

JSPS Grants-in-Aid for Scientific Research (S)
Understanding Persistent Deflation in Japan

Working Paper Series

No. 024

November 2013

Matching Indices for Thinly-Traded Commercial Real Estate in Singapore

Yongheng Deng
Daniel P. McMillen
Tien Foo Sing

UTokyo Price Project
702 Faculty of Economics, The University of Tokyo,
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
Tel: +81-3-5841-5595
E-mail: watlab@e.u-tokyo.ac.jp
<http://www.price.e.u-tokyo.ac.jp/english/>

Working Papers are a series of manuscripts in their draft form that are shared for discussion and comment purposes only. They are not intended for circulation or distribution, except as indicated by the author. For that reason, Working Papers may not be reproduced or distributed without the expressed consent of the author.

Matching Indices for Thinly-Traded Commercial Real Estate in Singapore

Yongheng Deng¹

Department of Real Estate, Department of Finance and
Institute of Real Estate Studies
National University of Singapore
21 Heng Mui Keng Terrace, #04-02
Singapore, 119613
ydeng@nus.edu.sg
+65 6516 8291

Daniel P. McMillen

Department of Economics
University of Illinois
1407 W. Gregory Dr.
Urbana, IL 61822
USA
mcmillen@illinois.edu
+01 217-333-8930

Tien Foo Sing

Department of Real Estate and Institute of Real Estate Studies
National University of Singapore
21 Heng Mui Keng Terrace, #04-02
Singapore, 119613
rststf@nus.edu.sg
+65 6516 4553

Revised: September 4, 2013

¹ *Corresponding Author.*

The authors wish to thank Daxuan Zhao for research assistance. We also than Diep Xuan Mai and Nam Chi Nguyen for their assistance with data collection.

Abstract

We use a matching procedure to construct three commercial real estate indices (office, shop and multiple-user factory) in Singapore using transaction sales from 1995Q1 to 2010Q4. The matching approach is less restrictive than the repeat sales estimator, which is restricted to properties sold at least twice during the sample period. The matching approach helps to overcome problems associated with thin markets and non-random sampling by pairing sales of similar but not necessarily identical properties across the control and treatment periods. We use the matched samples to estimate not just the mean changes in prices, but the full distribution of quality-adjusted sales prices over different target quantiles. The matched indices show three distinct cycles in commercial real estate markets in Singapore, including two booms in 1995-1996 and 2006-2011, and deep and prolonged recessions with declines in prices around the time from 1999-2005. We also use kernel density function to illustrate the shift in the distribution of house prices across the two post-crisis periods in 1998 and 2008.

Keywords: *Matching Approach, Repeat Sales, Commercial Real Estate Indices, Thin Market, Quantile Analysis*

Matching Indices for Thinly-Traded Commercial Real Estate in Singapore

1. Introduction

Unlike residential real estate markets where transactions are abundant, commercial real estate transactions are thin and lumpy. Many institutional owners hold commercial real estate for long-term investment purposes. The dearth of transaction data has led to the widespread use of appraisal based indices, such as the National Council of Real Estate Investment Fiduciaries (NCREIF) index, as an alternative to transaction-based indices in the U.S. However, appraisal-based indices are vulnerable to smoothing problems. Appraisers appear to systematically under-estimate the variance and correlation in real estate returns other asset returns (Webb, Miles and Guilkey, 1992). Despite various attempts to correct appraisal bias, it remains an *Achilles' heel* of appraisal-based indices. Corgel and deRoos (1999) found that recovering the true variance and correlation of appraisal-based returns reduces the weights of real estate in multi-asset portfolios.

Transaction-based price indices avoid appraisal smoothing problems. However, the number of private real estate transactions varies greatly over time. Investors hold onto their properties in “down” markets and sell them if they are priced above appraisal values in “up” markets. The volatility of transaction based indices is highly correlated with the liquidity (volume) in commercial real estate markets. Innovative methodologies have been proposed to correct for illiquidity in real estate transactions, including liquidity constant transaction based indices (Fisher, Gatzlaff, Geltner and Haurin, 2003), repeat sales regressions (RSR) adjusted for non-randomness in transactions (Gatzlaff and Haurin, 1997; Munneke and Slade, 2000 and 2001), among others.

This paper uses a propensity score matching technology to construct commercial real estate indices. The proposed matching indices are estimated based on comparable sales that occur in two different periods, one in a treatment period and another in a base (control) period. The matching approach has previously been used by McMillen (2012) and Deng, McMillen and Sing (2012) to estimate private residential real estate price indices in the U.S. and Singapore, respectively. Guo, Zheng, Geltner and Liu, (2012) propose a within-complex matching methodology, which pairs two different sales in the same building when constructing a pseudo repeat sale indices. Our matching indices differ from standard repeat sale price indices in three

ways. First, matching indices are not restricted to a small number of non-random repeat transactions when markets are illiquid. Second, the matching approach is less sensitive to changes in sample composition across transaction periods. Third, the matching indices have an advantage over the mean-based RSR estimator in explicitly accounting for variations across different price quantiles over time.²

Commercial real estate plays a key role in supporting various economic activities in Singapore, which is a global financial hub centered in Asia. Currently, the commercial real estate indices published by the Urban Redevelopment Authority (URA), a government agency overseeing urban development and planning, are the only publicly available indices for commercial real estate transactions in Singapore. The URA publishes quarterly indices for four commercial submarkets, including office, shop, multiple-user factory, and multiple-user warehouse. Despite the importance of commercial real estate markets, Tu, Yu and Sun (2004) is currently the only study devoted to developing robust commercial real estate indicators in Singapore. Tu, Yu and Sun (2004) use a spatiotemporal approach to model office price changes adjusted for different transaction activities in strata office space between the central and the suburban markets.

This paper develops matched-sample indices for three major commercial real estate submarkets in Singapore – office, shop, and factory – for the period 1995Q1-2010Q4. The matched price indices identify three distinct cyclical phases in commercial real estate prices in Singapore. Based on kernel density estimates of the price distributions, we observe significant differences in price trends over the two post-crisis periods in 1998 and 2008. The distributions of matched samples prices shifted to the right in the 2008 period. The quantile estimates reveal different price dynamics across the three submarkets. In the office market, the 90% percentile price index is more volatile than the index for the 10% percentile, which implies that the variation in sales prices in the top-tiered office market is highly sensitive to market shocks. The quantile distributions for factory prices shows a different picture, with both 10% and 90% percentiles price indices trending downward over the sample periods.

² McMillen (2012) finds that the upper end of housing price distribution in Chicago shifts further to the right than the lower end of distribution between 1995 and 2005.

The remainder of the paper is organized as follows. Section 2 reviews the literature on commercial real estate index methodologies and their limitations. Section 3 provides an overview of Singapore's commercial real estate market and the URA indices. Section 4 and 5 compare the standard RSR methodology with the matching approach. Section 6 summarizes the data used for our empirical estimation of commercial real estate indices in Singapore. Section 7 analyzes and compares the matching-based commercial real estate indices with other transaction-based indices. Section 8 concludes.

2. Commercial Real Estate Indices

The appraisal-based NCREIF property index and its predecessor, the Russell-NCREIF index, have been recognized as the *de facto* benchmarks of commercial real estate performance by institutional investors in the US (Fisher, Geltner and Webb, 1994; Geltner and Goetzmann, 2000; Pagliari, Lieblich, Schaner and Webb, 2001). The indices are estimated using appraised values of unleveraged commercial properties in the portfolios of NCREIF members. Appraisers' reliance on past information is a source of lagged errors that are embedded into current appraised values (Quan and Quigley, 1991; Cho and Megbolugbe, 1996; Chinloy, Cho and Megbolugbe, 1997; Clayton, Geltner and Hamilton, 2001; Diaz and Wolverton, 1998; Lai and Wang, 1998; Hansz and Diaz III, 2001; Bokhari and Geltner, 2011). Various de-lagging and de-smoothing procedures have been proposed to uncover the true volatility of appraisal-based indices (Geltner, 1989, 1991, 1993; Fisher, Geltner, and Webb, 1994; Cho, Kawaguchi and Shilling, 2003; Fu, 2003; An, Deng, Fisher and Hu, 2012, Bond, Hwang and Marcato, 2012).

While repeat sales regression (RSR) indices have been viewed as the gold standard for constructing house price indices for residential property markets, applications of RSR to illiquid commercial real estate indices are subject to serious disadvantages. First, the RSR sampling process restricts the data set to properties that sell twice within a sample period. This restriction reduces the size of the already thin commercial real estate sales sample.³ Second, potential sample selection bias occurs when more frequently transacted properties have higher price changes than the population of sample properties (Haurin and Hendershott, 1991;

³ Wu, Deng and Liu (2013) point out that a similar sampling limitation is faced in many booming nascent housing markets in emerging economies, where housing markets are dominated by new sales with single transactions only.

Munneke and Slade, 2000 and 2001). Third, asymmetry in the number of transaction activities (liquidity) during up and down markets is another potential source of non-randomness in the sample that distorts temporal changes in indices (Fisher, Gatzlaff, Geltner and Haurin, 2004).⁴ The first NCREIF transaction based index (NCREIF-TBI) was developed by the MIT Center for Real Estate in 2006 (Fisher, Geltner and Pollakowski, 2007). Derivations of monthly frequency commercial real estate indices were subsequently made possible by a frequency-conversion methodology proposed Bokhari and Geltner (2012), which can handle sparse and non-random commercial real estate transactions. Commercial real estate RSR indices constructed using the Bokhari-Geltner approach include Moody's/ Real Capital Analytics (RCA) commercial property price index (CPPI) and the CoStar commercial repeat-sales index (CCRSI) in the US.

Several attempts have been made to deal with thin market constraints in constructing RSR commercial real estate indices. Some use Bayesian estimators to simulate RSR commercial real estate indices from stock price data (Goetzmann, 1992; Kuo, 1997; Peng, 2002; Goetzmann and Peng, 2002). Goetzmann's ridge regression methodology was used in the original NCREIF-TBI. Other studies directly correct for heteroskedasticity in commercial real estate data using advanced econometric techniques (Schwann, 1998; McMillen and Dombrow, 2001; Hodgson, Slade and Vorkink, 2006; Graddy, Hamilton and Pownall, 2012). Heckman's two-stage methodology has also been used to correct for various types of selectivity bias in commercial sales. Munneke and Slade (2000 and 2001) correct for biases in unsold properties using a unique data set comprising both sold and unsold office properties in Phoenix. Gatzlaff and Haurin's (1997) selectivity adjusted RSR model uses property tax records from the Florida Department of Revenue to distinguish properties selling once from those selling twice. Fisher, Gatzlaff, Geltner and Haurin (2003) adjust for asymmetric liquidity in their proposed constant-liquidity RSR NCREIF commercial real estate indices. Spatial variations in price changes across different submarkets (Tu, Yu and Sun, 2004; Hayunga and Pace, 2010), information discovery between assessed value and transaction value (Gatzlaff and Holmes, 2011), and the small market capitalization effect (Gatzlaff and Holmes, 2011) are other issues discussed in the commercial real estate index literature.

⁴ Sales of starter or luxury homes (Englund, Quigley and Redfeare, 1999), and macro-economic shocks (Gatzlaff and Haurin, 1997) can cause variations in liquidity in the housing markets that can affect RSR price indices

This paper proposes a matching procedure for commercial real estate indices, where pairs of private commercial real estate transactions are matched using propensity scores derived from characteristics of the location and structure of properties that sold in the base period and later times. The matching process reduces bias by removing outliers and unrepresentative sales from the sample. The matching approach thus helps ensure that the estimated changes in price distributions capture true underlying market conditions. Matched pairs based on propensity scores are constructed from properties that were sold in the control (base) and treatment periods. The matching procedure avoids the backward adjustment problems found in revisions of hedonic and RSR indices (Clapham, Englund, Quigley and Redfearn, 2006; Deng and Quigley, 2008).⁵ In addition, matching indices are stable in thin trading markets and can be estimated using high-frequency data. If integrated into the two-stage frequency-conversion procedure of Bokhari and Geltner (2012), high-frequency matching indices could be derived to support tradable commercial real estate derivatives (Geltner and Ling, 2006).

⁵ Non-random sale sample also causes asymmetry resulting in more downward than upward revisions (Clapham, Englund, Quigley and Redfearn, 2006).

