

# Is There Excess Comovement in the U.S. Real Estate Markets?

Jarl Kallberg, Crocker H. Liu, and Paolo Pasquariello<sup>1</sup>

January 21, 2008

<sup>1</sup>Thunderbird School of International Management (jarl.kallberg@thunderbird.edu) and Arizona State University, W. P. Carey School of Business (jkallber@mainex1.asu.edu), Arizona State University, W. P. Carey School of Business (Crocker.Liu@asu.edu), and Ross School of Business, University of Michigan (ppasquar@bus.umich.edu), respectively. We are grateful to Sreedhar Bharath for helpful comments. The usual disclaimer applies.

## Abstract

This study addresses the recent performance of the U.S. residential real estate market. We investigate the comovement among Case-Shiller Home Price Indices for 14 metropolitan areas from January 1987 to October 2006. We identify the portion of this comovement deemed as *excessive*, which we define as the covariation that cannot be attributed to common fundamental factors (i.e., factors that directly influence real estate prices). We find that the degree of observed raw comovement in these markets increased over the sample period, most significantly so in the late 1990s, but that this increase is largely due to systemic shocks; the degree of excess comovement is a less important factor. Further analysis indicates that the dynamics of observed raw comovement among metropolitan U.S. residential real estate markets is mostly attributable to underlying systematic real and financial shocks. We conclude that *contagion* played only a minor role in the evolution of U.S real estate prices over the last two decades.

*JEL classification:* E30; R20

*Keywords:* Contagion; Excess Comovement; Correlation; Real Estate; Bubble

# 1 Introduction

The recent “bubble” in U.S. residential real estate has been the subject of an immense amount of public outcry and debate. In many publications, the housing crisis was compared to the *dot.com* bubble.<sup>1</sup> For instance, in July of 2007 Robert Shiller observed that housing prices were over-valued and a correction could cost trillions of dollars.<sup>2</sup> Numerous websites such as *housingdoom.com* have emerged adding to the housing bubble mania. Despite these passionate outpourings, very little substantive empirical analysis has been done to ascertain the degree to which these price dynamics were the result of irrationality of investors versus a more reasoned response to a diverse set of fundamental economic factors. As is typical, after the fact observers are quick to identify a bubble.

This study addresses this issue formally. We use a dataset of Case-Shiller Home Price Indices. These indices measure the path of residential housing prices in 14 major U.S. markets. Our analysis begins in January 1987 and runs through October 2006, giving us ample time to evaluate the dynamics of the recent price increases. We then decompose the comovement among these indices into fundamental covariation — which we define to be covariation arising from fundamental pricing factors (such as mortgage rates, stock market returns, etc.) — and excess covariation — which we define as the amount of covariation beyond this fundamental level — and analyze their properties. In the recent literature on financial contagion, the degree of excess correlation is used as a measure of financial contagion. For example a high level of excess correlation in these markets over the last decade would be consistent with a period of market irrationality.

Briefly noted, our results are the following. First, our analysis of comovement indicates that while, in general, the covariation in residential real estate prices rose over this period, much of such increase stems from fundamental correlation; much less can be attributed to excess correlation. This evidence is robust to various estimation procedures and statistical tests. In particular, we show that housing price fluctuations in geographically diverse U.S. metropolitan areas appear to be mostly driven by common fundamental shocks to mortgage rates, realized and expected inflation, and GDP growth. Secondly, despite economically significant heterogeneity in the extent and dynamics of observed and excess comovement across those markets, the dynamics of the former appear to be only weakly related to those of the latter. Accordingly, we find that observed raw comovement among U.S. residential real estate prices experiences economically and

---

<sup>1</sup>For instance, the popular web-based encyclopedia Wikipedia reports that “The housing bubble in the U.S. was caused by historically low interest rates, poor lending standards, and a mania for purchasing houses. This bubble is related to the stock market or dot-com bubble of the 1990s.”

<sup>2</sup>“Subprime Shockwave,” Bloomberg News, July 19, 2007.

statistically significant structural breaks upward over the sample period. Most of these breaks occur (with a single exception) in 1999, well before observers were calling the residential price increases a bubble, and cannot be explained by an increase in excess comovement. Thirdly, observed raw comovement of prices of single-home residences in the U.S. is to a large extent, yet not exclusively, driven by underlying systematic shocks in real, financial, and real estate markets.

Broadly interpreted, our results provide little or no support for the notion that the significant increase in comovement among prices of metropolitan U.S. residential real estate markets — a commonly mentioned feature of the so-called “housing bubble” — can be deemed “excessive.” Rather, that increase seems to be related to historically low interest rates, relatively poor performance of equity indices, and momentum in real estate prices.

The outline of the remainder of this study is as follows. Section 2 describes our data and our empirical methodology. Section 3 presents and interprets our results. Section 4 concludes.

## 2 Measuring excess comovement

There is some consensus in the economics and finance literature that not only periods of uncertainty but also more tranquil times may be accompanied by excess comovement among asset prices within and across both developed and emerging *financial* markets.<sup>3</sup> The objective of this study is to determine whether there is excess comovement within and across *real* markets as well. In particular, we focus on the market for single-family residences in the U.S. We define *excess comovement* in this market as comovement among housing prices beyond the degree that is justified by economic fundamentals — i.e., by factors affecting houses’ long-term valuation — and *contagion* as the circumstance of its occurrence. In this section we amend the multi-step methodology of Kallberg and Pasquariello (2007) to estimate the degree of intertemporal excess comovement among a set of  $K$  real asset prices.

### 2.1 Housing price data

The basic dataset we use in this paper consists of monthly returns for  $K = 14$  S&P/Case-Shiller Home Price Indices (CSIs,  $r_{kt}$ ) between January 1987 and October 2006. These indices correspond to 14 individual metropolitan markets: Los Angeles, San Diego, San Francisco, Denver,

---

<sup>3</sup>See, for example, the empirical studies by Pindyck and Rotemberg (1990, 1993), Karolyi and Stulz (1996), Fleming et al. (1998), Barberis et al. (2005), Bekaert, Harvey, and Ng (2005), Bekaert et al. (2005), and Kallberg et al. (2005). Many theoretical models (e.g., King and Wadhvani, 1990; Kodres and Pritsker, 2002; Veldkamp, 2005; Pasquariello, 2007) also describe contagion as a pervasive equilibrium property of financial markets.

Washington (DC), Miami, Tampa (FL), Chicago, Boston, Charlotte (NC), Las Vegas, New York, Cleveland (OH), and Portland (OR). CSI returns measure nominal price changes of individual single-family residences in the U.S. over time. Specifically, S&P/Case-Shiller collects data on all properties that have been traded more than once in each of those metropolitan areas from local deed recording offices. All resulting price changes between two arms-length sales of the same single-family home are then filtered and weighted to reflect the average change in market prices for constant-quality homes in a geographic market. For instance, CSIs do not include sales of properties that may have undergone substantial physical changes in proximity of the recorded transaction, and underweight (but do not eliminate) extreme price changes. S&P/Case-Shiller also employs an interval weighting procedure to account for the greater relevance of idiosyncratic factors (e.g., physical changes, local neighborhood effects) for price changes over longer time intervals.

Table 1 presents summary statistics for each of the 14 metropolitan indices  $r_{kt}$ , as well as for a market capitalization-weighted *composite* index,  $r_{mt}$ , based on 10 of these individual cities.<sup>4</sup> Not surprisingly, given the growth experienced by the U.S. real estate market in the past two decades, all mean monthly returns are positive, statistically significant, and greater than the average monthly nationwide inflation (0.25%, based on the CPI, all items excluding shelter). CSI returns display little or no skewness and some leptokurtosis, but are positively autocorrelated: The estimated first-order autocorrelation coefficients  $\hat{\rho}_1$  are significantly different from zero for each of the individual and the composite indices, and the corresponding values for the Ljung-Box portmanteau test for up to the fifth-order serial correlation,  $LB(5)$ , reject the null hypothesis that those  $r_{kt}$  are white noise.

## 2.2 Fundamental comovement

The starting point of the methodology in Kallberg and Pasquariello (2007) is the specification of a multi-factor model of the above real assets' returns with time-varying sensitivities. Let  $r_{kt}$  be a  $N \times 1$  vector of returns for real asset  $k$  over the interval  $[t - N + 1, t]$ . We assume that, for each asset  $k = 1, \dots, K$ , the return  $r_{kt}$  is characterized by the following linear factor structure:

$$r_{kt} = \alpha_{kt} + f_t \beta_{kt} + e_{kt}, \quad (1)$$

where  $f_t$  is a  $N \times N_f$  matrix of systematic shocks  $f_{it}$  affecting all assets and  $\beta_{kt}$  is a  $(N_f \times 1)$  vector of factor loadings. In this setting, comovement between any pairs of returns  $r_{kt}$  and  $r_{nt}$  is

---

<sup>4</sup>The S&P/CSI Composite-10 includes the metropolitan areas of Los Angeles, San Diego, San Francisco, Denver, Washington, Miami, Chicago, Boston, Las Vegas, and New York.

deemed excessive if, even after controlling for  $f_t$ , those returns are still correlated.

