.353	90.915	−64.280	64.631	−97.812	−69.824
Number of bedrooms	91.981	13.744	88.976	96.671	−50.243	60.627	−35.921	−53.203
Number of bathrooms	80.479	20.123	90.280	89.844	−107.446	92.215	−157.036	−112.035
Central air conditioning	79.596	41.770	88.670	76.332	33.903	13.122	80.783	32.013
Fireplace	76.298	29.669	97.685	84.770	34.898	41.054	70.972	33.531
Brick construction	95.858	−322.669	96.543	93.472	86.413	−532.546	66.552	99.418
Garage, 1 car	92.709	19.165	96.176	92.097	94.391	12.080	77.026	95.120
Garage, 2+ car	93.400	−27.525	94.882	95.865	96.579	−64.262	71.390	97.469
Distance from city center	91.527	39.499	55.053	85.428	−288.139	−43.646	39.534	−1.051
Within 1/4 mile of EL stop	83.443	20.052	83.880	99.605	83.443	−41.139	69.966	55.148
Within 1/2 mile of Lake Michigan	57.071	52.279	81.760	69.241	68.204	44.159	87.143	34.266
With 1/4 mile of rail line	81.728	38.123	31.506	58.263	57.489	−217.122	−455.648	−6.351
Year built	99.016	50.152	96.507	93.063	−11.617	52.317	53.386	8.877
Latitude	81.881	40.940	97.680	86.328	−16.313	−9.406	60.235	23.233
Longitude	77.131	37.373	99.452	88.322	13.345	8.751	74.062	64.616
Number of observations in base time	1650	3049	1019	1000	1650	3049	1019	1000

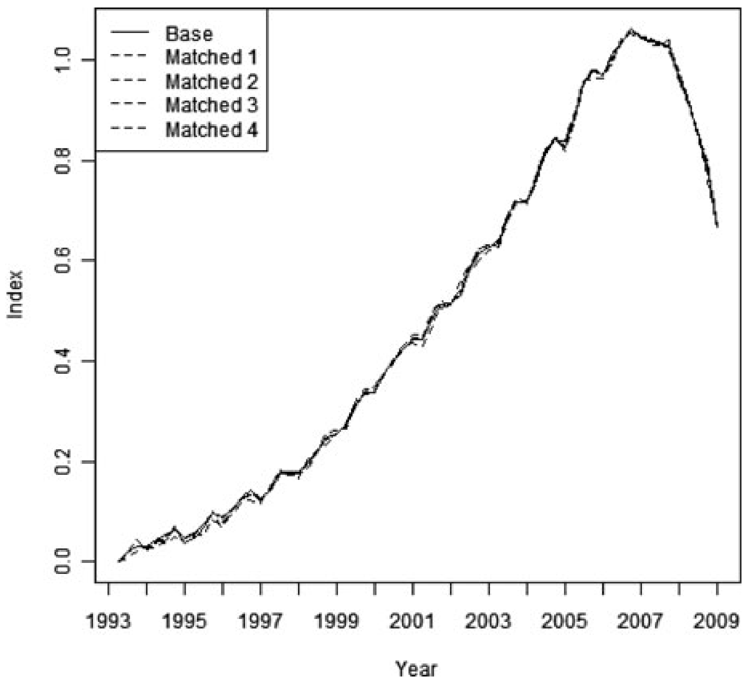
Figure 8 ■ Means for matched and unmatched observations: Log of lot size.



number of observations. However, the key is to have relatively homogeneous samples; a time period with a relatively large number of observations with similar values for the explanatory variables may also serve as a good base.

Model 4 is intended to investigate this issue of homogeneity. Rather than using all 1,651 observations from 1993:1 as the base, Model 4 matches only the 1,000 observations with the closest propensity scores. Thus, the 1,000 observations from 1993:2 with propensity scores closest to an observation from 1993:1 are included in the 1993:2 matched samples. Since the process is repeated for each subsequent year, all 1,651 observations from 1993:1 are likely to eventually find a close match with an observation from a later time, but only 1,000 observations are ever included from the 1993:2 to 2008:4 samples. The percentage improvement in the mean balance is only marginally better than when the matched samples from later time periods are allowed to have as many as the full 1,651 observations. Overall, the results in Table 4 suggest that propensity score matching is generally successful at balancing means across samples when the base time has a relatively small sample size.

