WP/14/72

IMF Working Paper

Commercial Property Price Indexes: Problems of Sparse Data, Spatial Spillovers, and Weighting

Mick Silver and Brian Graf

INTERNATIONAL MONETARY FUND

 WP/14/72

IMF Working Paper

Statistics Department

Commercial Property Price Indexes: Problems of Sparse Data, Spatial Spillovers, and Weighting

Prepared by Mick Silver and Brian Graf[1]

Authorized for distribution by Kimberly Zieschang

April 2014

Abstract

Transaction-price residential (house) and commercial property price indexes (RPPIs and CPPIs) have inherent problems of sparse data on heterogeneous properties, more so CPPIs. In an attempt to control for heterogeneity, (repeat-sales and hedonic) panel data regression frameworks are typically used for estimating overall price change. We address the problem of sparse data, demonstrate the need to include spatial price spillovers to remove bias, and propose an innovative approach to effectively weight regional CPPIs along with improvements to higher-level weighting systems. The study uses spatial panel regressions on granular CPPIs for the United States (US).

JEL Classification Numbers: C33, C43, E30, E31, R32.

Keywords: Commercial Property Price Indexes; House Price Indexes; Errors in Measurement; Spatial Econometrics; Spatial Panel Regressions; Index Number Weights

Author's E-Mail Address: msilver@imf.org and bgraf@imf.org

[1] Acknowledgements are due to Real Capital Analytics (RCA) for the supply of the data used in this study and subsequent technical advice on the features of the data. We note our due recognition of caveats as to the confidentiality of the data. We acknowledge helpful comments from Jeff Fisher (RCA), David Geltner (MIT), and Xingwei Hu (IMF).

Contents

I. INTRODUCTION

In principle, residential (house) and commercial property price indexes (RPPIs and CPPIs) should be at least quarterly measures of change in actual transaction prices of a constant-quality basket of representative properties. In practice, residential and commercial properties are heterogeneous and transactions irregular thus complicating comparisons of average transaction prices for a fixed-quality bundle of properties over time—the type of methodology that would be applied for a consumer price index. Hedonic regressions of price on the quality characteristics of the properties and use of only matched (repeat) transactions are the two primary methods used to control for changes in the quality-mix of properties transacted. Both can be estimated as panel regressions using the parameter estimates of dummy variables on time (see Eurostat *et al*., 2013) to derive the index.

Accurate measures of RPPIs and CPPIs are recognized as important for economic analysis, monetary policy, financial stability and prudential supervision. The International Monetary Fund (IMF)/Financial Stability Board (FSB) G-20 Data Gaps Initiative (DGI), endorsed by the G-20 Finance Ministers and Central Bank Governors, as well as the IMF's International Monetary and Financial Committee (IMFC), includes CPPIs in its 20 recommendations on data gaps (see Heath, 2013). Yet their measurement remains highly problematic.

We outline methodological improvements that apply to transaction-based RPPIs and CPPIs, though the focus hereafter is CPPIs. Regression formulations are commonly used in the aggregation of RPPIs and CPPIs in order to control for the afore-mentioned quality-mix changes of transactions data. The concern of this paper is not with these quality-mix methods *per se*, but the efficiency and bias of estimates of aggregate price movements arising from the panel regression frameworks often used. We base the empirical work on panel granular transaction-based CPPIs by metropolitan area and type of transaction for the United States (US) from 2000Q4–2012Q4.

Several methodological innovations are proposed: (i) the concern is with sparse data and, akin to normal statistical practice when faced with limited sample sizes, increasing the efficiency of the estimator. Data on "counts"—number of transactions—in each quarter for each area/type of property are used to improve the efficiency of the estimates of US commercial property price inflation; (ii) a spatial autoregressive (SAR) term is included in the regression thus removing potential bias by incorporating spillover effects—a SAR model; (iii) improvements are sought for parameter estimates by allowing the spatial weighting matrix and autoregressive parameter to vary over time; (iv) the $n \times n$ matrix (for n cross-sectional areas) that comprise each parameter estimate for each period is deconstructed and explicitly-weighted (direct, total, and indirect) averages of matrix elements provide estimates of commercial property price inflation, as opposed to the unweighted ones provided by standard software; (v) fixed and chained-weighted parameter estimates are provided for comparison; and (vi) the use of weights based on perpetual-inventory stock estimates are considered for higher level aggregation.

Section II briefly outlines select features of appraisal and transaction-based CPPIs with a focus on the United States (US). Section III discusses the methodological issues, with empirical results in section IV, and conclusions in section V.

II. Appraisal and Transaction-Based Indexes

A. Appraisal-based indexes

While the problem of sparse transactions on heterogeneous properties applies to both RPPIs and CPPIs, such shortcomings are particularly severe for CPPIs, especially when measurement really matters, as we go into and during recessions. As a result, valuation (appraisal-based) CPPIs are commonly used as a proxy to CPPIs since valuations of properties can be made irrespective of whether a transaction has taken place. An appraisal-based index is the ratio of a simple aggregation of appraised values in each period of a selected group of properties. The National Council of Real Estate Investment Fiduciaries (NCREIF) Property Index (NPI), for example, is a quarterly time series of the total rate of return of individual commercial real estate properties acquired in the private market for investment purposes.[2] The index covers office,

[2] Detailed information is given in the FAQ on NCREIF's website: http://www.ncreif.org/faqsproperty.aspx.

retail, industrial, apartment/hotel properties. The *total return* of a property includes both *net operating income* (NOI) and the c*apital return*, as if properties were purchased at the beginning of the quarter and sold at the end of the quarter with the investor receiving all net cash flows during the quarter. At the end of each quarter, (NCREIF data contributing-) members submit their appraisal of their properties' fair market values and NOI from which the aggregate NPI is compiled.

Such concepts and measures differ from a transaction price and suffer from limitations, as outlined in Diewert and Shimizu (2013), Geltner and Fisher (2007), Silver (2013), and Kanutin (2013). Appraisal-based indexes have an (albeit informed) subjective basis that does not accord with the (transaction-based) valuation principles of the *System of National Accounts* (see *SNA 2008*, chapter 2, C3). Further, valuations made by an appraisal firm are largely conducted irregularly, say annually, and quarterly data may in part be (stale) estimates by the manager/owner of the property largely based on the last formal appraisal. European experience with appraisal data is that appraisals generally take place annually and quarterly indexes compiled there from use interpolated values. European experience is that valuation methods differ across countries and, to a lesser extent, by different valuers within countries (Kanutin, 2013). There is evidence of a dampening or smoothing of market price volatility and a tendency of appraisal indexes to lag market price indexes, see Devaney and Diaz (2011), Fisher and Geltner (2007), and Geltner *et al*., (2003). Notably, users have an established preference for transaction-based indexes. The European Central Bank (ECB), as part of a stocktaking exercise on CPPIs, asked end-users their views as to their needs: the relatively uniform response was for commercial property price index based on transaction prices; valuation indexes were, as noted by Kanutin (2013, page 4), "..a second best option." The European Central Bank has a program of developing quarterly valuation-based CPPIs as a "second-best" solution given the sparse transaction data.

B. **Transaction-based indexes**

Real Capital Analytics (RCA CPPI) provided the transaction-based CPPI series used in this US study and we focus on its characteristics.[3] The data extend from 2000:Q4 to 2012:Q4 and comprise relatively high-value transactions; RCA data from 2000 covered transactions of over $5 million but in 2005 extended to transactions over $2.5 million (at constant dollars inflation-adjusted to December 2010). Applied filters exclude "flipped" properties (sold twice in 12-months or less), transactions not at arms-length, properties where size or use has changed, and properties with extreme price movements (more than 50 percent annual gain/loss).