The selection of the appropriate set of systematic sources of risk for housing returns is a crucial step in our analysis. Any subsequent test for excess comovement is unavoidably also a test of the validity of the specification we use to control for fundamental comovement in metropolitan housing prices at each point in time  $t$ . The challenge is therefore to design a model that is both comprehensive in its scope and general in its structure. To that purpose, we propose the following specification of Eq. (1):

$$r_{kt} = \alpha_{kt} + \beta_{mt}^k r_{mt} + \beta_{SPt}^k r_{SPt} + \beta_{CPIt}^k CPI_t + \beta_{MTGt}^k MTG_t + \beta_{SLPt}^k SLP_t + \beta_{UNEt}^k UNE_t + \beta_{POPt}^k POP_t + \beta_{INCt}^k INC_t + \beta_{GDPt}^k GDP_t + e_{kt}, \quad (2)$$

for each  $k = 1, \dots, K$ . Our evidence, discussed in Section 3, indicates that the above model, albeit parsimonious, performs satisfactorily in the data.

Housing price changes in individual metropolitan areas may be affected by nationwide shocks in real estate, financial, and real markets. We employ the time series of returns for the composite CSI ( $r_{mt}$ ) and S&P500 index ( $r_{SPt}$ ) to proxy for systematic real estate and stock market risks, respectively.  $MTG_t$  is the 30-year conventional mortgage rate (from the Federal Home Loan Mortgage Corporation).  $SLP_t$ , the slope of the U.S. Treasury yield curve — computed as the monthly difference between 10-year and 3-month U.S. Treasury constant-maturity rates (from the Board of Governors of the Federal Reserve System) — is a proxy for the *real-time*, marketwide perception of the current and future state of the economy (and its inflation risk). We also control for the (possibly heterogeneous) response of metropolitan housing prices to nationwide monthly percentage changes in CPI excluding shelter (i.e., without the housing component),  $CPI_t$ . Alternatively, we find the inference that follows to be virtually unaffected by deflating all CSI returns and nominal regressors in our study by the nationwide ex-shelter inflation  $CPI_t$ . Finally, Eq. (2) allows for nationwide economic and demographic shocks to impact the transaction prices of individual single-family residences:  $UNE_t$ , the monthly change in the civilian unemployment rate (from the Bureau of Labor Statistics);  $POP_t$ , the monthly change in total U.S. population;  $INC_t$ , the monthly percentage change in disposable personal income; and  $GDP_t$ , the interpolated monthly percentage change in nominal Gross Domestic Product (GDP), all from the U.S. Department of Commerce. Plots of each of these variables over our sample period (on the left axis in Figures 1a to 1i) display some of the familiar trends for the U.S. economy over the last two decades: The longest period of economic expansion in U.S. history (from 1991 to 2001, e.g., see  $r_{SPt}$ ,  $INC_t$ , and  $GDP_t$  in Figures 1b, 1h, and 1i) is accompanied by declining mortgage rates ( $MTG_t$  in Figure 1d), a significant increase in prices of single-family residences nationwide ( $r_{mt}$

in Figure 1a), yet only moderate expected and realized inflation (e.g., see  $SLP_t$  and  $CPI_t$  in Figures 1e and 1c).

### 2.3 Latent comovement

The next step in our methodology to measure excess comovement among housing price changes is the estimation of the parameters in Eq. (1), using sample data for the period  $[0, T]$  and the selected set of systematic factors described above. The resulting  $N \times 1$  vector of estimated residuals  $\widehat{e}_{kt}$ , where

$$\widehat{e}_{kt} = r_{kt} - \widehat{\alpha}_{kt} - f_t \widehat{\beta}_{kt}, \quad (3)$$

are in fact meant to capture the dynamics of individual CSIs that cannot be explained by those common factors.

To that purpose, we specify the following stacked version of the model of Eq. (1) over the interval  $[t - N + 1, t]$ :

$$\begin{bmatrix} r_{1t} \\ r_{2t} \\ \vdots \\ r_{Kt} \end{bmatrix} = \begin{bmatrix} F_t & O & \cdots & O \\ O & F_t & \cdots & O \\ & & \vdots & \\ O & O & \cdots & F_t \end{bmatrix} \begin{bmatrix} B_{1t} \\ B_{2t} \\ \vdots \\ B_{Kt} \end{bmatrix} + \begin{bmatrix} e_{1t} \\ e_{2t} \\ \vdots \\ e_{Kt} \end{bmatrix} = F_t B_t + e_t, \quad (4)$$

where  $F_t = [f_t, f_t]$  is a  $N \times M$  matrix of systematic factors affecting  $r_{kt}$  (in which  $f_t$  is a  $N \times 1$  unit vector and  $M = N_f + 1$ ),  $B_{kt} = [\alpha_{kt}, \beta'_{kt}]'$  is a  $M \times 1$  vector of factor loadings, and  $O$  is a zero matrix. We further assume that the  $N \times 1$  vectors of disturbances  $e_{kt}$  are uncorrelated across observations, i.e., that  $E[e_{kt}e'_{ns}] = \sigma_{knt}I_N$  (where  $I_N$  is a  $N \times N$  identity matrix) if  $t = s$  and  $E[e_{kt}e'_{ns}] = O$  otherwise. This implies that

$$E[e_t e_t'] = V_t = \begin{bmatrix} \sigma_{11t} & \sigma_{12t} & \cdots & \sigma_{1Kt} \\ \sigma_{21t} & \sigma_{22t} & \cdots & \sigma_{2Kt} \\ & & \vdots & \\ \sigma_{K1t} & \sigma_{K2t} & \cdots & \sigma_{KKt} \end{bmatrix} \otimes I = \Sigma_t \otimes I. \quad (5)$$

The seemingly unrelated regressions model of Eqs. (4) and (5) allows for the parameters controlling for fundamental comovement across assets to vary over time. It can be efficiently estimated via either ordinary least squares — OLS, i.e., separately, under the null hypothesis that the returns  $r_{kt}$ , after controlling for systematic sources of risk, are independent — or feasible generalized least squares — FGLS, i.e., jointly, under the alternative hypothesis that return

residuals do comove beyond what is justified by common economic fundamentals — since the  $K$  stacked regressions in Eq. (4) have identical explanatory variables  $F_t$  (see Greene, 1997, pp. 676). The efficient estimator of  $B_t$  is then given by

$$\widehat{B}_t^{OLS} = (F_t' F_t)^{-1} F_t' r_t, \quad (6)$$

where  $r_t = [r_{1t}, r_{2t}, \dots, r_{Kt}]'$ .

## 2.4 Excess comovement

Lastly, and consistently with the above discussion, we define comovement between any pairs of returns  $r_{kt}$  and  $r_{nt}$  as *excessive* if the corresponding return residuals from the OLS estimation of Eqs. (4) and (5),  $\widehat{e}_{kt}^{OLS} = r_{kt} - F_t \widehat{B}_{kt}^{OLS}$  and  $\widehat{e}_{nt}^{OLS} = r_{nt} - F_t \widehat{B}_{nt}^{OLS}$ , are correlated. Specifically, we use these residuals to produce an efficient estimate of the unknown matrix  $\Sigma_t$ ,  $\widehat{\Sigma}_t^{OLS}$ , whose individual elements are

$$\widehat{\sigma}_{knt}^{OLS} = \frac{(\widehat{e}_{kt}^{OLS})' \widehat{e}_{nt}^{OLS}}{N}, \quad (7)$$

and those individual elements to compute, for each  $k \neq n$ , excess correlation coefficients

$$\widehat{\rho}_{knt}^{OLS} = \frac{\widehat{\sigma}_{knt}^{OLS}}{[\widehat{\sigma}_{kkt}^{OLS} \widehat{\sigma}_{nnt}^{OLS}]^{\frac{1}{2}}}. \quad (8)$$