3. Singapore's Commercial Real Estate Markets and Indices

Singapore is an export-oriented economy that relies heavily on manufacturing and service industry outputs to generate its economic growth. As of 2011, 60% of the country's gross domestic product (GDP) is made up of goods and services produced by manufacturing (20%), wholesale and retail trade (16%), business services (13%), and finance and insurance (11%). The economic activities in the four key sectors create significant demand for commercial real estate. Private developers are the main suppliers of commercial real estate space. They operate in a *laissez-faire* market environment and are involved in a full range of development activities from land purchase, design and construction, financing and leasing activities. With the exception of the regulated sector in the provision of industry lands and single-user custom-built factories under the purview of the government's agency, Jurong Town Corporation (JTC), the public sector share in commercial real estate is relatively small: 19.5% office, 35.5% shop, and 22.0% multi-user factory stocks, based on the URA statistics on average stocks from 1Q2000 to 3Q20012. However, the Government is the main source of supply of development lands as it owns three-quarters of the land in Singapore.

Commercial real estate developers typically employ one of two strategies. First, developers usually adopt a "*build and hold*" strategy for prime-grade commercial properties, especially office buildings in central business districts and shopping malls along the tourist belt. These properties provide a steady stream of rental income and are held in developers' portfolios for long-term investment purposes. With the emergence of REITs in Singapore after July 2002, many developers sold their investment-grade investment properties to REITs. The second strategy involves "*build and sales*" of strata-titled commercial space⁶, such as condominium projects. Investors or space users who purchase undivided commercial space are issued strata-titles for the space, which are coupled with joint ownership rights of the land on which the property is annexed. Strata-title owners cannot sell joint interests in the land separately unless a resolution of a majority of owners is obtained. The process is known as *en bloc* sale, where strata-owners band together to jointly sell the land and property to a developer for

⁶ Unlike prime-grade commercial properties held by a single institutional investor, strata-titled space includes subdivided units and floors in commercial buildings that are sold separately to individual users/occupiers. The land on which the buildings are constructed is jointly owned by multiple strata-titled owners.

redevelopment purposes. The strata-commercial market is relatively small compared to prime commercial properties held in developers' and/or REITs' portfolios.

Currently, the only publicly available transaction based indices in Singapore for commercial real estate rely primarily on evidence from strata-sales evidence. Evidence of sales of strata commercial space is obtained from caveats lodged with the Registration of Land Title system in Singapore. The URA publishes four key indices, for offices, shops, multiple-user factories, and multiple-user warehouse.⁷ The URA also publishes sub-indices based on location for the four markets on a quarterly basis. The URA indices are unadjusted for quality, and they are median price indices that are weighted by a moving basket containing sales recorded in the last 12 months.

Figure 1 shows historical trends of the URA property price (nominal) indices for office, shop and multiple-user factory for 1990Q1 - 2011Q2. The base period is 1998Q4 = 100. The three markets move together, though short-term variations are observed over the sample period. In general, Singapore's commercial real estate market witnessed three distinct peaks in 1996Q3, 2000Q4, and 2008Q2. The positive economic growth and strong influx of funds into the property market coupled with low interest rates resulted in a buoyant property market in the mid-1990s. Fueled by exuberance in private residential markets, prices in the commercial real estate markets also increased significantly, reaching a peak in 1996. Over a 3-year period from 1993Q3 to 1996Q3, the property price indices for the office and industry sectors increased by 83.5% and 94.6%, respectively. The Government's intervention into the private residential market followed by the Asian Financial Crisis in 1997 caused prices of the commercial real estate to plunge deeply into recession in 1999. Two short booms were experienced in 2000 and 2008. The rise in prices in 2008 was derailed by the Subprime Crisis in the US. A strong rebound in all sectors was observed in 2009, with the industrial sector showing the strongest growth as the industry property price index surpassed the office and shop indices in 2011.⁸

[Insert Figure 1 here]

⁷ Warehouse index is excluded in our analysis due to limited transactions in this submarket.

⁸ The government has for the time ever intervene into the industry market on 11 January 2013 by introducing Seller's Stamp Duty (SSD) of 15%, 10% and 5% on industry lands and properties sold in the first year, second year and third year from the date of purchase, respectively.

The URA commercial indices are Laspeyres median price indices. Commercial real estate transactions in each quarter are first grouped by property type and locality, and the unit medium prices (\$ per square meter) for each group are used to compute different sub-indices. Based on the sub-indices, the indices for different sub-markets (including the three commercial sub-markets analyzed in this paper, office, shop, and factor) are computed using a weighted average of property sub-indices in different planning areas. Prior to the 4th Quarter of 1998, the weights are derived from the value of transactions for each property type in each locality for a fixed base year (= 1990). Current price indices are computed based on the moving average of the value of transactions over the previous 12 quarters. The weights in the price indices are therefore updated quarterly so that they are as current as possible.

The existing URA median price commercial real estate indices are subject to some technical limitations. The moving average may cause persistence in the quarterly revisions of the indices. Smoothing in the median price indices may occur if trading activities of the strata-titled commercial real estate are thin in some areas.. Since the indices are not quality adjusted, price changes may be dominated by either low priced or poor quality properties in some periods.

In this study, we propose to construct matching-based quality-adjusted price indices for offices, shops, and multiple-user factories using the transaction data in the URA Realis database. The matched sample approach is well-suited to dealing with problems related to thin markets and heterogeneous quality in the commercial real estate market. The procedure of matching sale samples in the base period to sales samples in the treatment periods using observed property characteristics is much more flexible than the standard repeat sales estimator. The matching approach eliminates outliers that could distort the estimations. It also expands the matched sample size by not limiting the sample to sales of the commercial property over time.. Another advantage of the matching approach is that it offers not just point estimates of mean or median price changes; it can stratify the price changes into different quantiles of the price distribution.

4. Transaction Based Index Methodologies

A typical hedonic price function expresses the relationship between real estate sales prices and characteristics of the structure, location, and time of sale:

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

where $y_{i,t}$ is the natural logarithm of the sale price of property i at time t ($t = 1, \dots, T$), $D_{i,t}$ is a time dummy that has a value of 1 if the property i sold at time t , $X_{i,t}$ is a set of observed attributes of the structure and location, and $Z_{i,t}$ is a set of unobserved characteristics that influence the sale price. The hedonic price equation (1) can be estimated directly using pooled data on sales over the full sample period. The vector $\delta = (\delta_1 \dots \delta_T)$ forms the quality-adjusted price index. Functional form misspecification and omitted variables correlated with the timing of sales can produce biased estimates of δ .

Bailey, Muth and Nourse (1963) and Case and Shiller (1987, 1989) introduced an innovative way of reducing omitted variable bias by restricting the sample to properties that have sold at least twice in the sample period.⁹ Subtracting the log-price of a property sold at time t from its log-price at an earlier time s leads to the following equation:

$$y_{it} - y_{is} = (\delta_t D_{it} - \delta_s D_{is}) + (\beta_t' X_{it} - \beta_s' X_{is}) + (\lambda_t' Z_{it} - \lambda_s' Z_{is}) + (u_{it} - u_{is}), \quad (2)$$

If both the coefficients and the values of X and Z do not change over time, the difference in log-prices is a function solely of the dates of sale:

$$y_{it} - y_{is} = (\delta_t D_{it} - \delta_s D_{is}) + (u_{it} - u_{is}) \quad (3)$$

Equation (3), which is the standard repeat sale equation, provides unbiased estimates of the coefficients for date of sale if X , Z , β , and λ are all constant over time. The estimates of δ will be biased if these assumptions are not met, i.e., 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 renovation and concentration of sales in specific locations with high appreciation rates are sources of omitted variables problems that cause bias in the estimation.

⁹ From commercial real estate investors' perspective, the repeat sale of a property is analogous to the realized trading returns of stock investors, where they could directly measure the "round-trip" experiences to date for completed investments, i.e. they must sell the same properties that they buy in the period covered by the index. Thanks for the comments by the referee.

5. Matching Strategy

The analogy between the repeat sales estimator and a matching model is illustrated using the following simple two-period version of the model with constant coefficients and housing characteristics:¹⁰

$$y_i = \delta_1 + (\delta_2 - \delta_1)D_{i,2} + \beta' X_i + \lambda' Z_i + u_{i,1}, \text{ or} \quad (4a)$$

$$y_{i,2} - y_{i,1} = (\delta_2 - \delta_1)D_{i,2} + u_{i,2} - u_{i,1} \quad (4b)$$

The hedonic approach directly regresses y_i on $D_{i,2}$, X_i , and Z_i . The repeat sales model is simply a regression of $(y_{i,2} - y_{i,1})$ on $D_{i,2}$ without the intercept. The time dummy $D_{i,2}$ estimates the difference in average log sales prices between period 1 and 2 for properties that have been sold in both periods.¹¹ Like the standard difference-in-means test, this difference is written as an *average treatment effect* (ATE) as follows: $ATE = \frac{1}{n_2} \sum_{i=1}^{n_2} D_{i,2} E[y_i(t_2) - y_i(t_1)]$, which represents the expected difference in sale price between the first and second period for the sample of property 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.¹² Since period 1 serves as the base for the price index, this expression applies directly to subsequent time periods:

$$ATE(t_j) = \frac{1}{n_j} \sum_{i=1}^{n_j} D_{ij} E[y_i(t_j) - y_i(t_1)] \quad (5)$$

The expected change in sale price relative to the base period for those properties that actually sold in time t_j is represented by equation (5), which is equivalent to a Laspeyres price index.

¹⁰ For simplicity, the time subscript t is dropped from the equations.

¹¹ Unlike RSR, the hedonic model uses all samples transacted in the base and the index periods in estimation. A time dummy, $D_{i,2}$, is used to capture the price change effects between the two periods after controlling for hedonic attributes of the sample sales.

¹² The literature on treatment effects is large and growing. Excellent overviews are presented in Ho, Imai, King and Stuart. (2007) and Imbens and Wooldridge (2009).

The repeat sales estimator leads directly to equation (4b) by removing the effects of X and Z by restricting the sample to properties that have sold twice over time. However, a difference in means approach can also be used to estimate average treatment effects when matches are not perfect. The requirements for estimating average treatment effects are less stringent than the requirements for unbiased estimates of a full set of parameters. Matching procedures pair each property with a similar property that sold in a different time period. If sales do not take place every period for all properties, but sales prices are appreciating at the same rate across property 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.¹³ Differences between matched observations will average out over many observations leading to an accurate estimate of the rate of appreciation across time periods.¹⁴ McMillen (2010) points out that the repeat sales approach can be considered to be an extreme version of a matching estimator.

Based on a propensity score approach, where “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 a set of predicted values from a probit or logit regression of I_t on the observed set of housing and location characteristics. Each observation in the base period is matched with an observation from the later time period that had a similar probability of sale.¹⁵ Our samples cover the periods from 1995Q1 to 2010Q4, and we use 1995Q1 as the base time. We use the following algorithm to construct the matched samples:

1. For each quarter q from 1995Q2 to 2010Q4 (excluding the base period 1995Q1), we estimate a logit model using all sales taking place in 1995Q1 and q . The dependent variable equals one if the sale took place in quarter q and zero if the sale is from 1995Q1.

¹³ As pointed out by a referee, the implicit price appreciation is analogous to the realized investment returns of a “round-trip” investor’s experience when the buy and sell events occurring in the two periods coincide with the index matching periods.

¹⁴ Following the approach of a Laspeyre price index, the matching procedure controls for quantity effects by pairing sales in different treatment periods to the same properties in the base period.

¹⁵ Ho, Imai, King and Stuart (2007) present a particularly useful discussion of the practical issues involved in matching estimators. Their “MatchIt” routine in R (Ho, Imai, King and Stuart, 2011) constructs matches very rapidly for a wide variety of algorithms. Good examples of matching estimator applications 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).

The explanatory variables for the logit regressions are the same as those used for the hedonic price function estimates.

2. We then use the estimated propensity score from each logit regression to match n_l observations from quarter q to sales from 1995Q1, where n_l is the number of sales in 1995Q1. Based on a random ordering of the 1995Q1 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 $63n_l$ observations – n_l matched sales for each quarter from 1995Q2 to 2010Q4, plus the n_l sales from the base year 1995Q1 (“approximately” because some observations may not have a close match).

After constructing the matched samples, the estimated price index is simply the series of differences between the average log-sales prices at time t and the average value in the base period, 2000Q1. Wang and Zorn (1997) demonstrate that a period-by-period comparison of average log sale price is *exactly* equivalent to the repeat sales model when applied to a data set with the same number of observations in each sample period. The repeat sales estimator controls for housing characteristics by matching each home to itself across time; the matching estimator accomplishes a similar objective by balancing the distributions of housing characteristics over time by pairing each observation with similar observations from other time periods. By eliminating extreme observations, the differences in housing characteristics average out, thereby leaving simple differences in means as an unbiased estimate of quality-constant price differences. Similar in spirit to the repeat sales estimator, period-by-period means from matched sample estimates are likely to be more efficient because sample sizes are larger when single-sale properties are included.