RCA National CPPIs are aggregated up from more detailed CPPIs. CPPIs for "Major" and "Non-Major" metro markets for each of five property types (apartment, retail, industrial, office CBD, and office suburban), as illustrated in Annex 1, form the building blocks for the aggregation of the national indexes. The CPPIs for the building blocks and the individual granular CPPIs use repeat sales regression to compile the estimates. For each quarter since 1988, data are collected on sales; if a record of an earlier transaction for the property is identified, the two transactions are paired and treated as a repeat sale. A repeat sales price index is estimated by including data on the prices of transaction pairs in a regression on dummy variables on the time of each sale (Shiller, 1991). By limiting the sample to price comparisons of transactions-pairs of the same properties, some control is established over changes in the quality mix of transactions. The primary disadvantages of repeat sales indexes are (i) the quality of a repeat purchase may depreciate, with wear and tear, or appreciate, with renovations; (ii) increased sampling error due to relatively small sample sizes and potential sample selectivity bias—properties not sold or new properties sold once are excluded. The sample would comprise an unduly higher proportion of atypical properties sold more frequently (see Mason and Pryce, 2011); (iii) implications for the estimator due to an asymmetric and positive variance of the error term in the repeat sales regression; longer gaps between sales may require less weight; alternative assumptions on the heteroskedasticity of the error term as a result of this relationship can have a major impact on the index (Leventis, 2008); and (iv) as new transaction pairs

[3] For details on methodology see: presentation by David Geltner under "MIT/CRE Historical Development of CPPI" at http://web.mit.edu/cre/research/credl/ and Geltner and Pollakowski (2007). A subset of 20 national level indices are co-branded with Moody's Investors Service as the Moody's/RCA CPPI.

become available with the addition of new historical data, the index may be subject to a volatile revision history. Further details on methodology are given in Geltner and Pollakowski (2007).

RCA estimates repeat sales regression CPPIs for each property type according to whether the transaction is for properties in areas described as “Major” or “Non-Major” metro markets. These building blocks are weighted together to form higher-level aggregates as depicted by Figure 1 in Annex 1.

CPPIs that make use of transaction *and* appraisal data can be estimated. Notably, the ECB “transaction-linked” indexes, the result of regressing the log of actual transaction prices for traded properties, when available, on the appraised capital value for the preceding two quarters, and dummy variables on country and property-type. The estimated coefficients are used to increase the sample by including predicted transaction prices for *non-traded* properties, though the problem of underlying sparse transaction data remains (Kanutin, 2013 and Picchetti, 2013).

III. METHODOLOGY

There is an increasing availability of more granular property price indexes. For example, the Federal Housing Finance Agency (FHFA) US House Price Index in its 2013Q1 release provided data for 75 metropolitan areas—an increase from 25 areas in previous quarters. The granular RCA US repeat sales CPPIs for 34 metros/markets used in this study was released in December 2012 extending back to 2000Q4. We use a two-way fixed effects panel regression estimator to derive estimates of higher-level commercial property price inflation. This paper focuses on methodological improvements to transaction-based CPPIs—see also Bokhari and Geltner (2012).

A. **Sparse data, the use of counts information, and CPPIs**

The requirement is for an efficient estimator given the sample size of a period’s CPPI is limited to (i) properties transacted in that period; (ii) for the repeat-sales approach, to properties that have had more than one sale during the period of the index; and (iii) transactions not deleted by data quality filters, as outlined in Section IIIA. Geltner and Pollakowski (2007, page 18) note that the RCA National All-Property CPPI averaged 285 monthly repeat sales in 2006, but only 29 in 2001, at its inception.

Consider a two-way fixed effects panel model:

$$\mathbf{Y}_{n,t} = \mathbf{Z}_{\mu}\boldsymbol{\mu}_n + \mathbf{Z}_{\gamma}\boldsymbol{\gamma}_t + \mathbf{V}_{n,t} \quad \text{...(1)}$$

where $\mathbf{Y}_{n,t} = \left(y_{1,t,} y_{2,t,} ..., y_{n,t,}\right)$ is a $n \times 1$ vector of commercial property price inflation (log-change of the index) for each of the periods $t = 1,.....,T$; $\boldsymbol{\mu}_n$ is the $n \times 1$ parameter vector of spatial (area) fixed effects and $\mathbf{Z}_{\mu}$ the associated $n \times n$ dummy variable matrix; $\boldsymbol{\gamma}_t$ is the $t \times 1$ parameter vector of fixed time effects and $\mathbf{Z}_{\gamma}$ the associated $n \times t$ dummy variable matrix; and $\mathbf{V}_{n,t} = \left(\nu_{1,t,} \nu_{2,t,} ..., \nu_{n,t,}\right) \sim \text{IID}(0, \sigma_V^2)$ are stochastic disturbances. The fixed time effects parameters $\boldsymbol{\gamma}_t$ are estimated using the least squares dummy variable (LSDV) method as opposed to demeaning, given a specific interest in $\boldsymbol{\gamma}_t$ and inflation estimates derived therefrom; the restriction is imposed that $\boldsymbol{\gamma}_1 = 0$ and a constant is included in equation (1). In addition, "counts data" $\mathbf{C}_{n,t}$, are the number of observed price transactions for each area *n* in each period *t*. We use ordinary least squares (OLS) and weighted least squares (WLS) estimators, the latter with $\sqrt{\mathbf{C}_{n,t}} = \left(\sqrt{c_{1,t}}, \sqrt{c_{2,t}} ..., \sqrt{c_{n,t}}\right)$ as explicit weights.

The assumption is that $V\left(\mathbf{V}_{n,t}\right) = \sigma^2 / \mathbf{C}_{n,t}$; as counts increase, the variance decreases. The $\hat{\boldsymbol{\gamma}}_t$ form the basis of estimates of property price inflation as outlined in Section III. Taken as a whole series, the more efficient WLS estimates can be argued to better estimate changes in commercial property price inflation. OLS gives less precisely measured observations more influence than they should have and more precisely measured ones too little influence. WLS using counts data assigns a weight to each observation that reflects the uncertainty of the measurement and thus improves the efficiency of the parameter estimates.

This focus on the efficiency of the estimator is in line with the literature on "errors in measurement" in the dependant variable—Hausman (2001). Such measurement errors result in

OLS parameter estimates that are unbiased, but inefficient, with reduced precision and associated lower *t*-statistics and $\bar{R}^2$. (This differs from the literature on measurement errors in the explanatory variable for which OLS parameter estimates are biased.) The measured value of $\mathbf{Y}_{n,t}$ is the sum of the true measure $\mathbf{Y}^*_{n,t}$ plus a measurement error $\mathbf{E}_Y$:

$\mathbf{Y}_{n,t} = \mathbf{Y}^*_{n,t} + \mathbf{E}_Y$ and the true measure is $\mathbf{Y}^*_{n,t} = \mathbf{Y}_{n,t} - \mathbf{E}_Y$.

Instead of estimating: $\mathbf{Y}^*_{n,t} = \mathbf{Z}_\mu \boldsymbol{\mu}_\mathbf{n} + \mathbf{Z}_\gamma \boldsymbol{\gamma}_\mathbf{t} + \mathbf{V}_{n,t}$, we estimate:

$$\mathbf{Y}^*_{n,t} = \mathbf{Z}_\mu \boldsymbol{\mu}_n + \mathbf{Z}_\gamma \boldsymbol{\gamma}_t + \mathbf{E}_Y + \mathbf{V}_{n,t} \quad \text{(2)}$$

Measurement error thus increases the variance of the error term from $\mathrm{var}(\mathbf{V}_{n,t})$ to $\mathrm{var}(\mathbf{V}_{n,t}) + \mathrm{var}(\mathbf{E}_Y)$ and the variance (standard error) of $\boldsymbol{\gamma}_t$ accordingly increases. We directly target the $\mathrm{var}(\mathbf{E}_Y)$ component with explicit WLS counts weights $\mathbf{C}_{n,t}$.