Several studies show that these correlation coefficients are *conditional* on return volatility; thus, in the presence of heteroskedasticity, contagion tests based on them may be biased toward rejection of the null hypothesis of no excess comovement (e.g., Boyer et al., 1999; Loretan and English, 2000; Forbes and Rigobon, 2002). Yet, these studies also argue that this bias can be corrected by computing an *unconditional* correlation measure for any pair of returns under the assumption of no omitted variables or endogeneity. The measure of excess comovement between any two CSI returns  $r_{kt}$  and  $r_{nt}$  that we ultimately adopt in this paper is based on their proposed adjustment and is given by

$$\widehat{\rho}_{knt}^{OLS*} = \frac{\widehat{\rho}_{knt}^{OLS}}{\left\{ 1 + \widehat{\delta}_{kt}^{OLS} \left[ 1 - \frac{(\widehat{\sigma}_{knt}^{OLS})^2}{\widehat{\sigma}_{kkt}^{OLS} \widehat{\sigma}_{nnt}^{OLS}} \right] \right\}^{\frac{1}{2}}}, \quad (9)$$

where the ratio  $\widehat{\delta}_{kt}^{OLS} = \frac{\widehat{\sigma}_{kkt}^{OLS}}{(\widehat{\sigma}_{kkt}^{OLS})_{LT}} - 1$ , when different from zero, corrects the conditional correlation  $\widehat{\rho}_{knt}^{OLS}$  of Eq. (8) for the relative difference between short-term volatility ( $\widehat{\sigma}_{kkt}^{OLS}$ ) and long-term volatility of  $r_{kt}$  ( $(\widehat{\sigma}_{kkt}^{OLS})_{LT}$ ), for any  $n \neq k$ . When computing Eq. (9), we alternatively assume

that the source of volatility shocks is either metropolitan market  $k$  (in  $\widehat{\rho}_{knt}^*$ ) or market  $n$  (in  $\widehat{\rho}_{nkt}^*$ ). This implies that  $\widehat{\rho}_{knt}^{OLS*}$  may be different from  $\widehat{\rho}_{nkt}^{OLS*}$ .

We then compute arithmetic means of pairwise adjusted correlation coefficients for each market  $k$ , along the lines of King et al. (1994). Much of the literature on financial contagion explores the circumstances in which correlation among asset prices becomes more positive (or less negative) during crisis periods. However, as is clear from Eq. (1), there is no reason to restrict the concept of excess comovement to a specific directional move in the correlation coefficients. In other words, both  $\widehat{\rho}_{knt}^{OLS*} \neq 0$  and  $\widehat{\rho}_{nkt}^{OLS*} \neq 0$  represent evidence of comovement between metropolitan markets  $k$  and  $n$  beyond what is implied by their fundamentals, regardless of their sign. Thus, we need a contagion measure that prevents such coefficients, if of different sign, from cancelling each other out in the aggregation. This measure needs to control for sample variation as well. Because of sample variation in the estimators  $\widehat{B}_{kt}^{OLS}$ , residuals' correlations  $\widehat{\rho}_{knt}$  are in fact estimated with error over  $N$  observations. Hence, failure to account for statistically insignificant  $\widehat{\rho}_{knt}$  may bias our analysis of the significance and extent of excess CSI return comovement. Consequently, we focus only on statistically significant conditional correlations for each market  $k$  among the  $K - 1$  possible  $\widehat{\rho}_{knt}$ , according to the  $t$ -ratio test  $\widehat{t}_{knt}^{OLS} = \widehat{\rho}_{knt}^{OLS} \left[ \frac{1 - (\widehat{\rho}_{knt}^{OLS})^2}{N - 2} \right]^{-\frac{1}{2}} \sim t[N - 2]$ , and measure excess comovement by computing the following means of excess square correlations:

$$\widehat{\rho}_{kt}^{OLS*} = \frac{1}{K - 1} \sum_{\substack{n=1 \\ n \neq k}}^K (\widehat{\rho}_{knt}^{OLS*})^2 I_{knt}^{OLS}, \quad (10)$$

where  $I_{knt}^{OLS} = 1$  if  $2 [1 - \Pr \{ |\widehat{t}_{knt}^{OLS}| \leq t_{\frac{\alpha}{2}} [N - 2] \}] \leq \alpha$  and  $I_{knt}^{OLS} = 0$  otherwise, for any  $k = 1, \dots, K$ . Eq. (10) implies that there is statistically significant excess comovement for metropolitan market  $k$  at time  $t$  if  $\widehat{\rho}_{kt}^{OLS*}$  is different from zero.<sup>5</sup> We also compute a nationwide measure of excess comovement among housing price changes as a mean of means of excess square correlation coefficients across all of the 14 metropolitan areas, i.e.,

$$\widehat{\rho}_t^{OLS*} = \frac{1}{K} \sum_{k=1}^K \widehat{\rho}_{kt}^{OLS*}. \quad (11)$$

We are further interested in the evolution of these measures over time while accounting for the dynamics of both the fundamental interdependence among CSI returns and their variances. Ignoring time-varying factor loadings and nonstationary variances in Eq. (4) may bias the inference on non-fundamental comovement from Eqs. (10) and (11). Parametric ARCH and stochastic volatility models, as well as their generalizations to multivariate settings, are frequently

---

<sup>5</sup>In Section 3.2, we analyze the robustness of our inference to this and other features of our empirical specification.

employed to describe such dynamics.<sup>6</sup> Nonetheless, these models are in general very difficult to estimate and do not offer a clear advantage over simpler, nonparametric approaches, especially when used to measure covariance rather than forecast it (e.g., Campbell et al., 2001). Further, Campbell et al. (1997) observe that rolling filters, like the rolling standard deviation measure used by Officer (1973), usually provide very accurate descriptions of historical variation (or comovement), in particular (as shown in Nelson, 1992) when volatility (or covariance) changes are not too gradual.

In light of these considerations, we construct time series of rolling realized excess square correlations for each metropolitan market  $k$  by treating the covariance matrix of return residuals as observable. To that purpose, we estimate  $\widehat{e}_{kt}^{OLS}$ ,  $\widehat{\delta}_{kt}^{OLS}$ , and  $\widehat{\rho}_{knt}^{OLS*}$  over rolling short-term and long-term intervals of the data of fixed length  $N$  and  $gN$  (with  $g > 1$ ), respectively, according to the following scheme, in the spirit of Campbell et al. (2001) and Kallberg and Pasquariello (2007):

$$\begin{array}{c}
 \widehat{e}_{kt}^{OLS}, \widehat{\sigma}_{kkt}^{OLS} \\
 \overbrace{\hspace{10em}} \\
 \dots \left| \hspace{2em} \left| \hspace{2em} \right. \right. \dots \\
 \hspace{2em} t - gN + 1 \hspace{4em} t - N + 1 \hspace{4em} t \\
 \underbrace{\hspace{10em}} \\
 (\widehat{e}_{kt}^{OLS})_{LT}, (\widehat{\sigma}_{kkt}^{OLS})_{LT}
 \end{array} \tag{12}$$

Specifically, at each point in time  $t$  and for each CSI  $k$ , the model of Eq. (4) is estimated twice, once over the short-term interval  $[t - N + 1, t]$  to compute  $\widehat{\rho}_{knt}^{OLS}$  as in Eq. (8), and once over the long-term interval  $[t - gN + 1, t]$  to compute  $(\widehat{\sigma}_{kkt}^{OLS})_{LT}$ , the resulting adjustment ratio  $\widehat{\delta}_{kt}^{OLS}$ , and eventually  $\widehat{\rho}_{knt}^{OLS*}$  as in Eq. (9), for each  $n \neq k$ . This rolling procedure generates time series of metropolitan and nationwide excess comovement measures  $\{\widehat{\rho}_{kt}^{OLS*}\}_{t=gN}^T$  and  $\{\widehat{\rho}_t^{OLS*}\}_{t=gN}^T$ , which we use in our analysis, without resorting to parametric specifications for the intertemporal dynamics of the covariance matrix of CSI returns.

---

<sup>6</sup>See Campbell et al. (1997) for a review of the literature on parametric models of changing volatility.

### 3 Empirical analysis

#### 3.1 OLS comovement

We determine the extent and dynamics of excess comovement among the housing price data described in Section 2.1 according to the procedure described in Sections 2.2 to 2.4. We begin by estimating the fundamental model of Eq. (2) for each metropolitan market in our database over rolling intervals of two and a half years ( $N = 30$ ) within the sample interval January 1987 to October 2006. We use the resulting estimated residuals to compute conditional measures of excess CSI return comovement ( $\widehat{\rho}_{knt}^{OLS}$  of Eq. (8)) for each pair of metropolitan markets  $k$  and  $n$ ; yet, in each month  $t$  we retain only the  $\widehat{\rho}_{knt}^{OLS}$  that are statistically significant at the 10% level (i.e.,  $\alpha = 0.10$  in  $I_{knt}$  of Eq. (10)). We then correct those correlations for shifts in conditional volatility (see Eq. (9)) by estimating long-term variances of CSI return residuals over a five-year interval (i.e.,  $g = 2$  in Eq. (12)). Therefore, the initial  $t = gN$  corresponds to January 1992.