The evolution of sale price distributions can be sorted further into different quantiles. McMillen (2012) uses this approach to compare the distribution of Chicago’s housing 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.

6. Data

This study uses commercial real estate transaction data from the URA's Real Estate Information System (REALIS), which is the same data source used to construct the URA commercial real estate indices. The transactions data cover the sample period from 1995Q1 to 2011Q2. A total of 22,842 sales were retrieved from the database, which comprise 4,646 office, 6,601 shop, and 11,595 multiple-user factory sales. We pool transactions from units sold at least twice during the sample period, which account for 2,434, 3,526 and 4,486 of the repeat sale pairs for office, shop and factory, respectively. For the matched pairs, we match 1,622, 4,103 and 7,298 over 65 n_t periods for the same three submarkets (office, shop and factory) over the sample periods from 1995Q2 to 2011Q2 using the sales in 1995Q1 as the base.

The property-level information recorded in the REALIS database includes the project name, address, floor area (m²), transaction price (SGD), contract date, tenure, postal district, and planning region. The properties sold are divided into five major planning regions: Central, East, North-East, North and West. We add several measures of location attributes, including accessibility to the Central Business District (CBD), the airport, seaport, the nearest expressway, and the nearest mass rapid transit (MRT) or light rail transit (LRT) station. All accessibility variables are measured as linear distances to the sample property in kilometers.¹⁶ We convert the linear distances into the natural logarithm forms to account for potential nonlinearity in their effects on sales prices. We measure neighborhood attributes using discrete variables, which have a value of 1 if the subject property is located within a 0.3 km radius of an amenity. The amenities include proximity to CBD, MRT/LRT stations, expressways, public housing estates, and shopping centers. The binary variables indirectly control for externalities associated with neighborhoods. For example, proximity to public housing estates may generate a positive externality to factories because residents in the housing estates are an important source of labor supply for manufacturing firms in factories. A dummy variable indicating the availability of a car-park within a building is also included in the regressions.

¹⁶ As Singapore is a densely built-up city state connected by closely knitted transportation network, linear distances to the employment center and the stations along the mass rapid transits and the light rail transit lines are good approximations of the travel distances to the respective nodes.

Table 1 lists the variables used in the empirical analysis along with descriptive statistics (mean and standard deviation). Figure 2 shows the number of observations by quarter for the three commercial real estate submarkets. The highest sales volumes in all three markets were recorded in 2007, prior to the US subprime crisis. The sales volumes also peaked in the three other earlier booms in 1996 and 2003. For the factory sector, 1999 also saw a delayed rise in sale volumes, shortly after the 1997 Asian Financial Crisis.

[Insert Table 1 and Figure 2]

7. Commercial Real Estate Price Indices for Singapore: 1995Q1-2010Q4

7.1. Implicit Pricing for Hedonic Attributes

We regress the natural logarithm of sale price on vectors of property, location and neighborhood characteristics, along with fixed effects for time of sale and the planning region. The regression estimates for the full-sample hedonic model and the matched sample model are summarized in Table 2. The adjusted R^2 ranges from 0.625 for the matched sample model for the shop sector (Model 3) to 0.927 for the full sample model for offices (Model 2). In general, the transaction data fit the office models the best, whereas the goodness of fit for the shop models is relatively lower. The regional fixed effects offer only marginal increases in explanatory power for commercial real estate prices in both the full sample and the matched sample models. Most of the estimated coefficients are highly significant and consistent across the full sample and matched sample models.

[Insert Table 2 here]

The coefficients for log-floor area indicate positive economies of scale effects for commercial real estate. The accessibility (distance) effects vary by property type, transportation connectivity, and business network. Distances to the airport and expressways have positive effects on commercial real estate prices, whereas proximity to seaport has a negative effect. Distances to the nearest MRT and LRT stations and to public housing estates have a positive effect on sales prices for factory users, but they have a negative effect on office and shop prices. These results may imply that labor supply and mobility are more important to factory users than to office and shop users. Given that our shop and office samples consist mainly of smaller

strata units, discounts to the two factors may reflect keen competition from other large office buildings and shopping malls near the MRT/LRT stations and in high catchment public housing estates. These large scale commercial properties are usually owned by developers and REITs, which have more resources for attracting tenants and shoppers. The neighborhood variable indicating a MRT/LRT station within a 0.3km radius shows similar results, with positive coefficients for offices and shops, and negative coefficients for the factory sector.

Office users value proximity to CBD positively partly because of network effects with business associates and clients. The coefficients on log-distance to CBD and the inside=CBD dummy are both positive and significant in the office models, but the coefficients are negative for factories. High land costs in the CBD partly explain the preference for suburban factory locations. The effects of proximity to CBD are mixed for shops. However, sales prices of shops are higher within 0.3km of other shopping centers. The coefficient on the availability of a car-park is highly significant and positive for shops; whereas car-parks lead to lower prices for factories, presumably due to the high cost of allocating car-park space in for a factory. The “car-park in a building” coefficient has mixed signs for the office sector. The results show heterogeneity in preference for location and neighborhood amenities among the three commercial real estate users.

7.2. *Price Indices*

The hedonic, repeat sale sample, and matched sample price indices are constructed using the estimated coefficients for the quarterly dummy variables, after normalizing the 1995Q1 index to zero. The repeat sales index is constructed by regressing the change in the natural log of sale price on a series of 63 indicator variables that equal -1 if the earlier sale occurs at the time, 1 for the second sale in the pair, and 0 if property did not sell at the time.

Figure 3 plots the hedonic, repeat sale, and matched sample indices along with the full sample and matched sample means for the three commercial real estate submarkets. The URA median price indices are also included in the figures. The full sample mean and the matched sample mean are more volatile than the URA median price indices. The indices estimated from the three different methodologies (hedonic, repeat sales, matched sample regression) also generate more short-term noise in the price trends. The office indices seem to closely track the URA

median price index, except for the periods after 2007. The three quality-adjusted shop indices trend above the URA median price indices after 1998. In the factory submarket, the URA indices drifted above the three indices throughout the time between 2000 and 2008.

Despite some variation among the indices, the quality-adjusted indices all indicate three distinct phases of cyclical fluctuations in prices over the sample period. Prices rose between 1995 and 1996. The 1997 Asian Financial Crisis dampened market confidence, causing prices to fall in all three commercial submarkets between 1997 and 2005. Commercial real estate prices recovered in mid2006, but the upward momentum was disrupted by the U.S. Subprime Crisis in 2007, which led to a decline in prices in 2008 and 2009. Prices rose again in the commercial real estate sub-market in 2009, with the rate of appreciation in the office and shop submarkets equaling, if not surpassing, their former peaks in 1996. The rate of decline in prices in the factory submarket has slowed since its nadir in 2006, but the prices continue to decline somewhat.

[Insert Figure 3 here]

Based on the number of transactions shown in Table 1, the factory submarket was the most active market followed by shops and offices. The level of market activity in the three submarkets appears to have a significant influence on price changes. For the office submarket, which has the smallest number of observations in the repeat and matched samples, the hedonic, repeat sale, and matched regression indices move closely over the sample period (Figure 3a). The full sample mean and matched sample means are highly variable in the office market when transactions are thin. In Figure 3b, where shop transactions are higher at 6,601 (with 3,526 repeat sales and 4,103 matched samples), all three indices and the matched sample means overestimate the full-sample means and the URA median indices. For the most actively traded factory submarket, the variation between the mean indices and the quality-adjusted indices are narrower; but the growth rates of the indices underestimate the full sample means and the median price growth of the URA indices (Figure 3c).

7.3. *Comparison of matched and unmatched samples*

The matching sample approach includes properties in the treatment periods that have the closest matches with properties in the base period as indicated by the propensity matching scores. The method keeps the quality of sample sales relatively constant, which is consistent with the RSR methodology that assumes no change in quality between the first period and the second period sales. The matching process is used to remove outliers and observations that are not comparable to those in the base periods. Table 3 presents descriptive statistics for the two samples. The results show that the differences in the means for the matched and unmatched indices are 0.07 and 0.139 for the office and shop submarkets, respectively. The difference is largest for the multiple-user factor submarket, which is estimated at 0.119.

[Insert Table 3 here]

The unmatched sample indices are more volatile compared with the matched sample indices and the URA median price indices as indicated by both the standard deviation and the range. The unmatched sample indices and the URA median price indices are highly positively skewed. For the matched sample indices, the office sub-index has a negative skewness of -0.074, and the positive skewness in the multiple-user factory sub-index is also relatively smaller. The kurtosis measure indicates that the unmatched sample indices have high “peakness” in the distributions, which is positively correlated with variances of the indices.

Figure 4 plots the matched sample indices, the unmatched sample indices, and the matched sample means. The unmatched sample indices show more quarter-to-quarter variation in the rate of price appreciation compared with the matched sample indices. However, the results show that the two indices closely track the major peaks and troughs throughout the full sample period.

[Insert Figure 4 here]

7.4. *Quantile distributions and kernel density of price changes*

Whereas the standard repeat sales estimator focuses on mean sale price, the matching approach can be used to show how the entire distribution of log sale price varies over time. Figure 5

presents kernel density estimates of the log sale price distributions for 1998 and 2008 for the full sample and the combined matched samples for the three commercial markets. The high density of sales prices in 2008 implies that the rate of price appreciation was much higher for high-priced properties than for low-priced properties during the two periods. The same effects were also observed in residential markets in the earlier studies by McMillen (2012) and Deng, McMillen and Sing (2012).

[Insert Figure 5 here]

Figures 6 and 7 illustrate the advantages of the matching approach relative to a mean or median based price index estimator. Figure 6 shows the distributions of the natural log of sale price in the full sample and the matched samples. The lines represent the 10%, 25%, 50%, 75%, and 90% percentiles of the distribution. Figure 7 shows the 10% and 90% percentiles distributions of price indices estimated using the full sample and the matched samples. We see more noise in the 90% percentile price distributions in the office sub-market relative to the 10% percentile price distributions. For the shop and factory submarkets, the log-prices fluctuate within a relatively tighter range in the shop market over the sample period. The 90% percentile factory indices show a large increase in prices between 1995 and 1997.

[Insert Figure 6 here]

[Insert Figure 7 here]

8. Conclusion

Dealing with non-random transaction samples in thin trading commercial real estate market has long been a challenge for researchers. Various advanced econometric methodologies have been applied to resolve biases in transaction based indices. Unlike the RSR methodology, which restricts the sample to pairs of identical properties sold at least twice over time, matching procedures offer an alternative way of dealing with problems associated with thin markets and non-random samples. The matches need not be the same property. The matching approach is particularly useful in thin markets in which the number of repeat sales is limited. The proposed matching algorithms pair the control ('base') observations with similar properties with sales in later periods. By keeping the same sample size in the control (base) and treatment periods, the

covariates capture only the temporal effects of implicit prices for the observed attributes. The matching process helps to control for selectivity bias by balancing covariates across time periods.

The conventional repeat sale approach to price index construction is a special case of a matching model. The only difference between the repeat sales estimator and period-by-period differences in means is that the average treatment effects in the repeat sales estimator are constructed from a sample that is limited to properties that have sold at least twice over the sample period. Since a transaction based matched sample price index is constructed from a much larger sample, a small number of unrepresentative repeat sales will not have an undue influence on the estimated price index. In addition to point estimates (median and mean), percentiles from any point in the distribution can also be generated to analyze within sample variations in overall distribution of sales prices over time.

In this paper, we construct matched sample commercial real estate price indices for offices, shops, and multiple-user factories using transaction data from Singapore for the period 1995Q1-2010Q4. The matched sample indices identify three distinct cyclical phases in price appreciation in the commercial real estate market in Singapore. The matched indices closely track the quality-unadjusted means and hedonic indices, which are estimated from the full sample. We observe that the distribution of sales prices in all three commercial submarkets shifted to the right between 1998 and 2008. The quantile estimates of the matched sample indices also show large variation in the price distributions in each of the three submarkets. The 90% percentile price index was more volatile than the 10% price index for the office sector. The 10% and 90% percentile price indices were more stable for shops and factories. The quantile results provide a much more complete picture of the evolution of sales prices in commercial property than a median or mean-based index.