B. Modeling spatial dependency

Unlike the measurement literature, modeling spatial dependency is not uncommon in the context of hedonic house price models, for example, Anselin (2008). We include a (first order) spatial autoregressive term in equation (1), a SAR model:[4]

$$\mathbf{Y}_{n,t} = \rho \mathbf{W}_n \mathbf{Y}_{n,t} + \mathbf{Z}_\mu \boldsymbol{\mu}_n + \mathbf{Z}_\gamma \boldsymbol{\gamma}_t + \mathbf{V}_{n,t} \boldsymbol{\mu}_n \quad \text{(3)}$$

where $\mathbf{W}_n$ is a $n \times n$ row-standardized spatial physical proximity weight matrix, outlined later, for the spatial autoregressive term and ρ the spatial autoregressive parameter to be estimated. Equation (3) expressed in its reduced form is given by:

$$\mathbf{Y}_{n,t} = \left(\mathbf{I}_\mathbf{n} - \rho \mathbf{W}_n\right)^{-1} \mathbf{Z}_\mu \boldsymbol{\mu}_n + \left(\mathbf{I}_\mathbf{n} - \rho \mathbf{W}_n\right)^{-1} \mathbf{Z}_\gamma \boldsymbol{\gamma}_t + \left(\mathbf{I}_\mathbf{n} - \rho \mathbf{W}_n\right)^{-1} \mathbf{V}_{n,t} \quad \text{(4)}$$

[4] We note from Manski (1993) that when a spatially lagged dependent variable, spatially lagged regressors, and a spatially autocorrelated error term are included simultaneously the parameters of the model are not indentified unless at least one of these interactions are excluded. We found no firm evidence for the spatial autocorrelated error (SEM) model and our explanatory variables of interest, the time dummies, *a priori*, have no spillover effect. In any event, we follow the more general advice by LeSage and Pace (2009, pp. 155–58), LeSage (2012), and Elhorst (2010) to adopt the SAR model and exclude the spatially autocorrelated error term to favor inclusion of the spatially autoregressive one.

The matrix of partial derivatives of $\mathbf{Y}_{n,t}$ with respect to a change in a dummy time variable, is given by Elhorst (2010) and Debarsy and Ertuur (2010) as:

$$\frac{\partial \mathbf{Y}_{n,t}}{\partial \mathbf{Z}_{\gamma}} = \boldsymbol{\gamma}_t \left(\mathbf{I_n} - \rho \mathbf{W}_n \right)^{-1} = \mathbf{B}_t \quad \text{……………………………………………………..…..(5)}$$

There is a resulting $n \times n$ matrix $\mathbf{B}_t$ for the marginal effect of each estimated parameter on a time dummy variable. It is apparent from equation (5) that in a fixed effect panel OLS model where $\rho = 0$, the diagonal elements would be $\boldsymbol{\gamma}_t$ and the non-diagonal effects zero, resulting in a single parameter estimate $\boldsymbol{\gamma}_t$ for each period *t*. In this SAR model, the spatial direct effects are not $\boldsymbol{\gamma}_t$, but are given for each area *n* by the diagonal elements of $\mathbf{B}_t$. The top-left element would be the effect on property price inflation of moving from one quarter to the next for say, Boston, but differs from the OLS estimate in that it includes the resulting feedback effects from proximate spatially dependent areas, arising from Boston's property price inflation. Direct effects can be seen from equation (5) to depend on (i) their proximity to other areas, as dictated by $\mathbf{W}_n$; (ii) the strength of spatial dependence, ρ; and (iii) the parameter $\boldsymbol{\gamma}_t$. The diagonal elements are estimates of the direct effect for each area *n*.

The indirect effects for each area *n* are given by the off-diagonal column elements of $\mathbf{B}_t$ and the total effect for area *n* is the column sum of area *n* and includes its direct and indirect effect. We follow LeSage and Pace (2009), and the output of standard software in this area, in reporting, for the marginal effect of each time dummy parameter estimate, one direct effect as the average of the diagonal of the elements, $tr\left(\mathbf{B}_t\right)/n$ and one total effect measured as the average of the column sums; the indirect effect is deduced as the difference between the two. The average total effect answers the question: what will be the average total impact on property price inflation of the typical area? (LeSage and Pace, 2009).

The spatial approach ameliorates omitted variable bias by its inclusion of $\rho \mathbf{W}_n \mathbf{Y}_{n,t}$ in equation (5). Debarsy and Ertur (2010), in a study of spillovers in a panel regression for 24 OECD countries of domestic savings on investment found, for 1971–1985, the coefficient from a conventional fixed effect panel estimator to be reduced from 0.609 to 0.452 when using a SAR model, a more reasonable estimate *a priori* in the context of the Feldstein-Horioka (1980) paradox.

C. Relaxing the restrictions of constant ρ and $\mathbf{W}_n$

In empirical work both ρ and $\mathbf{W}_n$ are generally held constant over time. Indeed the software noted and used in Section III for spatial econometric estimation does not allow any such variation within a panel. This restriction is relaxed by annually estimating equation (5) for 4 quarters in 2000, again for 2001, 2002,…., 2012, thus allowing ρ^{τ} and $\mathbf{W}_n^{\tau}$, $\tau = 1(2001),........,12(2012)$ to be fixed for quarters within each year, but vary between years. The spatial weight matrix is a proximity matrix generated from the longitudes and latitudes of the areas. As outlined in the empirical section below, the centers of these areas are in some cases better represented by a (relative employment) weighted average of the longitudes and latitudes of major conurbations within each area. There is a sense in which the center of gravity of an area with regard to commercial property can change over time and we take some account of this by estimating ρ^{τ} in annual segments and using $\mathbf{W}_n^{\tau}$. We compare the resulting measures.

D. Weighted averages of commercial property price inflation

The derivation of estimates of the direct effects as an average of the diagonals and total effects as an average of the column sums of equation (5) has an interesting useful index number application. The averaging applied by the software is unweighted over the *n* areas. We can deconstruct equation (5) into the total and direct effect for their *n* components and apply weights to the estimated area (direct and total) price changes; the weights may be relative values of the stock of, or transactions in, commercial property. Diewart (2005) previously proposed weighting systems within a regression framework via analytic weights using a WLS estimator. If using a SAR model, the framework advocated here provides an alternative explicit weighting mechanism that can be applied to each of the direct, total, and indirect spatial effects; thus, allowing WLS weights to be used for other purposes, say in relation to heteroscedasticity. For example, while the average unweighted direct effect for $\boldsymbol{\gamma}_1$ in Section IIB above is $tr(\mathbf{B}_1)/n$, the weighted average of direct effects for $\boldsymbol{\gamma}_1$ is given by:

$$tr(\mathbf{SB}_1) = \sum_n (\mathbf{SB}_1)_{nn} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(6)}$$

where $\mathbf{S}$ is a $n \times n$ matrix whose diagonal is the area *n* relative shares (weights) in stocks or transactions of commercial property.

E. **Relaxing the restriction of fixed weights**

In the previous section, the equation (6) weights, $\mathbf{S}$, were held constant for the estimation of each $\boldsymbol{\gamma}_t$. The assumption can be relaxed by estimating the *t*=48 commercial property price inflation rates in 12 independent windows of 4 quarters for successive years, thus allowing the weights to vary annually, that is using $\mathbf{S}^{\tau}$. We compare the resulting annually chain-weighted estimates with those resulting using fixed-base weights.

IV. Empirical Results

A. **Data**

The transaction-based CPPIs used in this US study were provided by Real Capital Analytics (RCA CPPI).[5] The coverage includes relatively high-value transactions; from 2000 commercial repeat-sale property transactions of over $5 million but extended in 2005 to transactions over $2.5 million (at constant dollars inflation-adjusted to December 2010). Applied filters exclude “flipped” properties (sold twice in 12-months or less), transactions not at arm’s-length, properties where size or use has changed, and properties with extreme price movements (more than 50 percent annual gain/loss). The empirical work uses two panel data sets: RCA CPPIs from 2000:Q1 to 2012:Q4 for “apartments” broken down by 34 metros/markets areas, and similarly for “other properties” that include industrial, office, and retail—hereafter “core

---

[5] Information on the RCA CPPI is at: https://www.rcanalytics.com/Public/rca_cppi.aspx; see Geltner and Pollakowski (2007) for methodological details.

commercial"—properties. RCA estimates each of the 34 granular series using repeat-sales regressions. In each section below, results will be presented for both apartments and core commercial properties. This paper does not replicate the RCA/Moody's aggregation process. We use all 34 area inflation rates to constitute $\mathbf{Y}_{n,t}$ on the left-hand-side of equation (1) above.[6]

B. **Sparse data**

Importantly, to gain insights into the problem of sparse data, RCA provided us with confidential data on the "counts"—number of transactions—underlying each of their granular indexes in each period. Our interest here focuses on whether by incorporating these counts data as weights in a WLS estimator, the measurement problems associated with sparse data can be ameliorated.