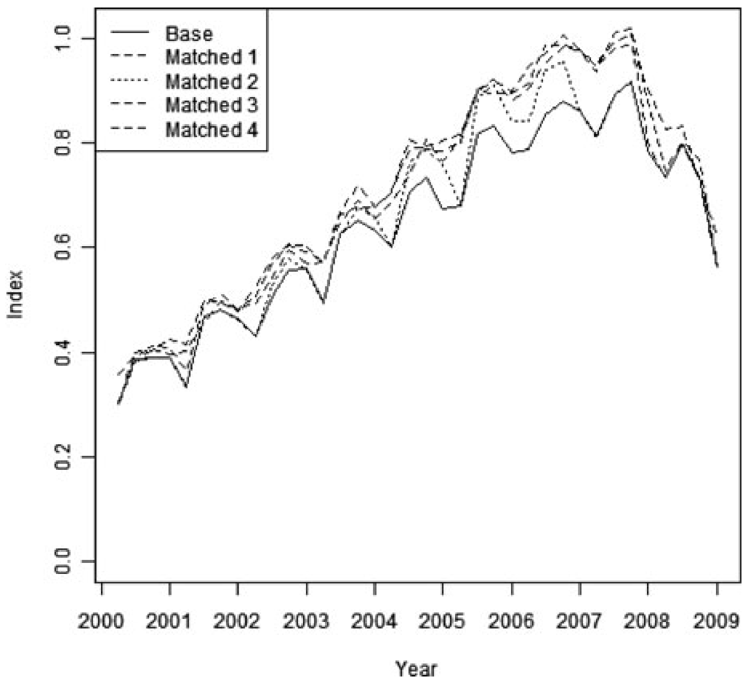
Figure 9 ■ Hedonic price indices.



The final four columns of results in Table 4 show the percentage improvements in the mean balance when a Mahalanobis measure of distance between the explanatory variables is used in place of the propensity score. Somewhat surprisingly in view of findings that Mahalanobis matching is "relatively robust" and that propensity score matching can perform poorly in small samples (Zhao 2004), the results in Table 4 suggest that propensity score matching leads to higher percentage improvements in mean balance in this example. Given Zhao's (2004) findings, this result may be explained by the relatively large sample sizes in this study and the tendency for a relatively large number of significant variables in the logit models of sale time.

Figures 9 and 10 compare the price indices calculated using the alternative propensity score matching dates. After normalizing the price indices to have a value of zero in 1993:1, Figure 9 shows that the results are virtually identical when the price indices are constructed using hedonic regressions of log sales prices on the full set of explanatory variables and a series of quarterly dummy variables. Figure 10 compares the results in the more interesting case where

Figure 10 ■ Sample means, 2000–2008.



price indices are constructed using simple period-by-period means of the log sales prices. To make the results easier to see, Figure 10 focuses on the post-2000 time period, which is when the various price indices begin to look somewhat different from one another. Figure 10 shows that the prices indices obtained by calculating quarterly means using the matched samples associated with Models 1, 3 and 4 are all very similar. In contrast, full-sample means and Model 2—matched samples constructed using a large base sample size—indicate relatively low appreciation rates in the later time periods when smaller, lower quality housing comprises a larger portion of the sample.

It is worth emphasizing again that the standard repeat sales estimator is *identical* to period-by-period means when the number of observations is the same across time. Thus, the fact that matching produces very stable price index estimates is very encouraging, particularly since the procedure is so easily adaptable to modeling changes across the full distribution of house prices. Although more research is warranted, the results here suggest that simple period-by-period matching based on propensity scores produces more reliable price index estimates than the standard repeat sales estimator, and it has the additional benefit that matched sample percentiles can easily be used to characterize changes in the full distribution of house prices.