References

- An, X., Y. Deng, J.D. Fisher and M. Hu. 2012. Commercial Real Estate Rental Index: A Dynamic Panel Data Model Estimation. *SSRN working paper wps 2205058*, <http://ssrn.com/abstract=2205058>.
- 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.
- Bokhari, S. and D. Geltner. 2011. Loss Aversion and Anchoring in Commercial Real Estate Pricing: Empirical Evidence and Price Index Implications. *Real Estate Economics* 39(4): 635-670.
- Bokhari, S. and D. Geltner. 2012. Estimating Real Estate Price Movements for High Frequency Tradable Indexes in a Scarce Data Environment. *Journal of Real Estate Finance and Economics* 45(2): 522-543.
- Bond, S.A., S. Hwang and G. Marcato. 2012. Commercial Real Estate Returns: An Anatomy of Smoothing in Asset and Index Returns. *Real Estate Economics*
- Bondonio, D. and J. Engberg. 2000. Enterprise zones and local employment: evidence from the states' programs. *Regional Science and Urban Economics* 30: 519-459.
- Bondonio, D. 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.
- Case, K. and R. Shiller. 1987. Prices of Single Family Homes since 1970: New Indexes for Four Cities. *New England Economics Review* September/October: 45-56.
- Chinloy, P., M. Choo and I. F. Megbolugbe. 1997. Appraisal, Transaction Incentives and Smoothing. *Journal of Real Estate Finance and Economics* 14(1): 89-112.
- Clayton, J., D. Geltner and S.W. Hamilton. 2001. Smoothing in Commercial Property Valuations: Evidence from Individual Appraisals. *Real Estate Economics* 29(3): 337-360.
- Chinloy, P., M. Cho and I.F. Megbolugbe, 1997. Appraisal, Transaction Incentives, and Smoothing. *Journal of Real Estate Finance and Economics* 14: 89-111.
- Cho, M. and I. F. Megbolugbe. 1996. An Empirical Analysis of Property Appraisal and Mortgage Redlining. *Journal of Real Estate Finance and Economics* 13(1): 45-55.
- Cho, H., Y. Kawaguchi and J.D. Shilling. 2003. Unsmoothing Commercial Property Returns: A Revision to Fisher-Geltner-Webb's Unsmoothing Methodology. *Journal of Real Estate Finance and Economics* 27(3): 393-405.
- Clapham, E., P. Englund, J.M. Quigley and C.L. Redfearn. 2006. Revisiting the Past and Setting the Score: Index Revision for House Price Derivatives. *Real Estate Economics* 34(2): 275-302.
- Clayton, J., D. Geltner and S.W. Hamilton. 2001. Smoothing in Commercial Property Valuations; Evidence from Individual Appraisals. *Real Estate Economics* 29(3):337-360.
- Corgel, J. and J. deRoss. 1999. Recovery of Real Estate Returns for Portfolio Allocations. *Journal of Real Estate Finance and Economics* 18(3): 279-296.
- Deng, Y., D.P. McMillen and T.F. Sing. 2012. Private Residential price Indices in Singapore: A Matching Approach," *Regional Science and Urban Economics* 42 (3): 485-494
- Deng, Y. and J.M. Quigley. 2008. Index revision, house price risk and the market for house price derivatives. *Journal of Real Estate Finance and Economics* 37: 191-209.
- Diaz, J. and M. Wolverton. 1998. A Longitudinal Examination of the Appraisal Smoothing Hypothesis. *Real Estate Economics* 26(2): 349-358.
- Englund, P., J.M. Quigley and C.L. Redfearn. 1999. The choice of methodology of computing housing price indexes: comparisons of temporal aggregation and sample definition. *Journal of Real Estate Finance and Economics* 19(2): 91-112.

- Fisher, J., D. Gatzlaff, D. Geltner and D. Haurin. 2003. Controlling for the Impact of Variable Liquidity in Commercial Real Estate Price Indices. *Real Estate Economics* 31(2): 259-303.
- Fisher, J., D. Gatzlaff, D. Geltner and D. Haurin. 2004. An Analysis of the Determinants of Transaction Frequency of Institutional Commercial Real Estate Investment Property. *Real Estate Economics* 32(2): 239-264.
- Fisher, J., D. Geltner and R. Webb. 1994. Value Indices of Commercial Real Estate: A Comparison of Index Construction Methods. *Journal of Real Estate Finance and Economics* 9: 137-164.
- Fisher, J., D. Geltner and H. Pollakowski. 2007. A Quarterly Transactions-based Index of Institutional Real Estate Investment Performance and Movements in Supply and Demand. *Journal of Real Estate Finance and Economics* 34: 5-33.
- Fu, Y. 2003. Estimating the Lagging Error in Real Estate Price Indices. *Real Estate Economics* 31(1): 75-98.
- Gatzlaff, D.H. and D.R. Haurin. 1997. Sample selection bias and repeat-sale index estimates. *Journal of Real Estate Finance and Economics* 14: 33-50.
- Gatzlaff, D. and C. Holmes. 2011. Estimating Transaction-Based Price Indices of Local Commercial Real Estate Markets Using Public Assessment Data. *Journal of Real Estate Finance and Economics*.
- Geltner, D. 1989. Estimating Real Estate's Systematic Risk from Aggregate Level Appraisal-Based Returns. *AREUEA Journal* 17(4): 463-481.
- Geltner, D. 1991. Smoothing in Appraisal-Based Returns. *Journal of Real Estate Finance and Economics* 4(3): 327-345.
- Geltner, D. 1993. Estimating Market Values from Appraisal Values Without Assuming an Efficient Market. *Journal of Real Estate Research* 8(3):325-346.
- Geltner, D. and W. Goetzmann. 2000. Two Decades of Commercial Property Returns: A Repeated-Measures Regression-Based Version of the NCREIF Index. *Journal of Real Estate Finance and Economics* 21(1): 5-21.
- Geltner, D. and D. Ling. 2006. Considerations in the Design & Construction of Investment Real Estate Research Indices. *Journal of Real Estate Research* 28(4): 411-444.
- Goetzmann, W.N. 1992. The Accuracy of Real Estate Indices: Repeat-Sale Estimators. *Journal of Real Estate Finance and Economics* 5(1): 5-54.
- Goetzmann, W.N. and L. Peng. 2002. The Bias of the RSR Estimator and the Accuracy of Some Alternatives. *Real Estate Economics* 30(1): 13-39.
- Graddy, K., J. Hamilton and R. Pownall. 2012. Repeat-Sales Indexes: Estimation without Assuming that Errors in Asset Returns Are Independently Distributed. *Real Estate Economics* 40(1):131-166.
- Guo, X., S. Zheng, D. Geltner and H.Liu. 2012. A New Approach for Constructing Home Price Indices in China: The Pseudo Repeat Sales Model. SSRN Working Paper, <http://ssrn.com/abstract=2127336>.
- Hansz, J.A. and J.Diaz III. 2001. Valuing Bias in Commercial Appraisal: A Transaction Price Feedback Experiment. *Real Estate Economics* 29(4): 553-565.
- Hayunga, D.K. and R.K. Pace. 2010. Spatial Statistics Applied to Commercial Real Estate. *Journal of Real Estate Finance and Economics* 41: 103-125.
- 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.
- Ho, D., K. Imai, G. King and E. Stuart. 2011. MatchIt: Nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software* 42(8) :1 -28.
- Hoag, J.W. 1980. Towards indices of Real Estate Value and Return. *Journal of Finance* 35(2): 569-580.

- Hodgson, D.J., B.A. Slade and K.P. Vorkink. 2006. Constructing Commercial Indices: A Semiparametric Adaptive Estimator Approach. *Journal of Real Estate Finance and Economics* 32:151-168.
- Imbens, G. and J. Wooldridge. 2009. Recent developments in the econometrics of program evaluation. *Journal of Economic Literature* 47, 5-86.
- Kiel, K., Zabel, J. 1997. Evaluating the usefulness of the American Housing Survey for creating house price indices. *Journal of Real Estate Finance and Economics* 14, 189-202.
- Kuo, C. 1997. A Bayesian Approach to the Construction and Comparison of Alternative House Price Indices. *Journal of Real Estate Finance and Economics* 14: 113-132.
- Lai, T.Y. and K. Wang. 1998. Appraisal Smoothing: The Other Side of the Story. *Real Estate Economics* 26(3): 511-535.
- List, J., McHone, W., Millimet, D. 2004. Effects of environmental regulation on foreign and domestic plant births: is there a home field advantage? *Journal of Urban Economics* 56, 303-326.
- McMillen, D.P. 2012. Repeat Sales as a Matching Estimator. *Real Estate Economics* 40(4): 743-771.
- McMillen, D.P. and J. Dombrow. 2001. A Flexible Fourier Approach to Repeat Sales Price Indexes. *Real Estate Economics* 29(2): 207-225.
- McMillen, D. and J. McDonald. 2002. Land values in a newly zoned city. *Review of Economics and Statistics* 84: 62-72.
- Miles, M., R.Cole and D.Guilkey. 1990. A Different Look at Commercial Real Estate Returns. *Journal of American Real Estate and Urban Economics Association* 18(4): 403-430.
- Munneke, H. J. and B.A. Slade. 2000. An Empirical Study of Sample-Selection Bias in Indices of Commercial Real Estate. *Journal of Real Estate Finance and Economics* 21(1): 45-64.
- Munneke, H. J. and B.A. Slade. 2001. A Metropolitan Transaction-Based Commercial Price Index: A Time-Varying Parameter Approach. *Real Estate Economics* 29(1): 55-84.
- 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.
- Pagliari, J. Jr, F. Lieblich, M. Schaner and J. R. Webb. 2001. Twenty Years of the NCREIF Property Index. *Real Estate Economics* 29(1):1-27.
- Palmquist, R. 1980. Alternative techniques for developing real estate price indexes. *Review of Economics and Statistics* 66: 394-404.
- Peng, L. 2002. GMM Repeat Sales Price Indices. *Real Estate Economics* 30(2): 239-261.
- Quan, D.C. and J.M. Quigley. 1991. Price Formation and the Appraisal Function in Real Estate Markets. *Journal of Real Estate Finance and Economics* 4:127-146.
- 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.
- Schwann, G. 1998. A Real Estate Price Index for Thin Markets. *Journal of Real Estate Finance and Economics* 16(3): 269-287.
- Tu, Y., S.M. Yu and H. Sun. 2004. Transaction-Based Office Price Indexes: A Spatiotemporal Modeling Approach. *Real Estate Economics* 32(2):297-328.
- Wang, F., Zorn, P. 1997. Estimating house price growth with repeat sales data: what’s the aim of the game? *Journal of Housing Economics* 6, 93-118.
- Webb, R.B., M. Miles and D. Guilkey. 1992. Transactions-Driven Commercial Real Estate Returns: The Panacea to Asset Allocation Models? *Journal of American Real Estate and Urban Economics Association* 20(1): 325-357.

Wu, J., Y. Deng and H. Liu. 2013. House Price Index Construction in the Nascent Housing Market: The Case of China. Forthcoming in *Journal of Real Estate Finance and Economics*, DOI 10.1007/s11146-013-9416-1.