A natural first step is to identify whether there is any association between the number of observed price quotes in any period constituting the index[7] and commercial property price inflation. We would expect a positive relationship; the collapse during this current recession in commercial property price inflation to be associated with a fall in transaction counts and, similarly, movement out of the recession with increasing inflation and increased counts. Table 1 uses simple bivariate correlation coefficients to show that counts matter when explaining variation in property price inflation.[8]

---

[6] We use inflation rates as opposed to index levels in the regressions. A panel unit root test across 34 metropolitan areas, using individual fixed effects as regressors, found the null hypothesis of individual unit roots for CPPIs to be rejected for core commercial property inflation: the Im, Pesaran and Shin *W*-statistic of -28.62 (*p*-value=0.0000), the ADF Fisher Chi-squared statistic of 773.7, (*p*-value=0.0000), and the Phillips and Perron Fisher Chi-squared statistic of 832.3 (*p*-value=0.0000). For apartments rejection was at the 10 percent level for IPS -1.33 (0.092), ADF 5.02 (0.081), but not rejected by the PP 2.92 (0.232) —Levin, Lin, and Chu (2002).

[7] For the RCA CPPI data the count in a given quarter is the number of transactions in that period for which there are earlier transaction-pairs on the same property. It may be argued that if there are more than two such transactions, the earlier data contribute to the effective sample, albeit less so given a heteroskedasticity correction designed to give less importance to earlier, noisier data.

[8] We note that assigning causality to commercial property inflation and counts is *a priori* problematic. We nonetheless report that estimated pooled regressions of the former on the latter produce for each property type (apartment, industrial, office, and retail) in each sub-period (pre-recession and during recession) significant positive associations on the coefficients for counts (with the sole exception of retail during the recession), along with significant differences (Chow F-statistic=113.1, *p*-value 0.0000) between the period before the recession and during it. There is no evidence of Granger-causality between inflation and counts during the pre-recession period and limited contradictory results during the recession: for office property we reject the null that counts does not cause inflation and for apartments we reject the null that inflation Granger-causes counts, with the remaining six

(continued…)

A consistent positive association exists between inflation and counts for each of office, retail, industrial and apartment properties and their respective counts for the period 2001:Q1 to 2012:Q4 and within pre-recession and recession sub-periods; the association is statistically significant at a 5 percent level for each type of property in the period of the recession and for the aggregate of all properties in each period considered.

Table 1, Correlation coefficients: count and commercial property price inflation:

	All	Office	Retail	Industrial	Apartment
Whole period: 2001Q1: 2012Q4	0.298**	0.215	0.286*	0.312**	0.239
Pre-recession: 2000Q4: 2007Q4	0.538***	0.702***	0.007	0.721***	0.168
Recession: 2008Q1: 2012Q4	0.612***	0.487**	0.601***	0.578***	0.622***

***,**,* denotes statistically signifcicant for two-tailed test at 1, 5, and 10 percent levels respectively.

We provide estimates of commercial property price inflation for both apartments and core commercial properties using OLS and counts-based WLS panel regressions, as outlined in Section IIA. Estimates of commercial property price inflation are based on $\hat{\gamma}_t$ in equation (1).[9] Figure 1 shows the results[10] along with the Moody's/RCA CPPI for comparison.

test results for these four property types not being significant at a 5 percent level. Results of the regression and Granger-causality tests are available from the authors on request.

[9] We follow Kennedy (1981) and use as the estimate of the proportionate impact of the period *t* time dummy, the consistent (and almost unbiased) approximation: $\left[\exp\left(\hat{\gamma}_t\right)/\exp(V(\hat{\gamma}_t/2))\right]-1$ where $\hat{\gamma}_t$ *i*s the OLS estimator of γ_t in equation (1) above and $V(\hat{\gamma}_t)$ its estimated variance. The approximation is shown by Giles (2011) to be extremely accurate, even for quite small samples.

[10] Sample sizes for the pooled regressions, of 48 monthly observations on 34 areas, are 1,632. $\bar{R}^2$ for OLS estimators of apartments and core commercial properties are 0.47 and 0.48 respectively while for core commercial properties the corresponding figures are 0.43 and 0.44. Individual results on parameter estimates are available from the authors.

There is some convergence between the WLS estimated inflation and OLS/RCA inflation for US apartments going into the recession, coming out, and thereafter (Figure 1A). However, the amplitude of the trough around 2009:Q1 as estimated by WLS was markedly smaller, by over 1.5 percentage points. Less precisely measured values due to smaller sample sizes, are more likely to have extreme values with relatively high leverage; WLS accordingly gave the extreme values in the downturn less weight.

The analysis is repeated here for "core-commercial properties" for which sparse data is less problematic—the CPPIs and counts are aggregated over industrial, retail, and office properties, due to the relatively small numbers of transactions for each. However, for the aggregate of these properties, core commercial properties, the counts are larger than for apartments. Figure 1B shows that for estimating core commercial property price inflation where data are not so sparse, the use of counts hardly matters. As outlined in Annex 1, this aggregation differs from that used by RCA for higher-level CPPIs, co-branded as Moody's/RCA CPPIs. They employ repeat-sales regressions over two sub-aggregates of the 34 areas— for major and non-major metro markets. Some estimates for these aggregates are presented in Annex 1.

C. SAR fixed effects panel models with fixed and varying $\rho = \rho^{\tau}$ and $\mathbf{W}_n = \mathbf{W}_n^{\tau}$

In panel data regressions individual fixed effects, in this study for the *n* metro areas, were included to account for the *individual heterogeneity* of each panel member due to unobserved specific spatial factors. But heterogeneity also arises from an area's interaction with other areas, its *interactive heterogeneity* resulting from spatial autocorrelation. Ignoring spatial dependence in the dependent variable has been shown to lead to biased and inefficient estimates of coefficients in the explanatory variables, that would include our estimates of $\hat{\boldsymbol{\gamma}}_t$ in equation (1)—LeSage and Pace (2009) and Elhorst (2010).

The panel regressions are for each of apartments and core commercial inflation on 34 areas by 48 quarters. An OLS model was estimated for the standard two-way fixed effects model, equation (1), and this is used as the benchmark to which improvements are sought. SAR models were estimated using **xsmle** and related commands in Stata (Belotti *et al*., 2013). All SAR models use maximum likelihood (ML) estimators with Driscoll- Kraay standard errors that are heteroskedasticity- and autocorrelation-consistent to remaining general forms of spatial and temporal dependence, though we found differences to be minimal (Driscoll and Kraay, 1998).

Figure 1, OLS and WLS panel regression estimates of commercial property price inflation: quarter-on-quarter inflation, percent

Figure 1A, apartment property inflation

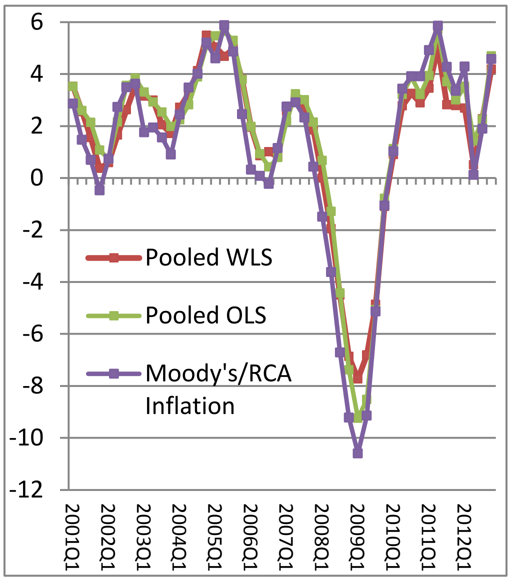
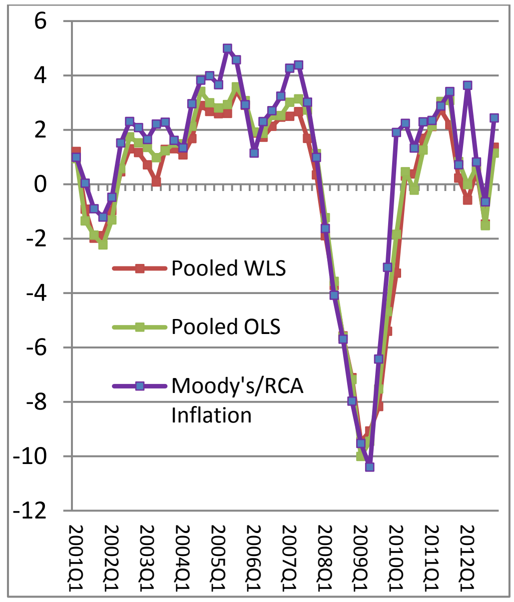