We plot the resulting time series of nationwide excess unconditional CSI return comovement —  $\{\widehat{\rho}_t^{OLS*}\}_{t=gN}^T$  of Eq. (11) — in Figure 2a. There we also plot two benchmark measures of raw nationwide square correlation:  $\{\widehat{\rho}_t^{BASE*}\}_{t=gN}^T$  — computed using unconditional correlations among all the raw CSI return series  $r_{kt}$  (instead of the OLS residuals  $\widehat{e}_{kt}^{OLS}$ ) in Eqs. (8) to (11) (i.e.,  $\widehat{\rho}_{knt}^{BASE}$ ) — and its conditional equivalent  $\{\widehat{\rho}_t^{BASE}\}_{t=gN}^T$  — computed using conditional correlations among those returns (i.e.,  $\widehat{\rho}_{knt}^{BASE}$ ). Figure 2a reveals that the adjustment of Eq. (9) for conditional variance essentially rescales the conditional correlations while preserving their intertemporal dynamics. By construction, both  $\widehat{\rho}_t^{BASE*}$  and  $\widehat{\rho}_t^{BASE}$  account for possibly non-stationary CSI return variances but not for the extent and dynamics of returns' fundamental interdependence. As such, both these benchmarks allow us to gauge the economic significance of levels and fluctuations of  $\widehat{\rho}_t^{OLS*}$ . We also plot our measures of excess comovement and their corresponding benchmarks across each of the 14 U.S. metropolitan markets listed in Table 1 (i.e.,  $\widehat{\rho}_{kt}^{OLS*}$  of Eq. (10) and  $\widehat{\rho}_{kt}^{BASE*}$ ) in Figures 3a to 3j (left axis). Table 2 reports summary statistics for each of these measures.

According to Figures 2 and 3, raw comovement among price changes of single-family homes traded in different geographical markets increases substantially in the late 1990s — i.e., almost simultaneously with the increase in residential real estate prices nationwide (e.g., see the CSI composite index, with base 100 in December 1991, on the right axis of Figure 2a) — and stays high afterwards, both nationally and locally. For example,  $\widehat{\rho}_t^{BASE*}$  in Figure 2a roughly doubles in the second half of the sample period (0.515 on average for the interval 1998-2006) as compared to its mean in 1997 (0.259). Urban areas in both the West and the East Coasts where prices

of single-family homes increase the most appear to experience the most dramatic increases in price comovement with the rest of the country (e.g., San Diego in Figure 3b and New York in Figure 3l); yet, this pattern is common to most of the U.S. metropolitan markets in our dataset. Similarly, square correlation of raw CSI returns  $\widehat{\rho}_{kt}^{BASE*}$  is generally decreasing during the period of stable or protractedly declining housing prices that followed the brief recession of 1990-1991 (e.g., Los Angeles in Figure 3a).

Further analysis reveals that an economically significant portion of the observed housing price comovement within the U.S. (i.e., of  $\widehat{\rho}_t^{BASE*}$  and  $\widehat{\rho}_{kt}^{BASE*}$ ) is explained by common fundamentals, although excess comovement does play a role. First, the fundamental OLS regression model of Eq. (2) performs quite successfully. For instance, the mean conventional adjusted  $R^2$  across all metropolitan markets ( $\overline{R}_{at}^2$  in Figure 2b) averages about 50% and is never lower than 33% over the entire sample.<sup>7</sup> More interestingly, the average explanatory power of the systematic shocks listed in Section 2.2 steadily increases toward 70% since 1997, i.e., almost simultaneously with the nationwide and local increase in raw CSI return comovement reported in Figures 2a and 3.

To assess the relative contribution of each of those shocks  $f_{it}$  to the increasing explanatory power of the multiple regression model of Eq. (2) for CSI returns, we compute two measures of their impact on that model's overall fit. The first one,  $\overline{R}_{f_{it}}^2$ , is the nationwide average of the  $R^2$  of the regression of *exclusively* factor  $f_{it}$  (and a constant term) on each CSI return series  $r_{kt}$  separately over each rolling interval of length  $N$  in Eq. (12). The second one,  $\overline{r}_{f_{it}}^2$ , addresses possible multicollinearity bias in  $\overline{R}_{f_{it}}^2$ : It is the nationwide average of factor  $f_{it}$ 's square partial correlation  $r_{r_{kt}, f_{it}}^2 = \frac{t_{r_{kt}, f_{it}}^2}{t_{r_{kt}, f_{it}}^2 + N - K}$  — where  $t_{r_{kt}, f_{it}}^2$  is the square of the  $t$ -ratio for testing the hypothesis that the coefficient on  $f_{it}$  is zero in Eq. (2) — over the same interval  $[t - N + 1, t]$ .<sup>8</sup> The extent and dynamics of both measures, on the right axis of Figures 1a to 1i, suggest that the simultaneous increase in raw CSI return comovement nationwide ( $\widehat{\rho}_t^{BASE*}$  in Figure 2a) and Eq. (2)'s overall fit ( $\overline{R}_{at}^2$  in Figure 2b) since 1998 may be attributed to the increasing importance of common real estate shocks ( $r_{mt}$  in Figure 1a) and fluctuations in realized inflation ( $CPI_t$  in Figure 2c), lending terms ( $MTG_t$  in Figure 1d), the slope of the U.S. Treasury yield curve ( $SLP_t$  in Figure 1e), and the growth of both the U.S. population and GDP ( $POP_t$  and  $GDP_t$ , in Figures

<sup>7</sup>In addition, the mean overall goodness-of-fit measure of McElroy (1977) for the model of Eq. (4) — computed as  $R_{*t}^2 = 1 - K \left\{ tr \left[ \left( \widehat{\Sigma}_t^{OLS} \right)^{-1} S_t \right] \right\}^{-1}$ , where  $S_t$  is a  $K \times K$  matrix in which  $S_{knt} = N^{-1} r'_{kt} r_{nt} - \overline{r}_{kt} \overline{r}_{nt}$  — is always greater than 99%, while the standard  $F$  test always strongly rejects the null hypothesis that all the slopes in that model ( $B_t$  in Eq. (4)) are zero.

<sup>8</sup>Thus, the ensuing  $\overline{r}_{f_{it}}^2$  can be interpreted as the average percentage of the *unique* variance in CSI returns that is *uniquely* accounted for by the common factor  $f_{it}$  after both the CSI returns and  $f_{it}$  have been controlled by all of the other common factors in Eq. (2).

1g and 1i) for the price changes of individual single-family residences in the U.S.

Nevertheless, our evidence also suggests that there is comovement beyond what can be explained by the fundamental model of Section 2.2 in the U.S. real estate markets. In particular, nationwide excess square correlation  $\widehat{\rho}_t^{OLS*}$  averages about 0.112 — equivalent to an average absolute return residual correlation of  $\left| \sqrt{\widehat{\rho}_t^{OLS*}} \right| = 0.335$  — and is statistically different from zero (according to either the  $t$ -square ratio test  $\widehat{t}_t^2 = \widehat{\rho}_t^* [1 - \widehat{\rho}_t^*]^{-1} \sim F[1, N - 2]$  in Table 2 or a standard  $t$ -test for means not reported here) over the entire sample and across all metropolitan markets in each month  $t$ . Excess unconditional comovement is economically significant as well, for it constitutes almost 30% of the benchmark raw square correlation of CSI returns,  $\widehat{\rho}_t^{BASE*}$  (Column  $R\rho$  in Table 2). Equivalently, the model of Eq. (4) can explain much — i.e., about  $1 - R\rho \approx 70\%$  — but not all of the observed raw CSI return comovement. The correlation between  $\widehat{\rho}_t^{BASE*}$  and  $\widehat{\rho}_t^{OLS*}$  (Column  $\beta\rho$  in Table 2) is low both nationwide and in most metropolitan markets. Yet, the extent and relative importance of excess comovement of single-family housing prices in the U.S. fluctuates considerably over time. For example,  $\widehat{\rho}_t^{OLS*}$  in Figure 2a is the highest in the early to mid-1990s, when it accounts for between 50% and 90% of the raw comovement measure  $\widehat{\rho}_t^{BASE*}$ . Subsequently,  $\widehat{\rho}_t^{OLS*}$  decouples from  $\widehat{\rho}_t^{BASE*}$ :  $\widehat{\rho}_t^{BASE*}$  increases and stays high in correspondence with the “real estate bubble” of the late 1990s-early 2000s, while excess square correlations drop to near-historical lows before increasing again by the end of the sample period, accompanying the sharp nationwide decline of residential housing prices in 2005 and 2006.