Conclusion

The repeat sales approach to estimating price indices is an extreme version of a matching estimator in which a house sale enters the sample only if it can be paired with a previous sale of the same home. Matching sales with similar but not necessarily identical homes has several advantages relative to either repeat sales or hedonic approaches for estimating price indices. Compared with the standard repeat sales estimator, a more general matching approach greatly increases the sample size, which tends to produce more efficient estimates and also allows price indices to be constructed for smaller geographic areas and shorter time intervals.

Wang and Zorn (1997) show that the standard repeat sales approach reduces to period-by-period means when the number of sales in the repeat sales sample does not vary over time. This observation has several interesting implications for constructing price indices. It shows that the simplest version of the matching approach—comparing means for a series of identically sized matched samples—is equivalent to a repeat sales estimator with the sample supplemented to include similar but not necessarily identical homes. Since the average predicted value of a regression is equal to the average value of the dependent variable by construction, the simple matching approach is also identical to averages of the predicted values of period-by-period estimates of hedonic price functions. The difference between the repeat sales and hedonic approach is that (1) the samples differ and (2) the most common version of the hedonic approach includes a restriction that the coefficients for most of the explanatory variables are constant over time. The matching approach serves as a middle ground between the two approaches. Like the repeat sales estimator, the matching approach uses a subset of the full hedonic sample, but the matching estimator can also closely follow the hedonic approach by including explicit controls for various explanatory variables and constraining some of the coefficients to be constant over time.

The matching approach facilitates the construction of price indices across the full range of the distribution of sales prices. The basic version of the estimator involves comparing means across time; an obvious extension is to compare other percentiles of the distribution. Little is known about the evolution of the full distribution of prices over time, at least in part because the standard tools for tracking price changes—the repeat sales and hedonic approaches—are mean-based estimators. The matching approach implicitly controls for quality differences by matching homes to similar properties selling at different times. Thus, the matching estimator can help track changes in the price distribution for a constant-quality set of homes.

References

Bailey, M., R. Muth and H. Nourse. 1963. A Regression Method for Real Estate Price Index Construction. *Journal of the American Statistical Association* 58: 933–942.

Bondonio, D. and J. Engberg. 2000. Enterprise Zones and Local Employment: Evidence from the States' Programs. *Regional Science and Urban Economics* 30: 519–459.

—— and R. Greenbaum. 2007. Do Local Tax Incentives Affect Economic Growth? What Mean Impacts Miss in the Analysis of Enterprise Zone Policies. *Regional Science and Urban Economics* 37: 121–136.

Cannaday, R., H. Munneke and T. Yang. 2005. A Multivariate Repeat-Sales Model for Estimating House Price Indices. *Journal of Urban Economics* 57: 320–342.

Case, B., P. Colwell, C. Leishman and C. Watkins. 2006. The Impact of Environmental Contamination on Condo Prices: A Hybrid Repeat-Sales/Hedonic Approach. *Real Estate Economics* 34: 77–107.

——, H. Pollakowski and S. Wachter. 1997. Frequency of Transaction and House Price Modeling. *Journal of Real Estate Finance and Economics* 14: 173–187.

—— and J. Quigley. 1991. The Dynamics of Real Estate Prices. *Review of Economics and Statistics* 73: 50–58.

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

—— and ——. 1989. The Efficiency of the Market for Single-Family Homes. *American Economic Review* 79: 125–137.

Cho, R. 2009. Impact of Maternal Imprisonment on Children's Probability of Grade Retention. *Journal of Urban Economics* 65: 11–23.

Clapp, J., C. Giaccotto and D. Tirtiroglu. 1991. Housing Price Indices: Based on All Transactions Compared to Repeat Subsamples. *American Real Estate and Urban Economics Association Journal* 19: 270–285.