Figure 1B, core commercial property inflation



The data were transformed according to the Lee and Yu (2010) using a sub-option in **xsmle**, but this had little substantive effect.[11]

SAR panel (two-way) fixed effects models, equation (4), were estimated first, for the 34 areas by 48 quarters for each of apartments and core commercial properties, and then for twelve annual panel (non-overlapping) windows of 34 areas by 4 quarters for each year. The ρ^{τ} are

[11] Lee and Yu (2010a) consider the estimation of SAR panel data models with fixed time and individual effects and SAR disturbances. Conditions under which some parameter estimates may be inconsistent and distributions not properly centered are outlined. For the transformed (de-meaning) approach used for individual effects for this study—fixed time effects are modeled through time dummies, are treated as common parameters (Elhorst, 2003, Lee and Yu, 2010a and 2010b)—along with the use of row-standardized weights, the sample sizes in this study are sufficient, as judged by the Monte Carlo results in Lee and Yu (2010a), to consider common parameter estimates to be consistent and the asymptotic distributions properly centered.

fixed for the quarters within each year $\tau = 1(2001),........,12(2012)$, but vary between years. The longitude and latitude of each of the metro areas, and their proximity to other areas, was obtained from geospatial mapping data and $\mathbf{W}_n$ was derived as a row-standardized inverse distance matrix. $\mathbf{W}_n$ is usually held constant over time in spatial econometric modeling, indeed **xsmle** does not (at least currently) allow it to vary. However, for nine less-concentrated metro areas, within-area employment-weighted averages of the longitude and latitudes were used; for example, for "Mid-West excluding Chicago" the average longitude and latitudes were calculated as weighted averages of longitudes and latitudes of 13 major conurbations in the area.[12] $\mathbf{W}_n^\tau$ varied annually in these areas to allow some representation of changes in the center of gravity, in terms of employment, of the areas.

Summary regression diagnostics are given in Table 2, both for panel regression over 2001:Q1 to 2012:Q4 and the quarterly data of individual years. The null hypothesis of $\rho = 0$ is rejected for the whole period for both apartments and core commercial properties. Spatial autocorrelation is more prevalent for apartments than core commercial properties; $\hat{\rho}^\tau$ is statistically significant in just over half of the 12 individual years for apartments, and in only 3 years for core commercial properties, and 12 of these at a 10 percent level. With a single exception, it is positive when significant.

Figures 2A and 2B show the SAR total (fixed ρ and $\mathbf{W}_n$) effect is primarily constituted by the direct effect, with little indirect difference, except for the trough in 2009. By way of example, for apartments in 2009:Q1, 8.6 of the 12.0 percent fall in prices was due to direct effects and the remaining 3.4 percent to indirect effects. The OLS estimates are biased upwards against the

[12] Longitude and latitude data were from the US Census Bureau, *2013 U.S. Gazetteer Files, Urban Areas*, at: http://www.census.gov/geo/maps-data/data/gazetteer2013.html, supplemented by: http://www.zipinfo.com/search/zipcode.htm, and l*ocal area employment data* at: http://www.bls.gov/sae/data.htm.

SAR total (fixed) for apartments (core commercial) by, on average 1.78 (1.68) percentage, and in 2009:Q1 by 2.47 (2.48) percentage points. However, the results for varying annual ρ^{τ} and $\mathbf{W}_n^{\tau}$ show volatile changes that are not unexpected given that the results for ρ^{τ} for core commercial properties are in many years not statistically significant; yet, its values were included in equation (5). The standard deviation of ρ^{τ} for apartments was greater than for core commercial, at 0.34 and 0.14 percent respectively, explaining some of the volatility. The exercise argues for the testing of $\rho^{\tau} = 0$ for sub-periods τ and wider windows for the panel regressions as applicable.

D. **SAR fixed effects panel models with fixed and chain weights**

Figures 2C and 2D show fixed and chain (metro area) weighted SAR panel inflation rates against unweighted SAR estimates. The share weights were applied to the *n* diagonal elements and average column sums of the matrix $\mathbf{B}_t$ as explained in Section IID. The

weights for 2007–12. The results show that the use of weights here makes little difference, on average. The mean absolute deviation between the unweighted and fixed base-weighted variants was 0.13 and 0.11 percentage points respectively for apartments and core commercial properties. Similarly, the annual updating of the weights against the use of fixed weights found average differences of 0.03 and 0.11 percentage points for apartments and core commercial properties respectively. It may be that incorporating a spatial component removes some of the inter-area price change and reducing the need for variable weights. As reiterated by Figures 2C, the two-way unweighted OLS panel fixed effects measures are, for apartments, substantially biased upwards.

Table 2, Spatial regression diagnostics for period and annual regressions								
	Apartments				Core commercial			
		Log-	Spatial			Log-	Spatial	
	R-sqrd	likelihood	rho	std.error	R-sqrd	likelihood	rho	std.error
2001Q1-2012Q12	**0.498**	**-4408.646**	**0.30059*****	**(0.0433741)**	**0.483**	**-4,427.10**	**0.24722*****	**(0.04370)**
2001	0.723	416.0655	0.31654***	(0.1414527)	0.5442	384.5384	0.66398***	(0.09448
2002	0.655	398.0335	0.24354*	(0.1469212)	0.6353	384.271	0.27202*	(0.14967)
2003	0.705	391.6213	0.501580***	(.119102)	0.5745	394.271	0.17193	(0.15535)
2004	0.691	385.8194	0.24029	(0.1656354)	0.5459	394.754	0.26022*	(0.142735)
2005	0.722	371.326	0.09197	(0.179746)	0.514	378.0487	0.06586	(0.16838)
2006	0.528	348.5704	0.51502***	(0.1321906)	0.456	348.677	-0.01671	(0.17451)
2007	0.481	348.6964	0.36278**	(0.1462062)	0.629	346.563	0.17426	(0.17000)
2008	0.668	283.7074	0.423143***	(0.1388487)	0.562	330.673	0.23001	(0.15683)
2009	0.755	267.055	0.25073	(0.1543489)	0.291	259.895	0.18346	(0.16151)
2010	0.533	294.320	-0.67647***	(0.2049032)	0.528	247.921	-0.03083	(0.18346)
2011	0.409	269.252	0.16994	(0.159545)	0.544	269.551	-0.148533	(0.200413)
2012	0.248	206.1177	0.08556	(0.16980)	0.378	214.147	0.15444	(0.165552)

***,**,* denote statistically significant at a 1, 5, and 10 percent level respectively.
The number of observations for 2001:Q1 to 2012:Q4 is $n \times t = 34 \times 48 = 1,632$, and for each year $34 \times 4 = 136$.
Driscoll-Kraay standard errors are in parentheses.

E. Higher-level weights

The above aggregation of granular RCA indexes into higher-level indexes for apartments and core commercial properties differs from that used by RCA/Moody's for its higher-level CPPIs. The national RCA/Moody's CPPIs, as outlined in Figure A1 (Annex 1), aggregates weighted "composites" of underlying indexes of major and non-major markets for each of the five property types (apartment, retail, industrial, office CBD, and office suburban), each estimated using repeat-sales regression. Observations within these "building-block" indexes are equally weighted so that each repeat sale reflects the same contribution to the total in the calculation of each index. The building blocks, and higher aggregates thereof, are weighted using a 10-year rolling average of transaction volume as tracked by RCA to approximate the share each sector represents in the investment market.