Similar conclusions are reached when examining the level and dynamics of raw and excess comovement for each of the metropolitan markets listed in Table 1, i.e. of  $\widehat{\rho}_{kt}^{BASE*}$  and  $\widehat{\rho}_{kt}^{OLS*}$ , respectively. However, these measures display significant cross-market variability. For instance, the observed raw comovement in housing price changes in Las Vegas with the rest of the U.S. is the lowest (0.275 in Table 2), but the corresponding ratio between  $\widehat{\rho}_{kt}^{OLS*}$  and  $\widehat{\rho}_{kt}^{BASE*}$  is the highest ( $R\rho = 47\%$  in Table 2). Both raw and excess price comovement are high in California, Denver, Washington, Boston, and New York, but each corresponding  $\widehat{\rho}_{kt}^{OLS*}$  appears to be only weakly related to its benchmark  $\widehat{\rho}_{kt}^{BASE*}$  (Column  $\beta\rho$  in Table 2). A notable exception is San Francisco (Figure 3c), arguably the urban area hit the hardest by the implosion of the “internet bubble” in early 2000s: During that time, observed single-home price changes shortly decline and comove significantly less with the rest of the country than in the 1990s; those prices’ subsequent increase is instead accompanied by sharply higher raw and excess comovement of their dynamics with those observed nationwide. Consistently,  $\beta\rho$  for San Francisco is the highest (0.225) in Table 2. We observe a similar trend in both  $\widehat{\rho}_{kt}^{BASE*}$  and  $\widehat{\rho}_{kt}^{OLS*}$  in proximity of the feverish increase in residential housing prices in Las Vegas, between 2003 and 2005 (Figure 3k). We explore in greater

depth the dichotomy in the dynamics of raw and excess comovement among U.S. residential real estate prices in Section 3.3.

## 3.2 Robustness tests

In this section we discuss whether altering any of the main features of the methodology described in Section 2 to measure raw and excess comovement among U.S. residential housing prices may affect the above inference significantly. We do not report most of these additional results for economy of space; yet, they are available on request.

### 3.2.1 Absolute correlation

We begin by computing an alternative proxy for raw and excess comovement based on means of excess *absolute* (rather than square) CSI return correlations. Specifically, we redefine  $\widehat{\rho}_{kt}^{OLS*} = \frac{1}{K-1} \sum_{\substack{n=1 \\ n \neq k}}^K |\widehat{\rho}_{knt}^{OLS*}| I_{knt}^{OLS}$  and  $\widehat{\rho}_{kt}^{BASE*} = \frac{1}{K-1} \sum_{\substack{n=1 \\ n \neq k}}^K |\widehat{\rho}_{knt}^{BASE*}| I_{knt}^{BASE}$ ,<sup>9</sup> and perform all subsequent steps listed in Section 2 accordingly. This definition is as effective as Eq. (10) at preventing excess correlation coefficients of opposite signs from canceling out when averaged across metropolitan markets, although its statistical properties are less well-known. Our main results nonetheless remain.

### 3.2.2 Sample variation

Next, we consider whether sample variation in the estimated correlations of CSI return residuals,  $\widehat{\rho}_{knt}^{OLS}$ , may bias our empirical tests. We preliminarily addressed this issue in Section 2.4 by computing each aggregate measure  $\widehat{\rho}_{kt}^{OLS*}$  only from statistically significant correlation coefficients at the 10% level ( $\alpha = 0.10$ ), as in Kallberg and Pasquariello (2007). The resulting percentage of statistically significant conditional correlations from  $\widehat{\Sigma}_t^{OLS}$  (column  $T\rho$  of Table 2) is relatively high, 33% (versus a starting point of 81% of  $\widehat{\rho}_{knt}^{BASE}$  among the raw CSI return series  $r_{kt}$ ), and fluctuates between 20% and 40% for most of the sample period. As importantly, the ensuing unconditional coefficients of determination  $(\widehat{\rho}_{knt}^{OLS*})^2$ , in column  $F\rho^{*2}$  of Table 2, are also statistically significant (according to the  $t$ -square ratio test  $\widehat{t}_{knt}^2$ ) and with almost identical frequencies across metropolitan areas and over time. Furthermore, the Lagrange multiplier (LM) statistic of Breusch and Pagan (1980) —  $\widehat{\lambda}_t = N \sum_{k=2}^K \sum_{n=1}^{k-1} (\widehat{\rho}_{knt}^{OLS})^2 \sim \chi^2 \left[ \frac{K(K-1)}{2} \right]$  — leads us to reject the

---

<sup>9</sup>Consistent with Section 2.4,  $I_{knt}^{BASE} = 1$  if  $2 [1 - \Pr \{ |\widehat{t}_{knt}^{BASE}| \leq t_{\frac{\alpha}{2}} [N-2] \}] \leq \alpha$  and zero otherwise, where

$$\widehat{t}_{knt}^{BASE} = \widehat{\rho}_{knt}^{BASE} \left[ \frac{1 - (\widehat{\rho}_{knt}^{BASE})^2}{N-2} \right]^{-\frac{1}{2}}.$$

null hypothesis that the matrix  $\Sigma_t$  in Eq. (5) is diagonal (at any conventional significance level) in each month  $t$ . Finally, we find our inference to be qualitatively unaffected by employing either a more restrictive ( $\alpha = 0.05$ ) or no significance threshold ( $\alpha = 1$ ) for the inclusion of pairwise conditional correlations  $\widehat{\rho}_{knt}^{OLS}$  and  $\widehat{\rho}_{knt}^{BASE}$  in the excess and benchmark comovement measures  $\widehat{\rho}_{kt}^{OLS*}$  (Eq. (10)) and  $\widehat{\rho}_{kt}^{BASE*}$ , respectively.

### 3.2.3 Conditional comovement

We also find our inference in Section 3.1 to be insensitive to employing unadjusted (i.e., conditional) correlations ( $\widehat{\rho}_{knt}^{OLS}$  of Eq. (8) and the corresponding  $\widehat{\rho}_{knt}^{BASE}$ ) or computing the adjustment ratio for conditional correlations ( $\widehat{\delta}_{kt}$  in Eq. (9)) over several alternative rolling intervals ( $N$  and  $g$ ) for short-term and long-term CSI return volatility. In the latter case, the resulting adjustment for heteroskedasticity may be incorrect — and the corresponding estimates for unconditional correlation inaccurate — in the presence of omitted variables or endogeneity between assets, unless when assets are “closely connected” (Forbes and Rigobon, 2002, p. 2255). This appears to be the case in our sample. For instance, unconditional correlations of raw CSI returns are large and statistically significant: E.g., the average  $\widehat{\rho}_t^{BASE*} = 0.396$ ,  $T\rho = 81\%$ , and  $F\rho^{*2} = 83\%$  in Table 2, while the corresponding LM statistic  $\widehat{\lambda}_t^{BASE}$  is strongly statistically significant (at the 1% level or less) in each month  $t$ .<sup>10</sup>

### 3.2.4 Alternative distributional assumptions

Lastly, we find our inference to be robust to plausible distributional assumptions of the elliptical class for CSI return residuals  $e_{kt}$ . Specifically, 10,000 Monte Carlo simulations of  $K = 14$  *independent* vectors of residuals  $e_{kt}$  with  $N = 30$  observations each indicate that our selection procedure based on pairwise  $t$ -ratio tests (Eq. (10)) does not reject the null hypothesis of no excess comovement “too often” (for the chosen  $\alpha = 0.10$ ) either under the assumption of either i.i.d. normality or under the multivariate  $t$  distribution with no dependence — a popular alternative (e.g., Zhou, 1993) when returns are leptokurtic (see Table 1). For instance, we find that  $T\rho = 10.0131\%$  when the vector  $e_{kt} \sim N(\underline{0}, I)$  and  $\underline{0}$  is a zero vector, while  $T\rho = 10.0052\%$  when  $e_{kt} = Z \left( \frac{x_{kk}}{v} \right)^{-\frac{1}{2}}$ ,  $Z \sim N(\underline{0}, I)$ ,  $x_{kk} \sim \chi^2[v]$ , and  $v = 7$  degrees of freedom.<sup>11</sup>

<sup>10</sup>In addition, according to Forbes and Rigobon (2002), unaccounted feedback from asset  $n$  to asset  $k$  may lead  $\widehat{\rho}_{knt}^{OLS*}$  to *underestimate* the true unconditional correlation of CSI return residuals, i.e., may bias our inference toward *acceptance*, rather than rejection, of the null hypothesis of no excess comovement.