Table 1: Descriptive Statistics

Descriptive Statistics	Office			Shop			Multiple-user Factory		
	Full	Repeat Sales	Matched	Full	Repeat Sales	Matched	Full	Repeat Sales	Matched
Number of Observations	4646.00	2434.00	1622.00	6601.00	3526.00	4103.00	11595.00	4486.00	7298.00
A) Property Attributes:									
Log Sale Price (in Singapore Dollars)	13.63 [1.20]	13.52 [1.09]	13.26 [1.04]	13.26 [0.91]	13.15 [0.82]	13.20 [0.86]	13.22 [0.72]	13.18 [0.71]	13.29 [0.72]
Log Floor Area of Unit (in square meters)	4.42 [1.00]	4.28 [0.87]	4.18 [0.85]	3.72 [0.85]	3.65 [0.73]	3.63 [0.77]	5.39 [0.73]	5.36 [0.74]	5.47 [0.71]
B) Location Attributes (km):									
Log Distance to CBD	0.75 [0.92]	0.76 [0.87]	0.71 [0.75]	1.17 [0.85]	1.11 [0.79]	1.07 [0.81]	2.25 [0.49]	2.21 [0.49]	2.22 [0.52]
Log Distance to Airport	2.78 [0.16]	2.79 [0.15]	2.77 [0.17]	2.77 [0.25]	2.76 [0.24]	2.78 [0.23]	2.81 [0.44]	2.78 [0.44]	2.72 [0.40]
Log Distance to Seaport	1.06 [0.66]	1.04 [0.64]	1.17 [0.67]	1.59 [0.58]	1.55 [0.54]	1.53 [0.53]	2.38 [0.46]	2.34 [0.47]	2.38 [0.49]
Log Distance to nearest MRT or LRT Station	-1.32 [1.22]	-1.34 [1.22]	-0.99 [0.85]	-1.16 [1.04]	-1.14 [1.03]	-1.22 [1.05]	0.02 [0.63]	0.02 [0.63]	-0.01 [0.63]
Log Distance to Expressway	-0.65 [0.71]	-0.67 [0.72]	-0.53 [0.59]	-0.42 [0.77]	-0.42 [0.74]	-0.38 [0.70]	-0.59 [1.27]	-0.69 [1.26]	-0.44 [1.24]
C) Neighborhood Attributes (Yes =1; No =0)									
Inside of CBD	0.52 [0.50]	0.52 [0.50]	0.54 [0.50]	0.28 [0.45]	0.31 [0.46]	0.29 [0.45]	0.00 [0.00]	0.00 [0.00]	0.00 [0.00]
MRT or LRT Station within 0.3km	0.92 [0.27]	0.94 [0.25]	0.93 [0.26]	0.87 [0.34]	0.87 [0.34]	0.89 [0.31]	0.32 [0.47]	0.34 [0.47]	0.36 [0.48]
Expressway with 0.3km	0.88 [0.33]	0.90 [0.30]	0.89 [0.31]	0.71 [0.45]	0.72 [0.45]	0.71 [0.46]	0.62 [0.49]	0.64 [0.48]	0.59 [0.49]
Shopping Center within 0.3km	1.00 [0.02]	1.00 [0.00]	1.00 [0.00]	0.98 [0.14]	0.98 [0.13]	1.00 [0.06]	0.34 [0.47]	0.35 [0.48]	0.31 [0.46]
Public Housing within 0.3km	0.84 [0.37]	0.83 [0.37]	0.85 [0.36]	0.82 [0.39]	0.80 [0.40]	0.80 [0.40]	0.69 [0.46]	0.75 [0.43]	0.74 [0.44]
Building with Car-park	0.64 [0.48]	0.67 [0.47]	0.80 [0.40]	0.71 [0.45]	0.75 [0.43]	0.77 [0.42]	0.11 [0.31]	0.10 [0.29]	0.08 [0.27]

Note: All distances are measured in straight-line (linear) kilometer. Standard deviations are in brackets below the sample means.

Table 2: Hedonic and Matched Sample Regressions

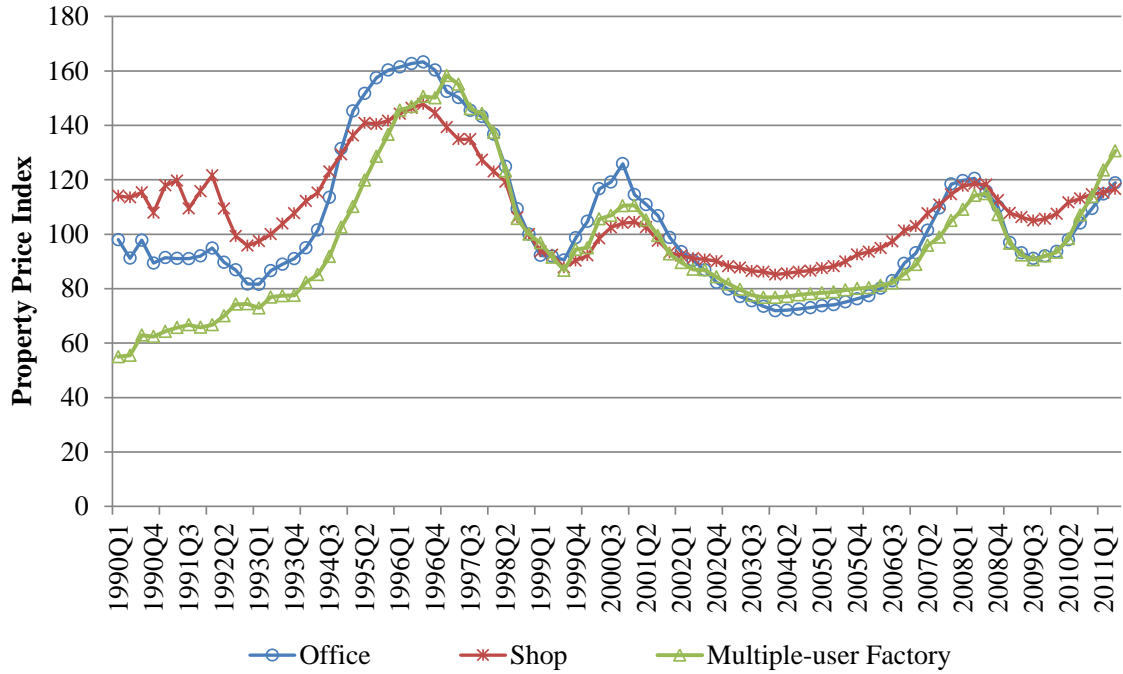
	Office				Shop				Multiple-User Factory			
	Full Sample		Matched Sample		Full Sample		Matched Sample		Full Sample		Matched Sample	
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
Log Area of Unit (square meters)	1.084*** -(0.006)	1.022*** -(0.006)	1.013*** -(0.012)	0.954*** -(0.011)	0.809*** -(0.008)	0.837*** -(0.007)	0.810*** -(0.011)	0.853*** -(0.010)	0.851*** -(0.004)	0.887*** -(0.005)	0.844*** -(0.006)	0.881*** -(0.006)
Log Distance to CBD	0.118*** -(0.012)	0.087*** -(0.015)	0.100*** -(0.035)	0.332*** -(0.041)	0.144*** -(0.023)	-0.084*** -(0.028)	0.194*** -(0.029)	-0.103*** -(0.038)	-0.237*** -(0.028)	-0.138*** -(0.033)	-0.361*** -(0.039)	-0.259*** -(0.045)
Log Distance to Airport	0.753*** -(0.061)	-1.189*** -(0.181)	0.582*** -(0.101)	-1.303*** -(0.358)	0.179*** -(0.034)	-0.508*** -(0.109)	0.075 -(0.047)	-0.738*** -(0.205)	0.120*** -(0.010)	0.153*** -(0.048)	0.227*** -(0.017)	0.276*** -(0.063)
Log Distance to Seaport	-0.230*** -(0.015)	-0.077*** -(0.021)	-0.137*** -(0.033)	-0.123*** -(0.043)	-0.192*** -(0.030)	0.007 -(0.037)	-0.248*** -(0.039)	-0.035 -(0.054)	-0.272*** -(0.027)	-0.686*** -(0.052)	-0.166*** -(0.037)	-0.632*** -(0.075)
Log Distance to MRT / LRT Station	-0.110*** -(0.007)	-0.154*** -(0.009)	-0.051*** -(0.016)	-0.157*** -(0.017)	-0.182*** -(0.009)	-0.232*** -(0.010)	-0.177*** -(0.011)	-0.256*** -(0.013)	0.046*** -(0.009)	0.100*** -(0.011)	0.014*** -(0.013)	0.024 -(0.015)
Log Distance to Expressway	0.120*** -(0.011)	0.027** -(0.011)	0.002 -(0.023)	0.017 -(0.021)	0.093*** -(0.011)	0.080*** -(0.012)	0.055*** -(0.016)	0.081*** -(0.017)	0.033*** -(0.005)	0.060*** -(0.005)	0.046*** -(0.007)	0.058*** -(0.008)
Inside the CBD	0.005*** -(0.023)	0.188*** -(0.023)	0.106*** -(0.037)	0.364*** -(0.036)	-0.129*** -(0.023)	0.086*** -(0.028)	-0.070** -(0.029)	0.194*** -(0.036)				
MRT/LRT Station within 0.3km	-0.591*** -(0.034)	-0.140*** -(0.119)	-0.508*** -(0.059)	0.915*** -(0.299)	-0.160*** -(0.029)	-0.229*** -(0.046)	-0.110*** -(0.037)	-0.323*** -(0.078)	0.042*** -(0.012)	0.036*** -(0.013)	0.035** -(0.016)	0.023 -(0.017)
Expressway with 0.3km	-0.034 -(0.022)	-0.155*** -(0.021)	-0.320*** -(0.038)	-0.157*** -(0.039)	-0.086*** -(0.020)	-0.085*** -(0.022)	-0.180*** -(0.027)	-0.099*** -(0.030)	0.082*** -(0.011)	0.103*** -(0.011)	0.012 -(0.016)	0.069*** -(0.016)
Building with car-park	-0.064*** -(0.015)	-0.030*** -(0.014)	0.116*** -(0.030)	0.101*** -(0.029)	0.147*** -(0.016)	0.138*** -(0.016)	0.173*** -(0.022)	0.141*** -(0.021)	0.019* -(0.011)	-0.086*** -(0.012)	-0.036** -(0.017)	-0.069*** -(0.017)
Shopping Center within 0.3km	0.957** -(0.395)	-0.550 -(0.357)			0.135** -(0.053)	-0.104 -(0.069)	0.227 -(0.141)	1.451*** -(0.329)	-0.007 -(0.008)	0.071*** -(0.009)	-0.055*** -(0.010)	0.005 -(0.012)
Public Housing within 0.3km	-0.274*** -(0.018)	-0.036** -(0.018)	-0.494*** -(0.031)	-0.204*** -(0.033)	-0.359*** -(0.020)	-0.129*** -(0.025)	-0.388*** -(0.025)	-0.179*** -(0.031)	0.042*** -(0.009)	-0.162*** -(0.011)	0.209*** -(0.015)	-0.060*** -(0.017)
Yr-Qtr Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
R-Square	0.897	0.927	0.876	0.912	0.660	0.716	0.625	0.742	0.816	0.847	0.792	0.823
Number of Obs.	4646		1622		6601		4103		11595		7298	

Note: The dependent variable is the natural log of the sale price. The explanatory variables include characteristics of the building and neighborhood listed in Table 1. '***' denotes significance at the 1% level; '**' denotes the 5% significance level; and '*' denotes the 10% significance level.

Table 3: Properties of the Matched Sample Mean and Regression-Based Indices and the URA Index

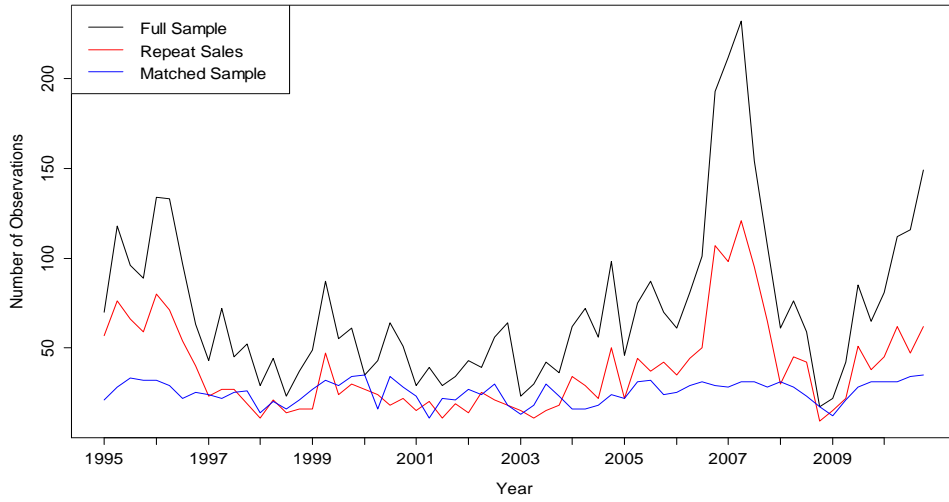
	Office			Shop			Multiple-User Factor		
	<i>matched sample means</i>	<i>matched sample regressions</i>	<i>unmatched sample URA Index</i>	<i>matched sample mean</i>	<i>matched sample regressions</i>	<i>unmatched sample URA Index</i>	<i>matched sample means</i>	<i>matched sample regressions</i>	<i>unmatched sample URA Index</i>
Mean	-0.166	-0.296	-0.223	0.135	0.039	-0.100	-0.235	-0.289	-0.408
Median	-0.165	-0.267	-0.186	0.123	-0.002	-0.115	-0.265	-0.295	-0.464
Standard Deviation	0.316	0.291	0.355	0.177	0.186	0.300	0.294	0.283	0.373
Sample Variance	0.100	0.085	0.126	0.031	0.034	0.090	0.086	0.080	0.139
Kurtosis	-0.319	-1.379	1.087	-0.577	-0.464	0.290	0.704	-0.652	2.155
Skewness	0.160	-0.074	0.417	-0.054	0.572	0.512	1.024	0.350	1.166
Range	1.445	0.995	2.069	0.751	0.789	1.454	1.254	1.098	2.055
Minimum	-0.862	-0.769	-1.043	-0.290	-0.259	-0.620	-0.674	-0.750	-1.136
Maximum	0.582	0.226	1.026	0.461	0.530	0.833	0.580	0.347	0.918
Count	63	63	63	63	63	63	63	63	63