Figure 2, SAR panel regression direct and total parameter estimates of core commercial property price inflation: quarter-on-quarter inflation, percent

Figure 2A, Apartments: with fixed and annual varying rho and spatial weight matrix

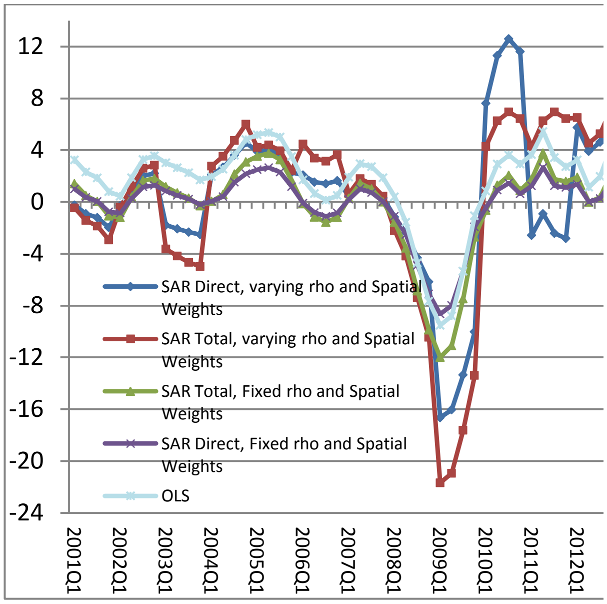


Figure 2B, Core commercial: with fixed and annual varying rho and spatial weight matrix

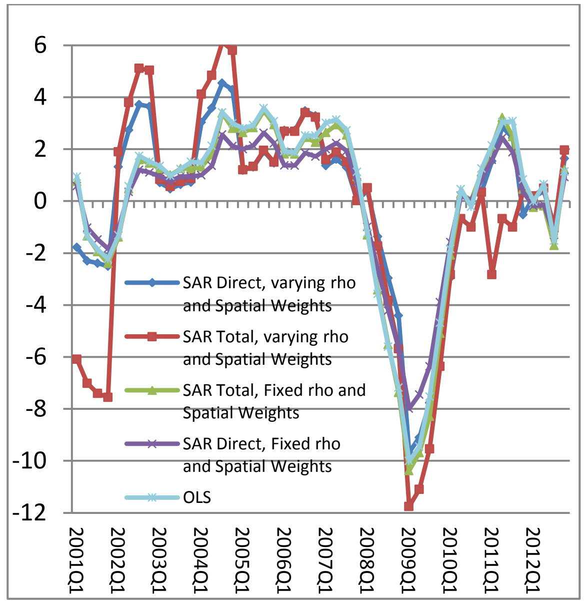


Figure 2C, Apartments: fixed and annual chain weights

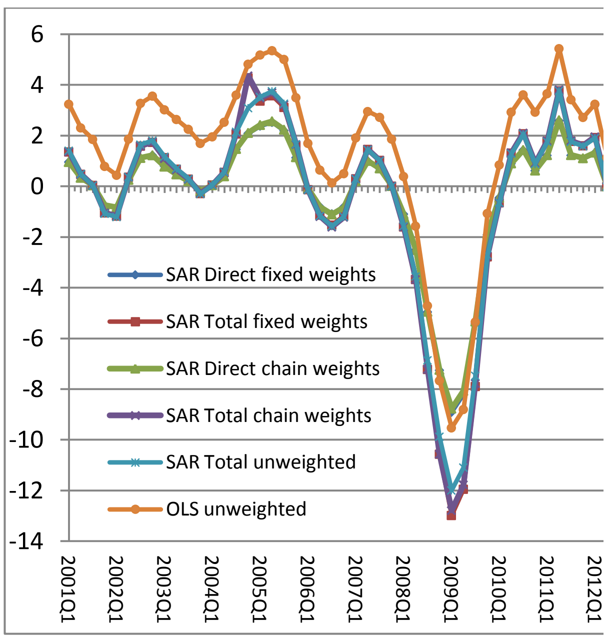
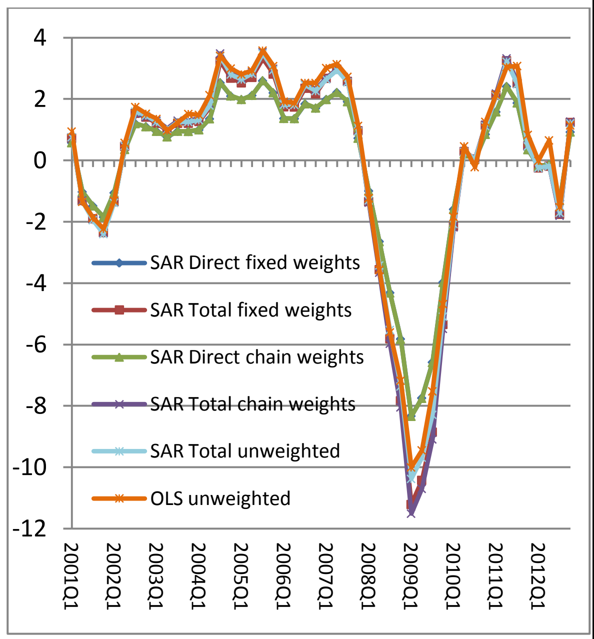


Figure 2D, Core commercial: fixed and annual chain weights



Determining the relative values for different types of the stock of the commercial real estate market in the United States is problematic—see Florance *et al*.(2010). There are three methods for estimating the market capital value of real estate: i) direct measurement of stock; ii) perpetual inventory calculations; and iii) extrapolation based on proxies (e.g. property tax assessments, population, etc). Direct measurement of stock is in principle preferable, though the decennial census conducted by the U.S. Bureau of the Census does not cover commercial real estate. The U.S. Bureau of Economic Analysis (BEA) uses the perpetual inventory method to estimate fixed capital stock, including structures by type. The method is based on investment flows and a geometric deprecation formula—see Seskin and Parker (1998)—though this approach is vulnerable to errors if the assumptions made about building permits representing flows or building economic longevity and depreciation estimates are inaccurate (Florance, *et al.*, 2010).

As described above, the Moody's/RCA CPPI weights are said to approximate the share of outstanding stock. In the table below, we compare the weights used to aggregate the Moody's/RCA building block indexes into upper level aggregate indexes, including the All Property National index,[13] with weights derived using BEA estimates of the current-cost net stock of structures by type.[14]

As seen from Table 3, the estimates for the relative stock weights for "apartments" against "core commercial" properties used by RCA/Moody's compare favorable with the BEA figures, and this continues to be the case over time, Table 4. Yet Table 3 shows the estimates to depart strikingly for more-detailed breakdowns, for example, for "industrial" properties,

[13] Data from the BEA refer to estimates for 2011. Moody's/RCA weights are taken from a relatively recent PowerPoint presentation detailing index enhancement and methodological updates; they are assumed to refer to a 10 year period ending 2011, or in close proximity to it. See www.creconsole.com/blog/wp-content/.../05/Moodys-RCA-CPPI.pdf.

[14] Moody's/RCA define an apartment as a multi-family structure with 10 units or more; whereas, for the BEA Fixed Asset Accounts, we use the break-down for a multi-family structure with 5 or more units. BEA does not disaggregate offices into central business district (CBD) versus suburban offices.

Table 3, Moody's/RCA and BEA weights: core, retail, office, industrial, and apartment					
Moody's/RCA CPPI Composite Index			BEA-based Index Weights (2011)		
Type	Weight		Type	Weight	
National All-Property Index	100.0		National All-Property Index	6213.6	100.0
Core Commercial	73.0		Core Commercial	4583.5	73.8
Retail	19.564		Retail	1155.5	18.6
Office CBD	19.491		Office	1670.6	26.9
Office Suburban	19.345				
Industrial	14.6		Industrial	1757.4	28.3
Apartment	27.0		Apartment	1630.1	26.2

Table 4, BEA and Moody's/RCA weights, percentages: apartment and core commercial CPPIs

	Moody's/RCA Weights		BEA Weights				
				Core commercial			
	Apartments	Core commercial	Apartments	All	Industrial	Office	Retail
2000	-	-	0.258	0.742	0.306	0.264	0.172
2001	0.273	0.727	0.256	0.744	0.306	0.265	0.173
2002	0.271	0.729	0.259	0.741	0.301	0.265	0.175
2003	0.267	0.733	0.261	0.739	0.296	0.266	0.178
2004	0.270	0.730	0.258	0.742	0.292	0.268	0.182
2005	0.270	0.730	0.262	0.738	0.284	0.272	0.182
2006	0.263	0.737	0.263	0.737	0.278	0.275	0.184
2007	0.264	0.736	0.262	0.738	0.277	0.274	0.186
2008	0.268	0.732	0.253	0.747	0.282	0.277	0.188
2009	0.268	0.732	0.260	0.740	0.280	0.275	0.186
2010	0.269	0.731	0.265	0.735	0.279	0.272	0.184
2011	0.275	0.725	0.262	0.738	0.280	0.273	0.185
2012	0.272	0.728	0.255	0.745	0.284	0.273	0.188

14.6 (RCA) compared to 28.3 (BEA) percent of the estimated stock in 2011 of all commercial properties. In the absence of commensurate BEA weights for all years to replicate the Moody's RCA aggregation[15] we assume the relative stability of the weights for apartments against core commercial properties in Table 3 carries over to more detailed breakdowns and use 2011 fixed base BEA weights to aggregate each of the Moody's/RCA indexes by property type (e.g. US Apartment, US Office, etc).