<sup>11</sup>Our inference is also unlikely to be affected by bias in the  $t$ -ratio test  $\widehat{t}_{knt}$  in Eq. (10) under the null hypothesis that CSI return residuals are *uncorrelated yet dependent*. For example, Monte Carlo analysis of

### 3.3 Regime shifts in excess comovement

Casual observation of Figures 1 and 2 suggests not only that comovement among U.S. metropolitan real estate markets for single-family residences has increased dramatically in the late 1990s ( $\widehat{\rho}_t^{BASE*}$  in Figure 2a), but also that *prima facie* such increase has not been accompanied by greater excess comovement in housing price changes ( $\widehat{\rho}_t^{OLS*}$  in Figures 1a and 2 and the ratio  $\widehat{\rho}_t^{OLS*}/\widehat{\rho}_t^{BASE*}$  in Figure 2b). In this section, we conduct a more rigorous analysis of the relationship between raw and residual comovement in our sample. Specifically, we test whether such relationship has experienced a regime shift over our sample period.

To that purpose, we first specify the following reduced-form model for the extent and dynamics of nationwide comovement among real estate price changes,

$$\widehat{\rho}_t^{BASE*} = a + b\widehat{\rho}_t^{OLS*} + \varepsilon_t, \quad (13)$$

as well as for each metropolitan market  $k$ ,

$$\widehat{\rho}_{kt}^{BASE*} = a_k + b_k\widehat{\rho}_{kt}^{OLS*} + \varepsilon_{kt}, \quad (14)$$

where  $k = 1, \dots, 14$ . We then test for breaks in the parameters of Eqs. (13) and (14). We do so by means of the statistical methodology of Bai et al. (1998), for it allows statistical inference about structural breaks (including the estimation of confidence interval around the break dates) with minimal restrictions on the underlying data generation process. Bai et al. (1998)'s non-parametric technique searches for a single break in univariate or multivariate time-series models (with or without stationary regressors) and generates asymptotic confidence intervals around their estimated break dates.

We start by amending the linear models of Eqs. (13) and (14) to allow for the possibility of a structural regime shift. If  $\tau$  is a potential break date,  $X_t$  is a  $1 \times 2$  vector of nationwide ( $1, \widehat{\rho}_t^{OLS*}$ ) or market-specific ( $1, \widehat{\rho}_{kt}^{OLS*}$ ) regressors, and  $\phi$  and  $\Delta\phi$  are  $2 \times 1$  vectors of nationwide ( $a, b$ ) or market-specific ( $a_k, b_k$ ) factor loadings, we specify the relation

$$y_t = X_t'\phi + d_t(\tau) X_t'S'\Delta\phi + \varepsilon_t, \quad (15)$$

where  $y_t = \widehat{\rho}_t^{BASE*}$  or  $\widehat{\rho}_{kt}^{BASE*}$ ,  $d_t(\tau) = 1$  if  $t \geq \tau$  and zero otherwise, and  $S$  is a binary selection matrix with unit diagonal elements corresponding to the coefficients in  $\phi$  which are allowed to

---

uncorrelated residuals  $e_{kt} = Z\left(\frac{x}{v}\right)^{-\frac{1}{2}}$  sharing a common shock  $x \sim \chi^2[v]$  shows that  $T\rho = 16.4712\%$  of the resulting correlations are rejected as being non-zero at the  $\alpha = 0.10$  confidence level for  $\widehat{t}_{knt}$ , i.e., too often but much less than the average number of non-zero correlations entering either the nation-wide excess comovement measure  $\widehat{\rho}_t^{OLS*}$  or each metropolitan-wide measure  $\widehat{\rho}_{kt}^{OLS*}$  (between 26% and 41%, in column  $T\rho$  of Table 2).

change. Hence, Eq. (15) is a model of full structural change if  $S$  is equal to the  $2 \times 2$  identity matrix  $I_2$ . The vector  $S\Delta\phi$  can be interpreted as the change in the corresponding subset of coefficients after a break occurred. In more compact form, the above equation is equivalent to

$$y_t = Z_t(\tau)' \Phi(\tau) + \varepsilon_t, \quad (16)$$

where  $Z_t(\tau)' = (X_t', d_t(\tau) X_t' S')$  and  $\Phi(\tau) = (\phi, S\Delta\phi)$ .

We are interested in testing the null hypothesis that  $S\Delta\phi = 0$  for each potential break date  $\tau$ . Bai et al. (1998)'s test for structural breaks to Eq. (15), based on Wald statistics (e.g., Quandt, 1958, 1960), considers the maximum of the following  $F$  process:

$$F(\tau) = T \left[ R\hat{\Phi}(\tau) \right]' \left\{ R \left[ T^{-1} \sum_{t=1}^T Z_t(\tau) \hat{\sigma}^2 Z_t(\tau)' \right]^{-1} R' \right\}^{-1} \left[ R\hat{\Phi}(\tau) \right], \quad (17)$$

where  $R = [0, S]$  is such that  $R\Phi(\tau) = S\Delta\phi$ , and  $\hat{\Phi}(\tau)$  and  $\hat{\sigma}^2$  are OLS estimators for  $\Phi(\tau)$  and  $\text{var}(\varepsilon)$ , respectively, under the alternative hypothesis of one break at date  $\tau$ . The ensuing estimated break date,  $\hat{\tau} = \arg \max F(\tau)$ , is statistically significant if  $F(\hat{\tau})$  is greater than its critical value at the chosen level of significance.<sup>12</sup>

If  $\hat{\tau}$  is statistically significant, a confidence interval around it is typically obtained by assuming that residuals in Eq. (15) are normally distributed.<sup>13</sup> However, Bai et al. (1998) propose an alternative estimator only requiring these disturbances to form a sequence of martingale differences with some moment conditions. This milder restriction is sufficient to specify the following asymptotic  $100(1 - \pi)\%$  confidence interval  $[\hat{\tau}^-, \hat{\tau}^+]$  for the true break date:

$$\hat{\tau}^\pm = \hat{\tau} \pm c_{\frac{1}{2}\pi} \left\{ (S\Delta\phi)' S \left[ (\hat{\sigma}^2 T)^{-1} \sum_{t=1}^T X_t X_t' \right] S' (S\Delta\phi) \right\}^{-1}, \quad (18)$$

where  $c_{\frac{1}{2}\pi}$  is the  $100(1 - \frac{\pi}{2})$ -th quantile of the Picard (1985) distribution.

Albeit constructed using limit theorems, the Wald statistic of Eq. (17) and its associated confidence interval around  $\hat{\tau}$  (Eq. (18)) display satisfactory finite-sample properties. In particular, according to Bai et al. (1998) and Bekaert, Harvey, and Lumsdaine (2002), these tests perform adequately, in terms of both size and power, under the null hypothesis of no break and the alternative hypothesis either of a single break in the mean of  $y_t$  or of structural breaks in all coefficients  $\phi$  for  $X_t$ , as in the general model of Eq. (15).

<sup>12</sup>To compute these critical values, Bai et al. (1998) suggest to approximate the limiting distribution of  $F(\hat{\tau})$  with partial sums of normal random variables for each possible rank of the selection matrix  $S$  in Eq. (15). Bekaert, Harvey, and Lumsdaine (2002) report a table with critical values for up to 68 parameters allowed to break.

<sup>13</sup>Bai et al. (1998) review the available literature on the topic.

The ensuing evidence in Table 3 is striking: The structural relationship between raw and excess comovement among real estate price changes experiences a statistically significant break both nationwide (Eq. (13)) and for each metropolitan market in our sample (Eq. (14)). With the exception of San Francisco, all the estimated break dates cluster in the late 1990s and in no circumstance is the corresponding estimated confidence interval wider than one month. The estimated breaks are economically significant as well: No absolute shift in  $\widehat{a}$ ,  $\widehat{a}_k$ ,  $\widehat{b}$ , or  $\widehat{b}_k$  is for less than 40% of their pre-break date levels. According to Column  $\widehat{\Delta a}$  of Table 3, in all cases the estimated regime shift is accompanied by an increase in mean correlation among the raw CSI return series  $r_{kt}$ . For instance, the average nationwide index  $\widehat{\rho}_t^{BASE*}$  is 144% higher between June 1988 and October 2006 ( $\widehat{a} + \widehat{\Delta a} = 0.548$ ) than in the previous six years ( $\widehat{a} = 0.224$ ); among the metropolitan markets in our sample, nominal price changes of individual single-family residences in San Diego, Las Vegas, Denver, and Tampa experience the greatest percentage increase in raw comovement with the rest of the U.S., consistent with Figure 3.