Figure 1: Commercial Real Estate Price Indices

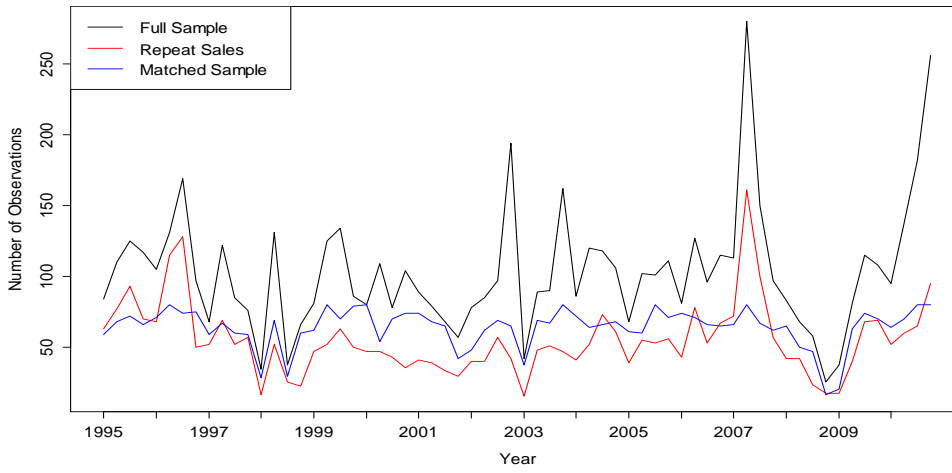


Source: URA

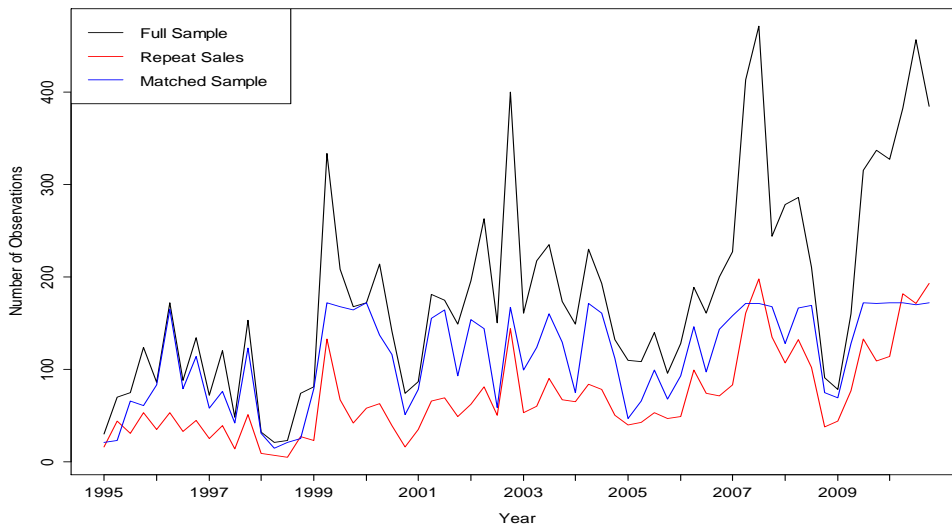
Figure 2: Number of Observations by Quarter



(a) Office

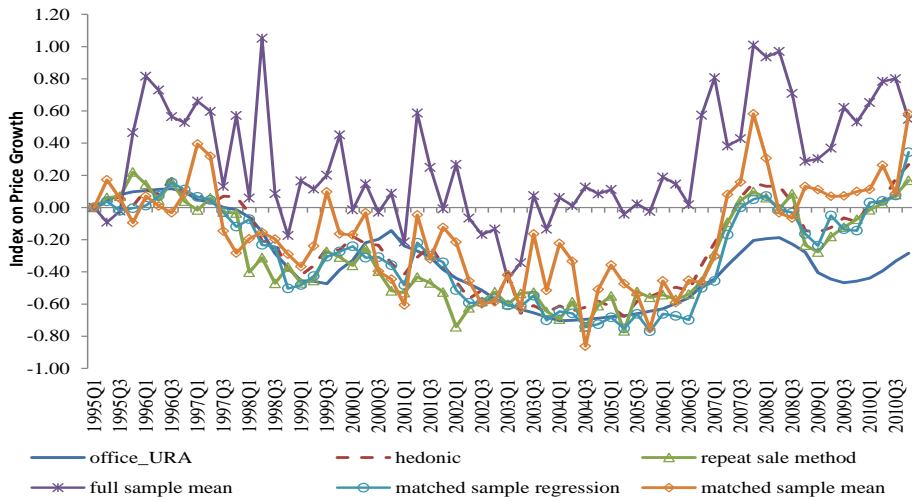


(b) Shop

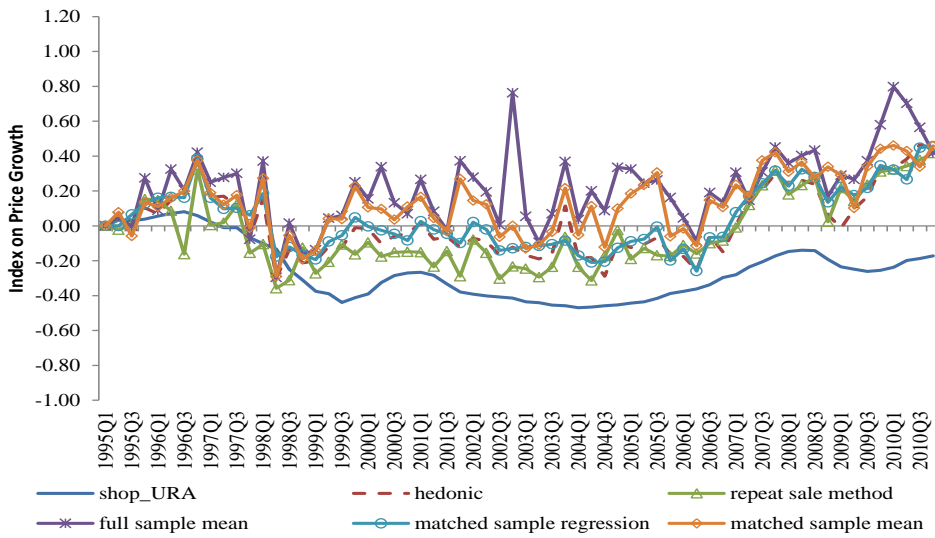


(c) Multiple-user Factory

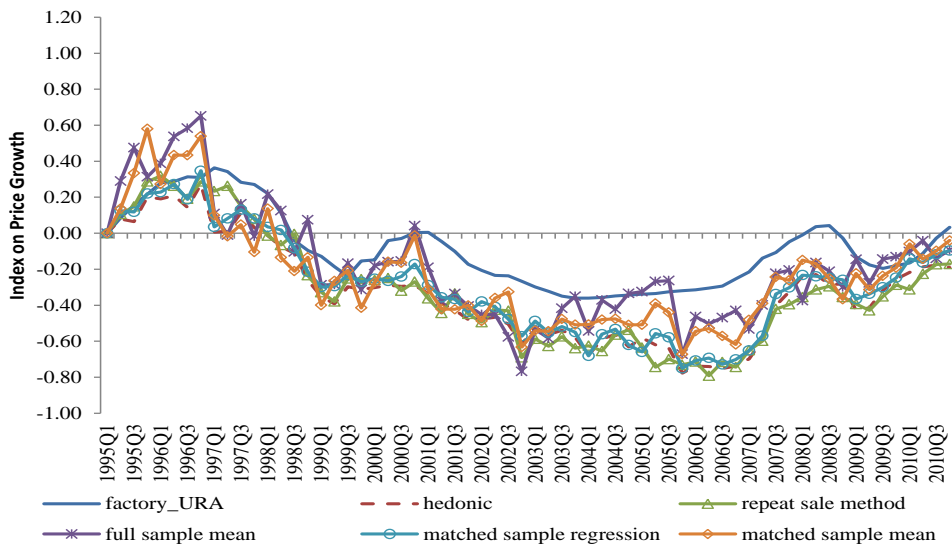
Figure 3: Standard Price Indices and Matched Sample Indices



(a) Office

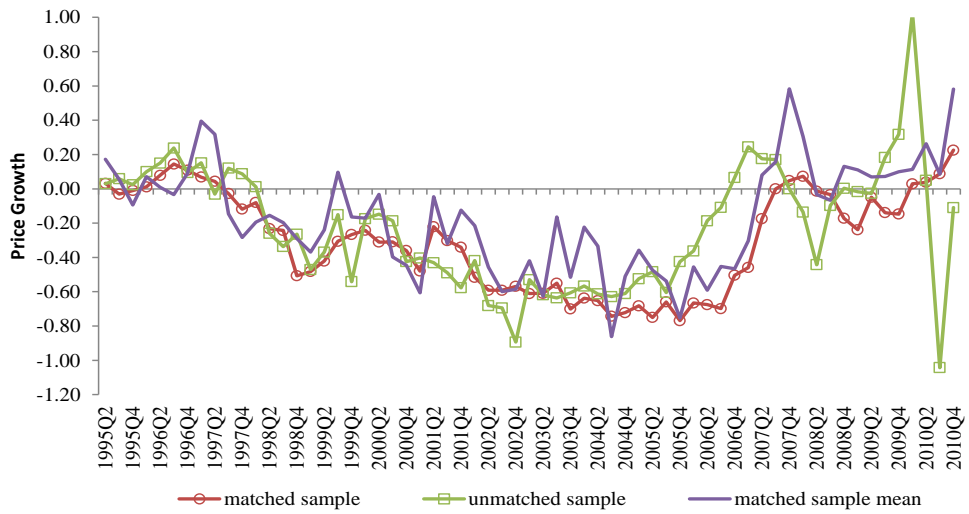


(b) Shop



(c) Multiple-User Factory

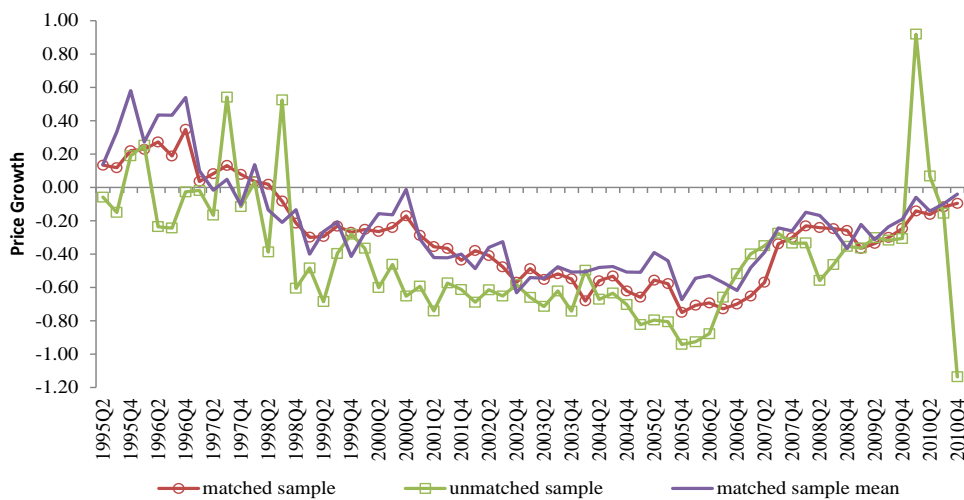
Figure 4: Indices based on matched and unmatched samples