On average, the quarter on quarter change for the recompiled indexes (2000:Q4 – 2012:Q4) using BEA weights was 3.51 percent, while the average changes over the same period for the Moody's/RCA index was 3.69 percent. Though the difference between these averages does not appear to be substantial, though for one period the difference in the quarter-on-quarter rates of change was -9.53 percent (BEA-based weighted index) versus -10.1 percent (Moody's/RCA weighted index); notable for a single two-way stage of aggregation.

V. CONCLUSIONS

In this paper we considered methods of improving estimates from transaction data based on the (increasing) availability of granular CPPI data. Two-way fixed effects panel models were used to generate estimates of higher-level CPPI aggregates. Issues of sparse data, spatial autocorrelation, and weights were considered. The efficiency of the inflation estimates was improved using a counts-based WLS estimator and a SAR model was used to illustrate how fixed effects bias can be removed by directly incorporating spatial spillovers into the model. Restrictions on fixed ρ and $\mathbf{W}_n$ were relaxed, an innovative formulation that allows weights to be applied within the SAR model outlined, chained and fixed-base weights introduced and their effects compared. The empirical work served to demonstrate the gains in efficiency and removal of bias in estimates and gains from less-constrained formulations of ρ and $\mathbf{W}_n$. While the empirical results for chain-weighting were, in this application, disappointing, the proposed methodology should be of interest.

---

[15] BEA stock data for commercial properties by structure type on major metro versus non-major metro area are not available. Derivation of the implicit weights at the level of the building block indexes proved to be problematic.

References

Anselin, L 2008. Errors in variables and spatial effects in hedonic house price models of ambient air quality". *Empirical Economics*, 34, 1, 5–34.

Anselin L., Le Gallo J., Jayet H. 2008. Spatial Panel Econometrics, in Matyas L. and Sevestre P. Eds., *The Econometrics of Panel Data, Fundamentals and Recent Developments in Theory and Practice,* 3rd Edition, chapter 18, 627-662, Berlin, Springer-Verlag.

Belotti, Federico, Hughes, Gordon, and Andrea Piano Mortari 2013. XSMLE - A command to estimate spatial panel models in Stata, German Stata Users Group Meeting, Potsdam, June 7, 2013.

Bokhari, Sheharyar and David Geltner 2012. Estimating real estate price movement for high-frequency tradable indexes in a scarce data environment, *Journal of Real Estate Finance and Economics* 45, 2, 522-43.

Debarsy, Nicolas and Cem Ertur 2010. Testing for spatial autocorrelation in a fixed effects panel data model, *Regional Science and Urban Economics* 40, 453–470.

Devaney, Steven and Roberto Martinez Diaz 2011. Transaction Based Indices for the UK Commercial Real Estate Market: An Exploration using IPD Transaction Data, *Journal of Property Research*, 28, 4, 269-289.

Diewert, W. Erwin 2005.Weighted country product dummy variable regressions and index number formulae, *Review of Income and Wealth*, 51, 561–70.

Diewert, W. Erwin and Chihiro Shimizu 2013. A conceptual framework for commercial property price indexes, University of British Columbia, *School of Economics Discussion Paper* 13-11.

Driscoll, John C. and Aart C. Kraay 1998. Consistent covariance matrix estimation with spatially dependent data, *Review of Economics and Statistics* 80, 4, 549-60.

Elhorst, J. Paul 2010. Applied Spatial Econometrics: Raising the Bar, *Spatial Economic Analysis* 5, 1, March.

Eurostat, European Union, International Labor Organization, International Monetary Fund, Organisation for Economic Co-operation and Development, United Nations Economic Commission for Europe, The World Bank 2013. *Handbook on Residential Property Prices Indices RPPIs,* Luxembourg, European Union.

Feldstein, Martin and Charles Yuji Horioka 1980. Domestic saving and international capital flows, *Economic Journal* 90, 314–29.

Florance, Andrew C., Miller, Norm G., Peng, Ruijue, and Jay Spivey 2010. Slicing, dicing, and scoping the size of the U.S. commercial real estate market, *Journal of Real Estate Portfolio Management*, 16, 2, 101-18.

Geltner, David and Jeffrey Fisher 2007. Pricing and index considerations in commercial real estate derivatives, *The Journal of Portfolio Management* 33, 5, 99–118.

Geltner, David and Henry Pollakowski 2007. A set of indexes for trading commercial real estate based on the Real Capital Analytics Transaction Prices Database, *MIT Center for Real Estate Working Paper*, Release 2.

Giles, David E. 2011. Interpreting dummy variables in semi-logarithmic regression models: exact distributional results, University of Victoria, Department of Economics, *Econometrics Working Paper* EWP1101, January.

Hann, Jan de and W. Erwin Diewert, Hedonic regression methods, and Jan de Hann, Repeat sales methods. In Eurostat *et al.* 2013 *op. cit.*.

Hausman, J. 2001. Mismeasured variables in econometric analysis, problems from the right and problems from the left. *Journal of Economic Perspectives* 15, 4, 57-67.

Heath, Robert 2013. Why are the G-20 Data Gaps Initiative and the SDDS Plus Relevant for Financial Stability Analysis? *Journal of International Commerce, Economics and Policy*, *4*, *3*.

Kanutin, Andrew 2013. ECB progress towards a European commercial property price index. Paper presented at the 13th Ottawa Group Meeting held from 1–3 May, 2013, Copenhagen, Denmark. Available at: http://www.dst.dk/da/Sites/ottawa-group/agenda.aspx.

Kennedy, Peter, E. 1981. Estimation with correctly interpreted dummy variables in semilogarithmic equations, *American Economic Review* 71, 4, 801.

Lee, Lung-fei and Jihai Yu 2010a. Estimation of spatial autoregressive panel data models with fixed effects, *Journal of Econometrics* 154, 165–185.

Lee, Lung-fei and Jihai Yu 2010b. Some recent developments in spatial panel data models. Regional Science and Urban Economics 40, 255–271.

LeSage JP, Pace RK 2009. *Introduction to Spatial Econometrics.* Taylor & Francis, New York, FL

Leventis, Andrew 2008. Revisiting the Differences between the OFHEO and S&P/Case-Shiller House Price Indexes: New Explanations, Office of Federal Housing Enterprise Oversight, January, Available at: http://www.fhfa.gov/webfiles/1163/OFHEOSPCS12008.pdf.

Levin, Andrew, Chien-Fu Lin, and Chia-Shang James Chu 2002.Unit Root Tests In Panel Data: Asymptotic and Finite-Sample Properties, *Journal of Econometrics*, 108, 1–24.

Manski, C. F. 1993. Identification of endogenous social effects: The reflexion problem. *Review of Economic Studies* 60: 531-42.

Mason, Phil and Gwilym Pryce 2011. Controlling for transactions bias in regional house price indices, *Housing Studies*, 26, 5, July, 639-660.

Picchetti, Paulo 2013. Estimating and smoothing appraisal-based commercial real-estate performance indexes. Paper presented at the 13th Ottawa Group Meeting held from 1–3 May, 2013, Copenhagen, Denmark. Available at: http://www.dst.dk/da/Sites/ottawa-group/agenda.aspx.

Shiller, R.J., 1991. Arithmetic repeat sales price estimators, *Journal Housing Economics,* 1, 1, 110–126.

Silver, Mick 2012. Why house price indexes differ: measurement and analysis, *IMF Working Paper* WP/12/125.

Silver, Mick 2013. Understanding commercial property price indexes, *World Economics* 14, 3, September, 27-41.