Deng, Y., D. McMillen and T. Sing. 2012. Private Residential Price Indices in Singapore: A Matching Approach. *Regional Science and Urban Economics* 42: 485–494.

Englund, P., J. Quigley and C. Redfearn. 1999. The Choice of Methodology for Computing Housing Price Indexes: Comparisons of Temporal Aggregation and Sample Definition. *Journal of Real Estate Finance and Economics* 19: 91–112.

Gatzlaff, D. and D. Haurin. 1997. Sample Selection Bias and Repeat-Sales Index Estimates. *Journal of Real Estate Finance and Economics* 14: 33–50.

Guerrieri, V., D. Hartley and E. Hurst. 2010. Endogenous Gentrification and Housing Price Dynamics. Working Paper. University of Chicago.

Ho, D., K. Imai, G. King and E. Stuart. 2007. Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference. *Political Analysis* 15: 199–236.

——, ——, —— and ——. 2011. Matchit: Nonparametric Preprocessing for Parametric Causal Inference. *Journal of Statistical Software* 42: 1–28.

Imbens, G. and J. Wooldridge. 2009. Recent Developments in the Econometrics of Program Evaluation. *Journal of Economic Literature* 47: 5–86.

Kiel, K. and J. Zabel. 1997. Evaluating the Usefulness of the American Housing Survey for Creating House Price Indices. *Journal of Real Estate Finance and Economics* 14: 189–202.

List, J., W. McHone and D. Millimet. 2004. Effects of Environmental Regulation on Foreign and Domestic Plant Births: Is There a Home Field Advantage? *Journal of Urban Economics* 56: 303–326.

Longford, N. 2009. A House Price Index Defined in the Potential Outcomes Framework. Working Paper. Universitat Pompeu Fabra.

Mark, J. and M. Goldberg. 1984. Alternative House Price Indices: An Evaluation. *American Real Estate and Economics Association Journal* 12: 30–49.

McMillen, D. 2003. The Return of Centralization to Chicago: Using Repeat Sales to Identify Changes in House Price Distance Gradients. *Regional Science and Urban Economics* 33: 287–304.

——. 2008. Changes in the Distribution of House Prices over Time: Structural Characteristics, Neighborhood or Coefficients? *Journal of Urban Economic* 64: 573–589.

—— and J. McDonald. 2002. Land Values in a Newly Zoned City. *Review of Economics and Statistics* 84: 62–72.

—— and P. Thorsnes. 2006. Housing Renovations and the Quantile Repeat Sales Price Index. *Real Estate Economics* 34: 567–587.

O'Keefe, S. 2004. Job Creation in California's Enterprise Zones: A Comparison using a Propensity Score Matching Model. *Journal of Urban Economics* 55: 131–150.

Palmquist, R. 1980. Alternative Techniques for Developing Real Estate Price Indexes. *Review of Economics and Statistics* 66: 394–404.

Reed, W. and C. Rogers. 2003. A Study of Quasi-Experimental Control Group Methods for Estimating Policy Impacts. *Regional Science and Urban Economics* 33: 3–25.

Romero, R. 2009. Estimating the Impact of England's Area-Based Intervention "New Deal for Communities" on Employment. *Regional Science and Urban Economics* 39: 323–331.

Rosenbaum, P. and D. Rubin. 1983. The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika* 70: 41–55.

—— and ——. 1984. Reducing Bias in Observational Studies Using Subclassification Based on the Propensity Score. *Journal of the American Statistical Association* 79: 516–524.

Thibodeau, T. 1989. Housing Price Indices from the 1974–83 SMSA Annual Housing Surveys. *American Real Estate and Urban Economics Association Journal* 17: 110–117.

Wang, F. and P. Zorn. 1997. Estimating House Price Growth with Repeat Sales Data: What's the Aim of the Game? *Journal of Housing Economics* 6: 93–118.

Zhao, Z. 2004. Using Matching to Estimate Treatment Effects: Data Requirements, Matching Metrics, and Monte Carlo Evidence. *Review of Economics and Statistics* 86: 91–107.