(a) Office

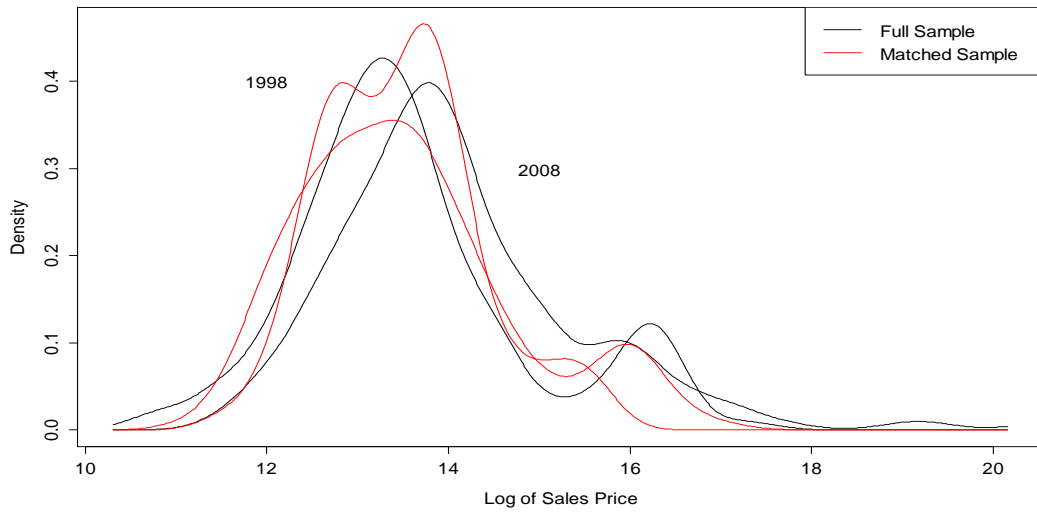


(b) Shop

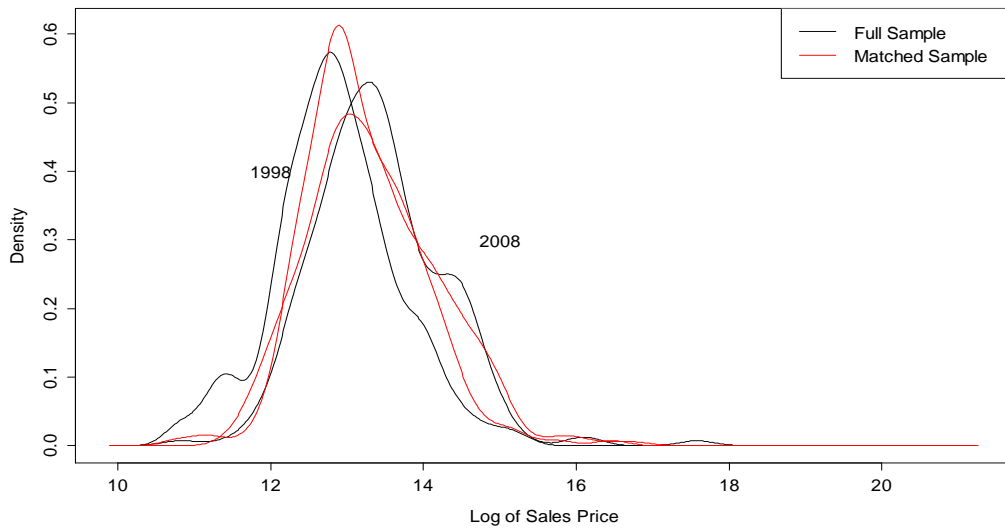


(c) Multiple-user factory

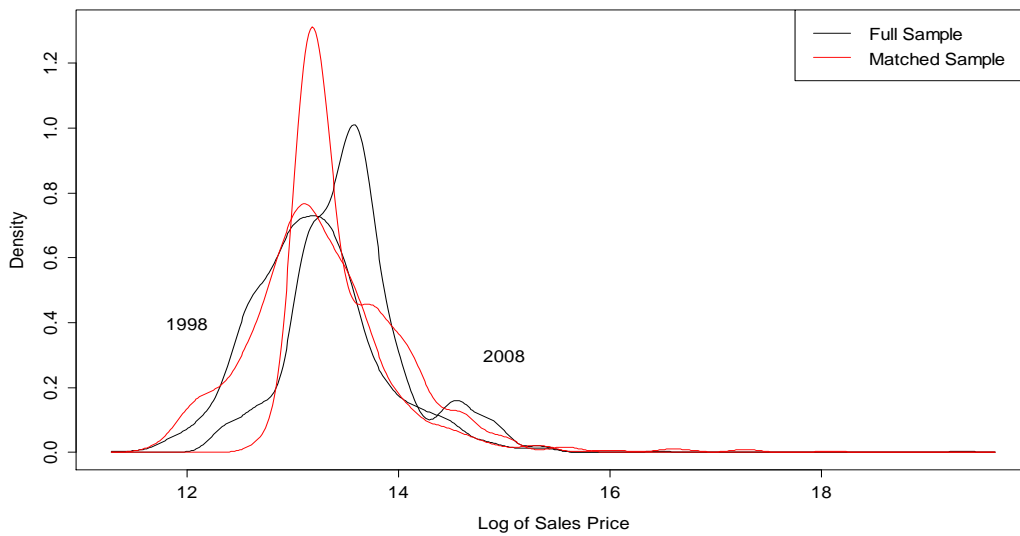
Figure 5: Sale Price Densities



(a) Office



(b) Shop



(c) Multiple-User Factory

Figure 6(a): Price Quantiles for the Full Sample and the Matched Sample - Office

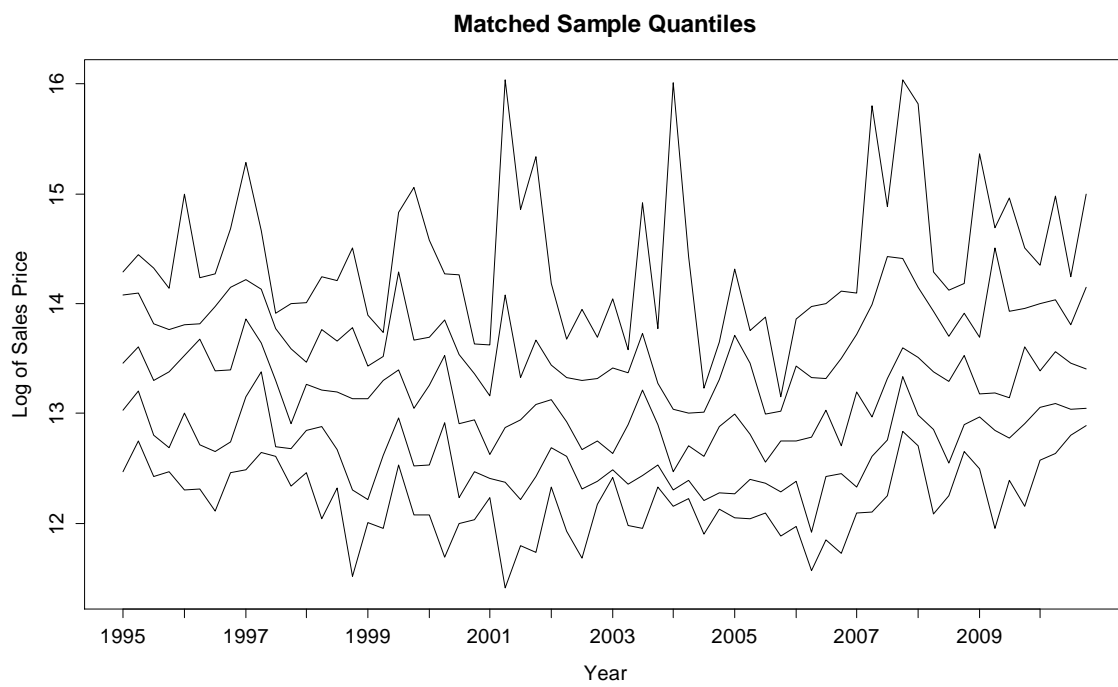
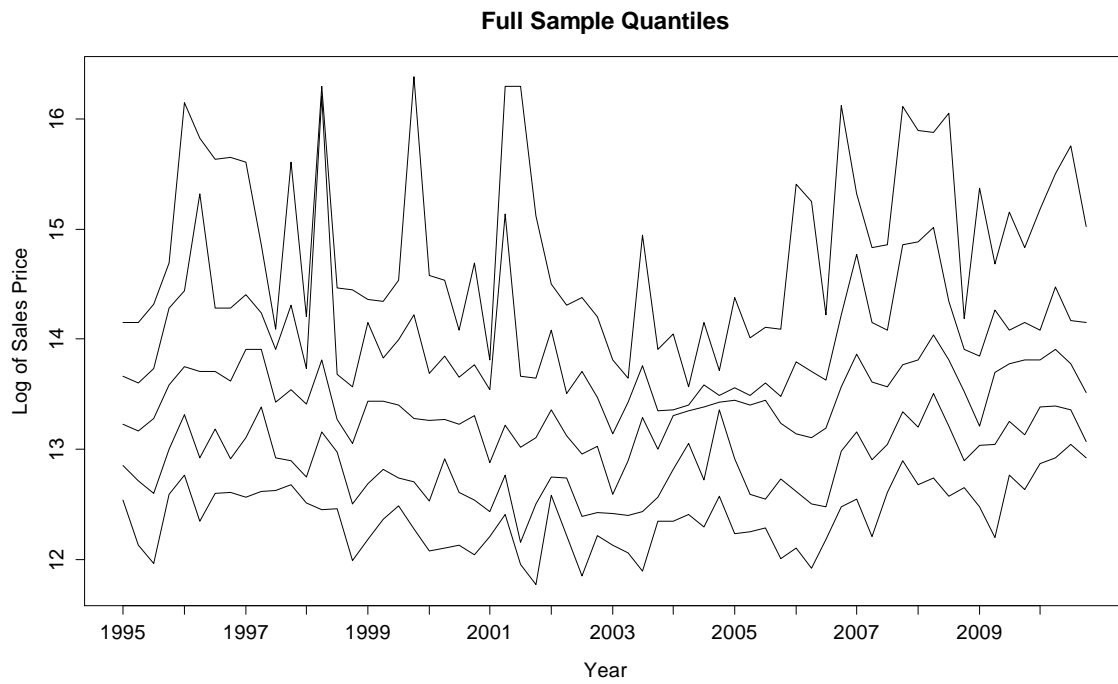


Figure 6(b): Price Quantiles for the Full Sample and the Matched Sample - Shop

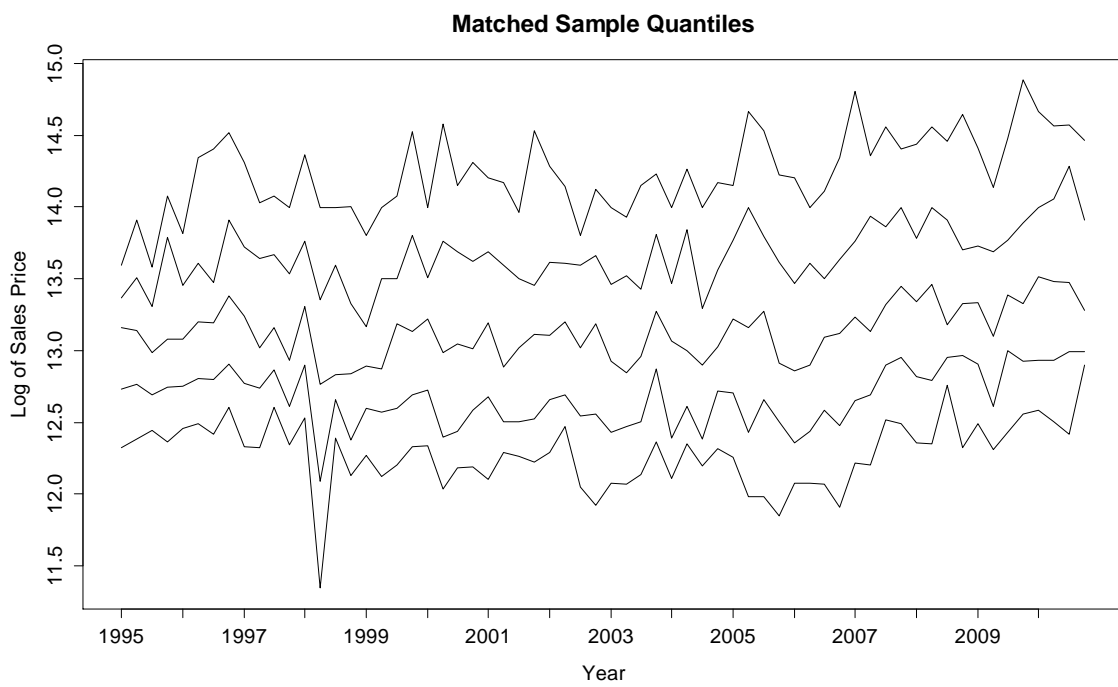
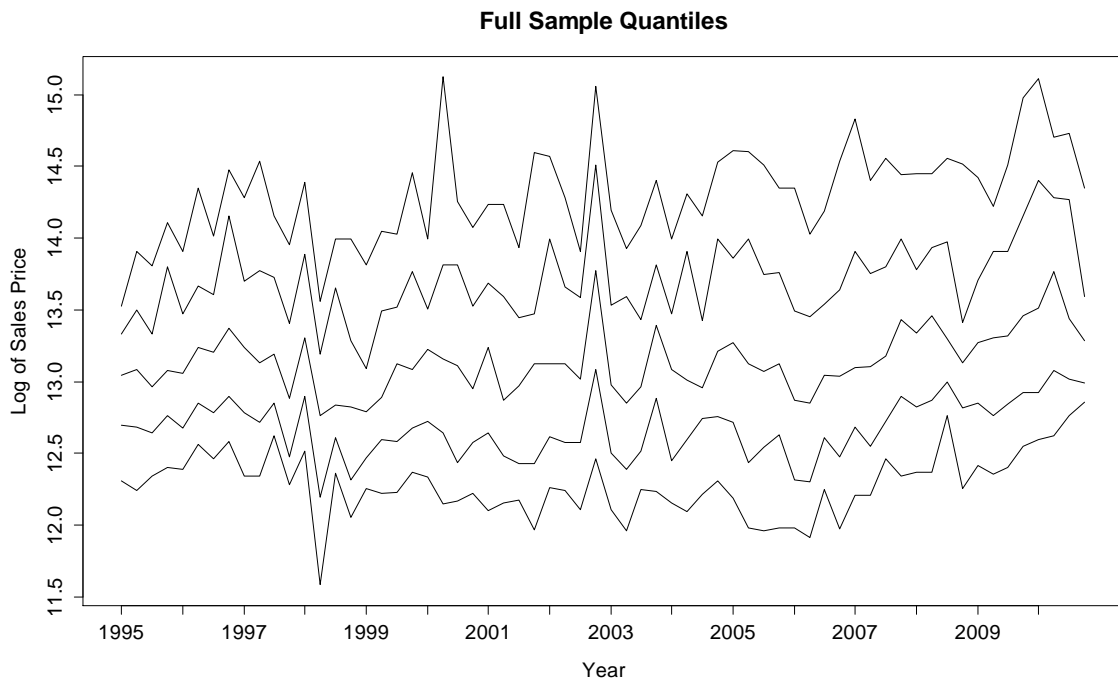