Van Garderen, K. J. and C. Shah 2002. Exact interpretation of dummy variables in

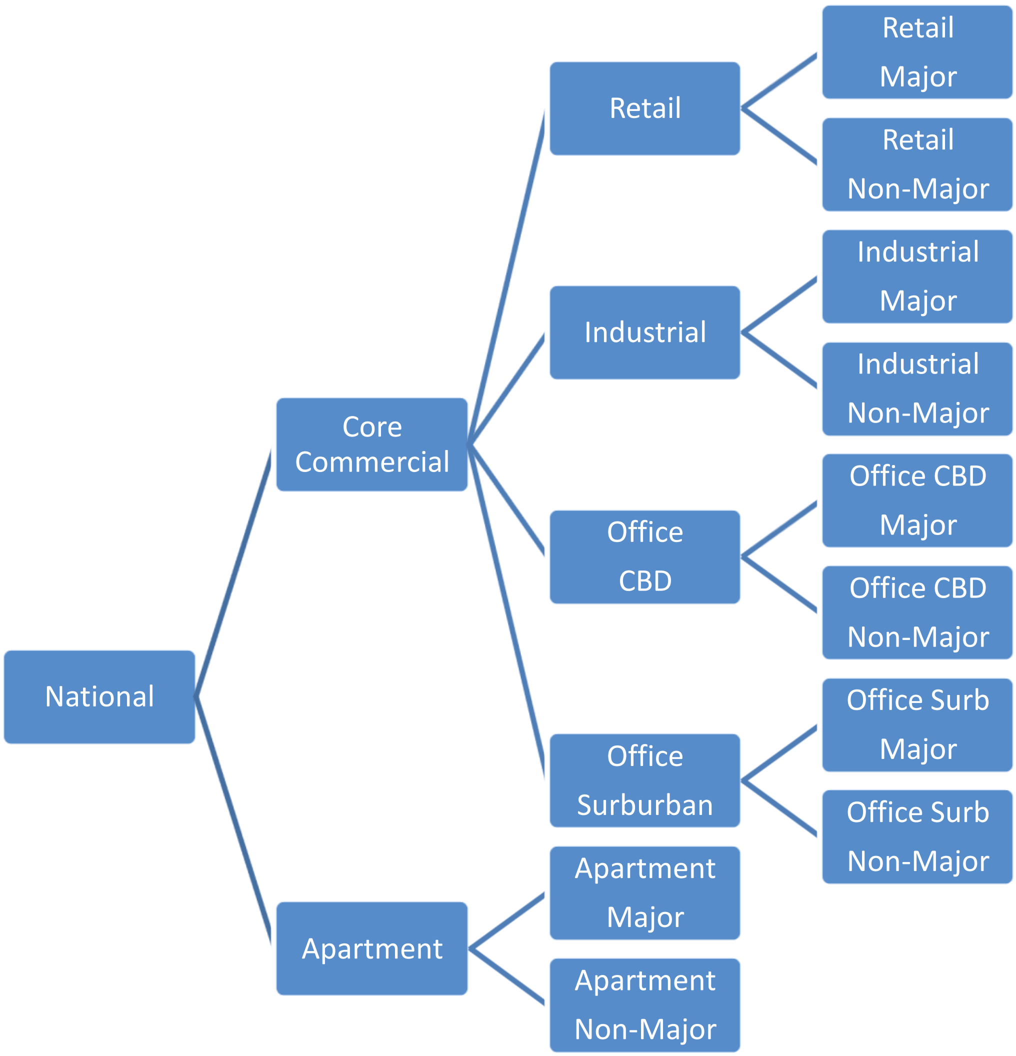
semilogarithmic equations, *Econometrics Journal*, 5, 149-159.

Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data*, Cambridge, Massachusetts and London, England: The MIT Press.

Zellner, Arnold 1962. An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias, *Journal of the American Statistical Association*, 57, 298, June, 348-368.

Annex 1, the Aggregation of Moody's/RCA CPPIs and use of Counts Data

Figure A1, Schemata of the Aggregation of Moody's/RCA CPPIs

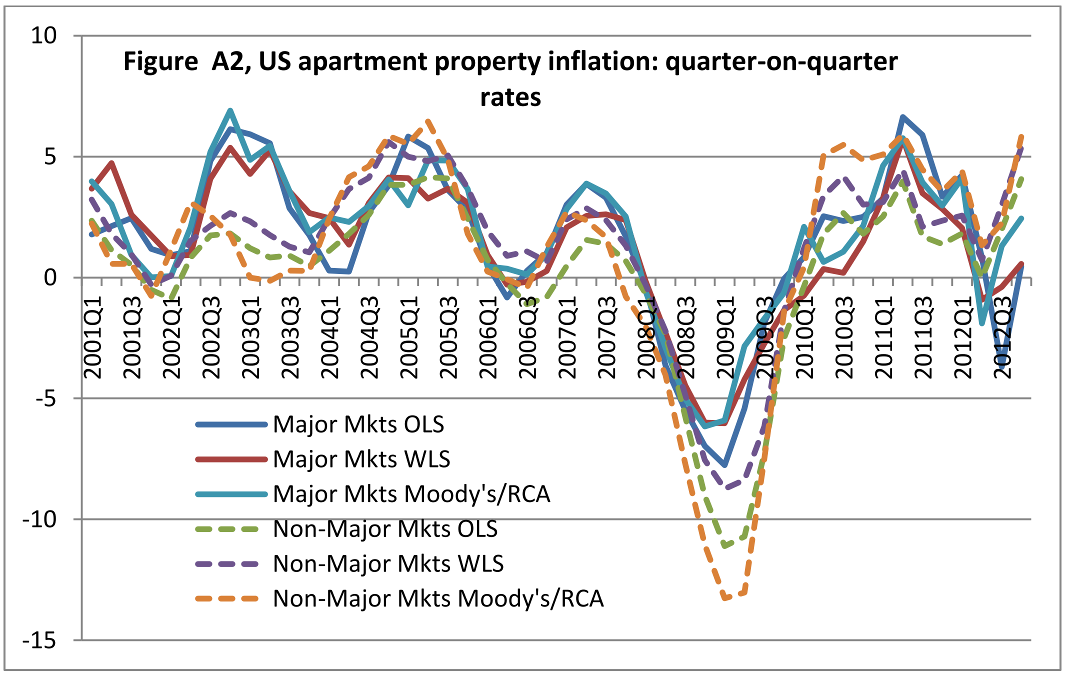


Apartment CPPIs: broken down by major and non-major markets

The RCA/Moody's aggregate CPPI for "apartments" is a weighted average of two building-block indexes: the CPPI for apartments in "Major Markets" and the CPPI for apartments in "Non-Major Markets." Major markets comprise six of the 34 MMs, namely Boston, Chicago, District of Columbia (DC), Los Angeles, New York, and San Francisco. Moody's/RCA CPPIs for each of the 34 building blocks are estimated using repeat-sales regression, as are their CPPIs for major and non-major markets.

Following the methodology outlined in section III we use equation (1) to estimate commercial property price inflation based on $\hat{\lambda}_t$ (though see *ff*. 10) for US apartments using OLS and counts-based WLS to take account of the sparse data.

We estimated apartment building property price inflation for major and non-major markets using OLS and WLS as pooled regressions on the six major-market CPPI series and again on the remaining 28 non-major market CPPIs. The results are shown in Figure A2 along with the corresponding Moody's/RCA CPPIs.



However, as noted above, Moody's/RCA CPPIs for each of major and non-major markets are estimated using repeat sales regression, and the CPPI for apartments is a weighted average of the two. Figure A2 shows all three estimates in each of major and non-major market CPPIs to track each other fairly closely with respect to the turning points, but major differences are apparent. The bottom of the trough in 2009:Q1 is underestimated by OLS, compared with WLS, by 2.4 percentage points for non-major markets and by 1.7 percentage points for major markets. This accords with our expectation for discrepancies to be smaller for major markets given their larger sample sizes, but in both cases they are sizable.

Core-commercial property price inflation

A similar analysis for commercial property price inflation disaggregated by major and non-major markets in Figure A3 shows little difference for major markets, as expected due to an aggregation across major regional metropolitan markets and property types (industrial, office, and retail) in which there are relatively large numbers of transactions. However, a marked difference arose when comparing OLS and WLS estimates for non-major markets. The trough in commercial property price inflation was overestimated (deeper) using OLS by 1.34, 1.70, and 2.19 percentage points in 2001:Q4, 2009:Q2 and 2012:Q4 respectively, though in 2010:Q3 OLS inflation *exceeded* WLS inflation by 1.49 percentage points.