Excess comovement plays no role in the dynamics of raw CSI return comovement nationwide:  $\widehat{b} \approx 0$  and  $\widehat{\Delta b} \approx 0$  in Table 3 for the model for  $\widehat{\rho}_t^{BASE*}$  (Eq. (13), consistent with Column  $\beta\rho$  in Table 2). However, the estimated shift in the impact of excess comovement on metropolitan raw CSI return correlations is much more heterogeneous. Specifically, Table 3 reports that  $\widehat{\rho}_{kt}^{OLS*}$  continued to contribute to  $\widehat{\rho}_{kt}^{BASE*}$  throughout the sample only for Chicago (i.e., where both  $\widehat{b}_k > 0$  and  $\widehat{\Delta b}_k > 0$ ). Vice versa, the observed increase in CSI return correlations in Washington, Miami, and Boston cannot be attributed to  $\widehat{\rho}_{kt}^{OLS*}$ . Interestingly, excess comovement does not appear to explain the significant increase in CSI return correlations in the Western United States, i.e., the metropolitan areas where residential real estate prices increased the most over the sample period (see Table 1): Either  $\widehat{\Delta b}_k < 0$  or  $\widehat{\Delta b}_k \approx 0$  in California (Los Angeles, San Diego, San Francisco), Colorado (Denver), Nevada (Las Vegas), and Oregon (Portland). Excess unconditional correlation  $\widehat{\rho}_{kt}^{OLS*}$  becomes instead more relevant for  $\widehat{\rho}_{kt}^{BASE*}$  in the East:  $\widehat{b}_k \leq 0$  but  $\widehat{\Delta b}_k > 0$  for the comovement of CSI price changes in New York, Tampa, Charlotte, or Cleveland with the rest of the U.S. Overall, the evidence in Table 3 indicates that the observed increase in the correlation among the price changes of single-family residences traded in different U.S. metropolitan real estate markets since the late 1990s can be only partially attributed to excess comovement.

### 3.4 Further properties of CSI comovement

The evidence presented so far indicates that (i) observed raw comovement among prices of single-family homes traded in different U.S. urban areas is both statistically and economically

significant; (ii), raw comovement is primarily related to common fundamental shocks in U.S. real estate, financial, and real markets, especially to mortgage rates, realized and expected inflation, and GDP growth; (iii) comovement in excess of those systematic shocks is nonetheless of non-trivial extent; (iv) there is economically significant heterogeneity in the extent and dynamics of raw and excess comovement across U.S. metropolitan markets; and (v) the dynamics of the former are only weakly related to those of the latter. In this section, we describe and investigate further properties of these phenomena. In particular, we explore whether the time-series behavior of both  $\hat{\rho}_t^{BASE*}$  (and  $\hat{\rho}_{kt}^{BASE*}$ ) and  $\hat{\rho}_t^{OLS*}$  (and  $\hat{\rho}_{kt}^{OLS*}$ ) discussed in Section 3.1, as well as the inference on their dichotomy presented in Section 3.3, is sensitive to the state of the U.S. economy, nationwide and local trends in real estate prices, and the performance of the U.S. stock market.

In a recent paper, Pavlova and Rigobon (2007) argue that output or productivity shocks can increase price comovement, even among fundamentally unrelated traded assets, by altering consumers' relative demands for these assets, hence their relative prices. This effect may be more intense in circumstances when those shocks are more likely to occur, i.e., during either recessions or expansions. We define U.S. economic conditions using the business cycle dates provided by the National Bureau of Economic Research (NBER).<sup>14</sup> NBER expansions (recessions) begin at the trough (peak) of the cycles and end at the peak (trough). We then construct a dummy variable  $d_t^R$  equal to one if the U.S. economy is in a NBER recession in month  $t$  and zero otherwise. However, only one such event takes place over our sample, between March and October 2001. We therefore employ an additional set of discrete proxies for significant trends in the health of the U.S. economy, based on the performance of the U.S. stock market. Specifically, we define two dummy variables  $d_{SPt}^+$  ( $d_{SPt}^-$ ), equal to one if  $sign(r_{SPt}) = sign(r_{SPt-1}) = sign(r_{SPt-2}) = +$  ( $-$ ) and zero otherwise. In other words, these dummies are positive when the S&P500 index is experiencing a positive or negative run of length of (at least) three months, respectively.

Trends in raw and excess comovement in real estate prices may also be related to momentum trading. De Bondt and Thaler (1985, 1987), Jegadeesh and Titman (1993), Wermers (1999), and Sias (2004), among others, examine the extent of such activity in the U.S. stock markets. Generalized purchases or sales of single-family homes across different metropolitan markets, motivated either by herding, imitation, the activity of momentum speculators, or rational or irrational bubbles, may indeed link prices of real estate assets in markets otherwise sharing very little in common. We proxy for the extent of nationwide and local momentum of either sign in the U.S. residential housing market with two sets of dummy variables:  $d_t^+$  ( $d_t^-$ ), equal to one if  $sign(r_{mt}) = sign(r_{mt-1}) = sign(r_{mt-2}) = +$  ( $-$ ), and  $d_{kt}^+$  ( $d_{kt}^-$ ), equal to one if

---

<sup>14</sup>These dates are reported in the NBER's website, [www.nber.org/cycles.html](http://www.nber.org/cycles.html).

$sign(r_{kt}) = sign(r_{kt-1}) = sign(r_{kt-2}) = + (-)$ , i.e., when either the composite CSI index or the corresponding metropolitan index is experiencing a positive or negative run of length of (at least) three months, and zero otherwise.

We regress our measures of nationwide and local raw and excess square correlation on all of the above proxies jointly in Panel A ( $\hat{\rho}_t^{BASE*}$  and  $\hat{\rho}_{kt}^{BASE*}$ ) and Panel B ( $\hat{\rho}_t^{OLS*}$  and  $\hat{\rho}_{kt}^{OLS*}$ ) of Table 4, respectively. Regressions are estimated via OLS, but we evaluate the statistical significance of the coefficients' estimates with Newey-West standard errors to correct for heteroskedasticity and autocorrelation.<sup>15</sup> According to Papke and Wooldridge (1996), inference from these regressions may be biased since our comovement measures are between zero and one by construction. However, the estimation of either OLS regressions of *logit* transformations mapping the corresponding dependent variable to the real line — e.g.,  $\ln\left(\frac{\hat{\rho}_t^{BASE*}}{1-\hat{\rho}_t^{BASE*}}\right)$ , as in Greene, 1997, pp. 894-896 — or generalized linear models via quasi-maximum likelihood — as in Papke and Wooldridge (1996) — leads to virtually identical inference.

In general, the proxies devised above perform well in explaining the dynamics of square correlations of raw CSI returns within the U.S. over the sample period 1992-2006 (Panel A of Table 4). For instance, the adjusted  $R^2$  is greater than 41% for  $\hat{\rho}_t^{BASE*}$  and between 24% (San Francisco) and 54% (Los Angeles) for  $\hat{\rho}_{kt}^{BASE*}$ . A clear pattern also emerges from the analysis of the estimated coefficients of these regressions. Comovement among unadjusted, unconditional single-family housing price changes, either at the aggregate or local level, is greater during periods of weaker U.S. economic and financial conditions — i.e., during both NBER recessions and prolonged stock market declines. These effects are both statistically and economically significant: For example, ceteris paribus,  $\hat{\rho}_t^{BASE*}$  is 27% higher than its conditional mean ( $\frac{0.071}{0.264}$ ) when  $d_t^R = 1$  and 23% higher when  $d_{SPt}^- = 1$  ( $\frac{0.060}{0.085}$ ) at the 1% significance level. Among the metropolitan areas in our sample, raw comovement of CSI returns for Los Angeles, Las Vegas, and Boston is the most sensitive to fluctuations in the U.S. business cycle and stock market downturns. Vice versa, only raw comovement of residential real estate prices in Miami with the rest of the U.S. is (positively) related to prolonged stock market increases. Furthermore, both  $\hat{\rho}_t^{BASE*}$  and  $\hat{\rho}_{kt}^{BASE*}$  are positively related to upward momentum in either the composite ( $d_t^+ = 1$ ) or the corresponding CSI price indices ( $d_{kt}^+ = 1$ ): In those circumstances, prolonged upswings in  $r_{mt}$  can add about 72% to nationwide raw comovement ( $\frac{0.191}{0.264}$ ), and between 37% ( $\frac{0.115}{0.313}$  for Cleveland) and as much as 167% ( $\frac{0.141}{0.085}$  for Las Vegas) to raw metropolitan comovement. The estimated impact of prolonged upswings in  $r_{kt}$  on the corresponding measure  $\hat{\rho}_{kt}^{BASE*}$  is of similar absolute

---

<sup>15</sup>King et al. (1994) and Carrieri et al. (2006), among others, estimate linear regressions whose dependent variables are correlation coefficients.

and relative magnitude.