Figure 6(c): Price Quantiles for the Full Sample and the Matched Sample - Factory

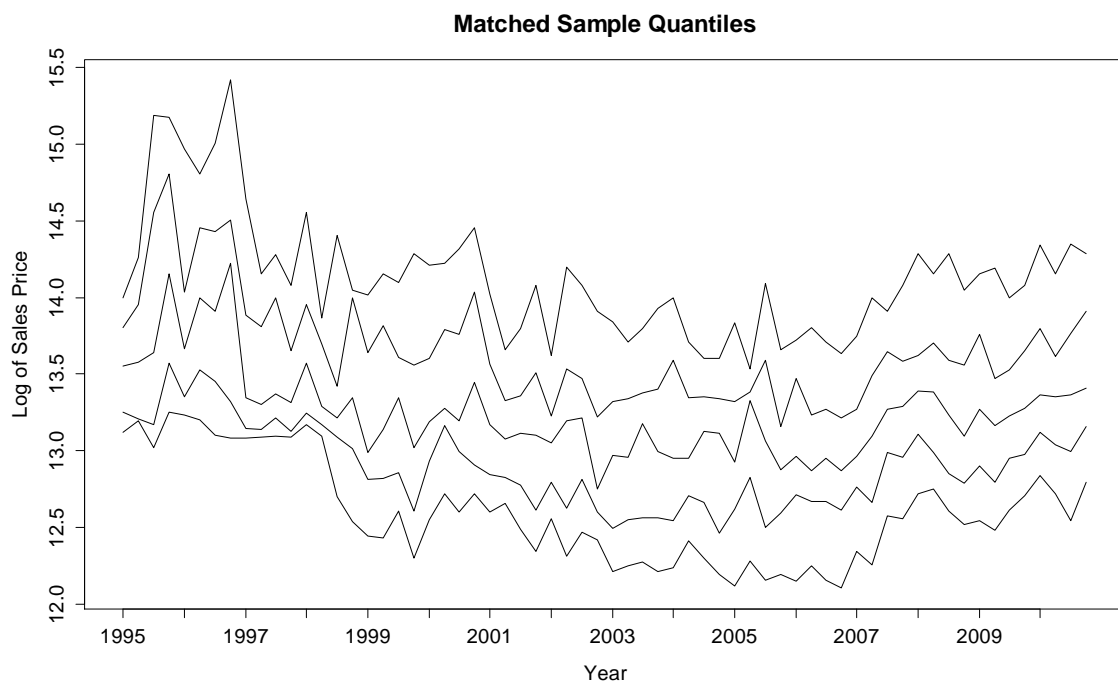
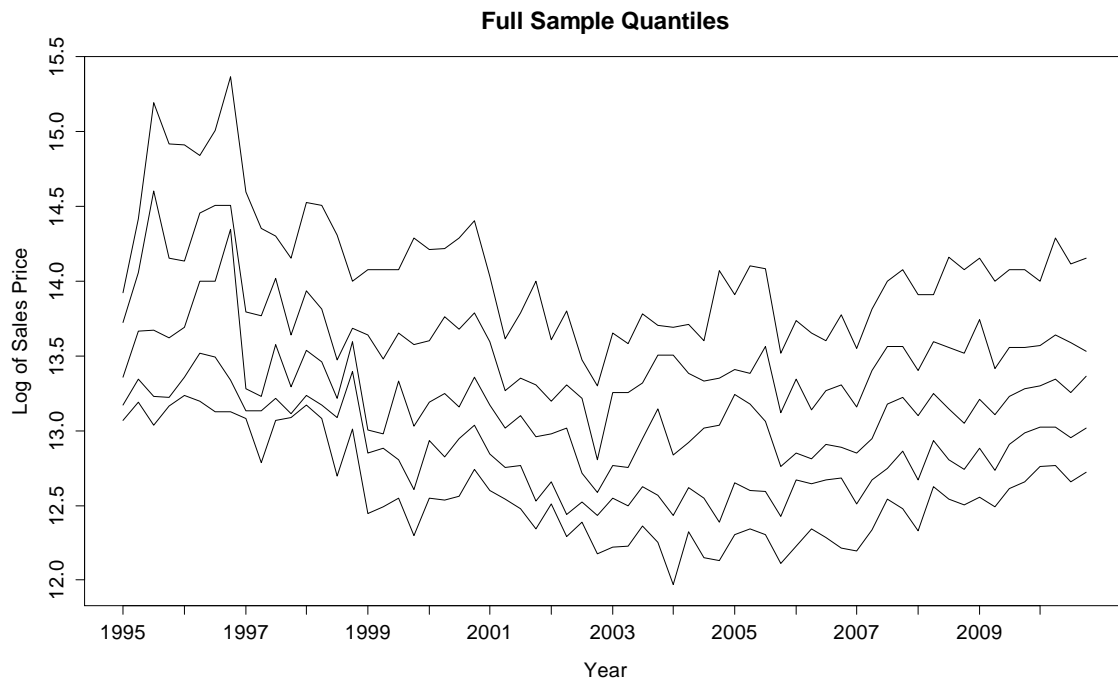
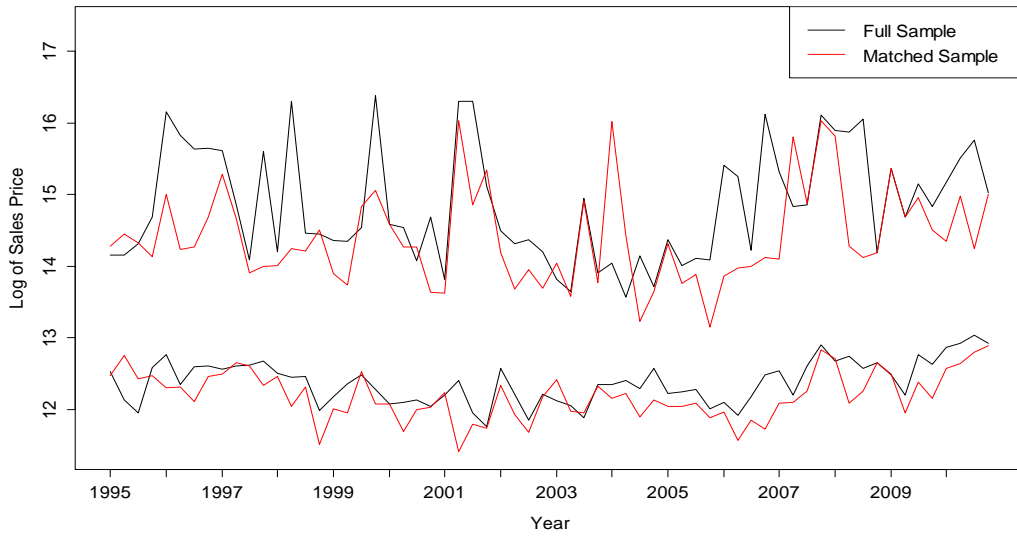
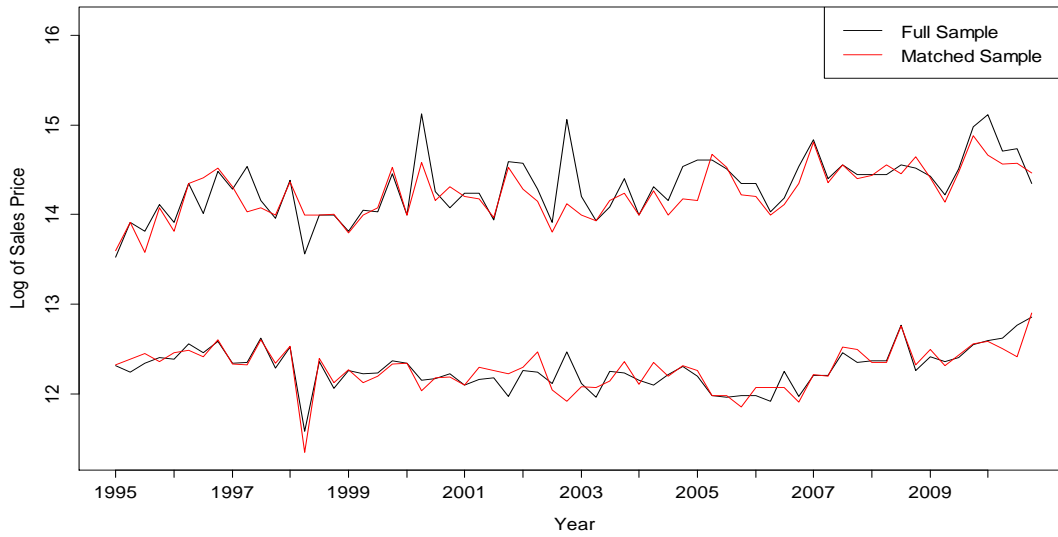


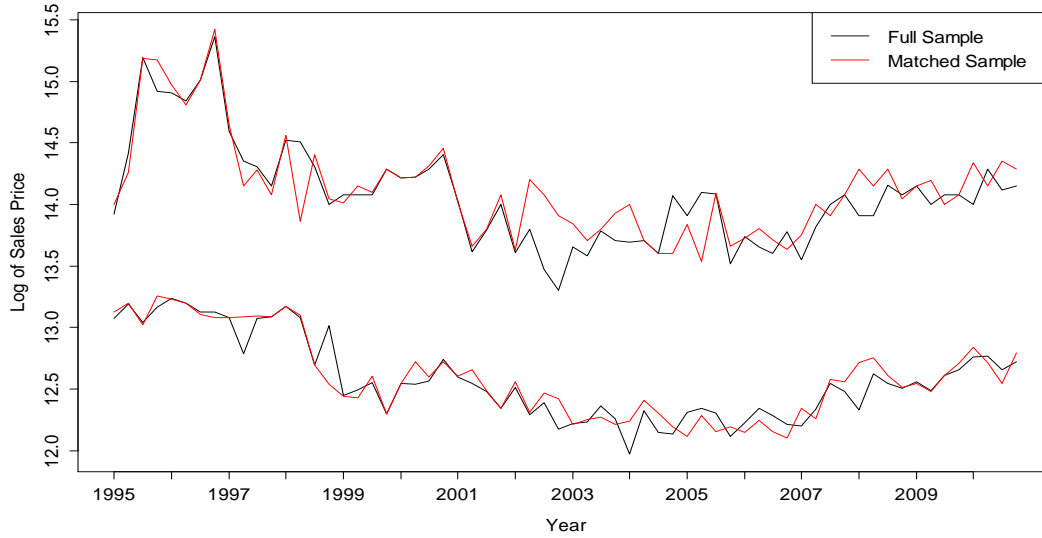
Figure 7: The 10th and 90th Percentiles of Log-Sales Prices



(a) Office



(b) Shop



(c) Multiple-User Factory