A less clear, and often more heterogeneous picture emerges from the analysis of the extent and dynamics of nationwide and metropolitan excess CSI return comovement conditional upon the state of the U.S. economy and real estate markets, in Panel B of Table 4. For example, the corresponding adjusted  $R^2$  (Column  $R_a^2$ ) are much lower than those reported in Panel A and vary considerably across the 14 urban areas in our sample. In addition, and contrary to the evidence above for raw CSI return comovement, real estate and stock market momentum play a much more limited role in explaining  $\hat{\rho}_t^{OLS*}$  and  $\hat{\rho}_{kt}^{OLS*}$ . Aggregate excess square correlation  $\hat{\rho}_t^{OLS*}$  tends to be *higher* in correspondence with prolonged declines in residential housing prices nationwide — albeit on average by less than 19% ( $\frac{0.021}{0.114}$  when  $d_t^- = 1$ ); yet, the local excess comovement measures  $\hat{\rho}_{kt}^{OLS*}$  are mostly insensitive to fluctuations in both  $r_{mt}$  and  $r_{kt}$ . Interestingly, a substitution effect between the two most important financial centers in the country appears to take place during bearish stock market performance: Excess CSI return comovement in Boston with the rest of the country increases, while  $\hat{\rho}_{kt}^{OLS*}$  for New York decreases, when  $d_{SPt}^- = 1$ . Lastly, and again contrary to what reported in Panel A of Table 4 for  $\hat{\rho}_t^{BASE*}$  and  $\hat{\rho}_{kt}^{BASE*}$ , excess comovement among single-family housing price changes *decreases* during NBER recessions — e.g., on average by 28% nationwide ( $\frac{-0.031}{0.114}$  when  $d_t^R = 1$ ) — but not during stock market swings, especially for urban areas in the West (San Diego, San Francisco, Denver, and Las Vegas) and Midwest (Cleveland).

In short, the evidence in Table 4 suggests that both raw and excess square CSI return correlation between 1992 and 2006 vary with various proxies for the state of the U.S. economy, although mostly in a dichotomous fashion, consistent with the inference in Section 3.3. We conclude that the observed comovement of prices of single-home residences in the U.S. is to a large extent, yet not exclusively, driven by underlying systematic shocks in real, financial, and real estate markets.

## 4 Conclusions

Despite the common wisdom in the popular press and beliefs shared by many economists, our empirical analysis fails to support the notion that U.S. residential real estate has gone through a bubble over the last two decades. Using housing price index data for 14 among the largest U.S. metropolitan areas, we find that observed price comovement among these markets increased between 1987 and 2006, but that this increase stems more from shocks to fundamental factors driving real estate returns than from an irrational response by investors. In essence, we find that domestic residential real estate markets within the U.S. have become more *fundamentally*

integrated, in a manner parallel to the increasing convergence of international financial markets (e.g., Bekaert, Harvey, and Lumsdaine, 2002; Bekaert, Harvey, and Ng, 2005; Bekaert et al., 2005).

Our empirical conclusions arise from a distinct approach to analyzing these data. It relies on the specification of a comprehensive linear model for describing residential real estate returns. We find that this model performs well over the time frame of our study; furthermore, its performance in fact improves over the latter part of our data. We then use this model to decompose the observed comovement among these 14 indices into fundamental covariation (arising from our return generating process) and excess covariation (the observed comovement net of this fundamental covariation). If markets were becoming "less rational," we would expect an increasing amount of the total covariation to be due to such excess covariation over our sample period. We do not find this; comovement among these regional real estate markets is indeed increasing over our sample period, but the explanatory power of the common factors driving real estate prices is also increasing. Thus, excess comovement explains a *declining* portion of the observed comovement in the latter part of the sample.

We also analyze the possible determinants of the extent and dynamics of observed and excess comovement among U.S. metropolitan real estate markets. First, a structural break analysis indicates that observed comovement among price changes of single-home residences in all but one of the 14 metropolitan markets in our sample experiences a (statistically and economically) significant upward regime shift in 1999, well before analysts were decrying a bubble in these markets; yet none of these breaks can be attributed to an increase in excess comovement. Second, we show that both observed and excessive comovement are sensitive to systematic real and financial shocks — e.g., NBER recessions or prolonged stock and real estate market declines — albeit often heterogeneously so in sign and magnitude.

In summary, the evidence reported in this study contributes to, and possibly tempers the debate about the U.S. residential real estate bubble. It shows that a parsimonious fundamental model of residential real estate returns can explain a great deal of the comovement of prices of single-home residences among different U.S. metropolitan areas over the past two decades and, perhaps more strikingly, that it in fact performs better — and excess covariation matters less — over the latter part of our sample, the period most commonly associated with the bubble.

## References

- Bai, J., Lumsdaine, R., Stock, J. (1998). Testing for and Dating Common Breaks in Multivariate Time Series. *Review of Economic Studies*, 65, pp. 395-432.
- Barberis, N., Shleifer, A., Wurgler, J. (2005). Comovement. *Journal of Financial Economics*, 75, pp. 283-317.
- Bekaert, G., Harvey, C., Lumsdaine, R. (2002). Dating the Integration of World Equity Markets. *Journal of Financial Economics*, 65, pp. 203-247.
- Bekaert, G, Harvey, C., Ng, A. (2005). Market Integration and Contagion. *Journal of Business*, 78, pp. 39-69.
- Bekaert, G., Hodrick, R. , Zhang, X. (2005). International Stock Return Comovements. Working Paper, Columbia University.
- Boyer, B., Gibson, M., Loretan, M. (1999). Pitfalls in Tests for Changes in Correlations. Working Paper No. 597R, Federal Reserve Board.
- Breusch, T., Pagan, A. (1980). The LM Test and Its Application to Model Specification in Econometrics. *Review of Economic Studies*, 47, pp. 239-254.
- Campbell, J., Lo, A., MacKinley, C. (1997). *The Econometrics of Financial Markets*. Princeton University Press, New Jersey.
- Campbell, J., Lettau, M., Malkiel, B., Xu, Y. (2001). Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk. *Journal of Finance*, 56, pp. 1-43.
- Carrieri, F., Errunza, V., Sarkissian, S. (2006). Economic Integration, Industrial Structure, and International Portfolio Diversification, Working Paper, McGill University.
- Corsetti, G., Pericoli, M., Sbracia, M. (2005). Some Contagion, Some Interdependence: More Pitfalls in Tests of Financial Contagion. *Journal of International Money and Finance*, 8, pp. 1177-1199.
- De Bondt, W., Thaler, R. (1985). Does the Stock Market Overreact? *Journal of Finance*, 40, pp. 793-805.
- De Bondt, W., Thaler, R. (1987). Further Evidence on Investor Overreaction and Stock Market Seasonality. *Journal of Finance*, 42, pp. 557-581.

- Fleming, J., Kirby, C., Ostdiek, B. (1998). Information and Volatility Linkages in the Stock, Bond, and Money Markets. *Journal of Financial Economics*, 49, pp. 111-137.
- Forbes, K., Rigobon, R. (2001). Measuring Contagion: Conceptual and Empirical Issues. In S. Claessens and K. Forbes (eds.): *International Financial Contagion*. Kluwert Academic Publishers, Norwell.
- Forbes, K., Rigobon, R. (2002). No Contagion, Only Interdependence: Measuring Stock Market Co-Movements. *Journal of Finance*, 57, pp. 2223-2261.
- Greene, W. (1997). *Econometric Analysis*. Prentice Hall, New Jersey.
- Jegadeesh, N., Titman, S. (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *Journal of Finance*, 48, pp. 65-91.
- Kallberg, J., Liu, C., Pasquariello, P. (2005). An Examination of the Asian Crisis: Regime Shifts in Currency and Equity Markets. *Journal of Business*, 78, pp. 169-211.
- Kallberg, J., Pasquariello, P. (2007). Time Series and Cross-Sectional Excess Comovement in Stock Indexes. *Journal of Empirical Finance*, forthcoming.
- Karolyi, A., Stulz, R. (1996). Why Do Markets Move Together? An Investigation of U.S.-Japan Stock Return Co-movements. *Journal of Finance*, 51, pp. 951-986.
- King, M., Wadhvani, S. (1990). Transmission of Volatility between Stock Markets. *Review of Financial Studies*, 3, pp. 5-33.
- King, M., Sentana, E., Wadhvani, S. (1994). Volatility and Links between National Stock Markets. *Econometrica*, 62, pp. 901-934.
- Kodres, L., Pritsker, M. (2002). A Rational Expectations Model of Financial Contagion. *Journal of Finance*, 57, pp. 769-799.
- McElroy, M. (1977). Goodness of Fit for Seemingly Unrelated Regressions: Glahn's  $R_{y,x}^2$  and Hooper's  $\bar{r}^2$ . *Journal of Econometrics*, 6, pp. 381-387.
- Loretan M., English, W. (2000). Evaluating "Correlation Breakdowns" during Periods of Market Volatility. In Bank of International Settlements (BIS), *International Financial Markets and the Implications for Monetary and Financial Stability*, BIS, Switzerland.

- Nelson, C. (1992). Filtering and Forecasting with Misspecified ARCH Models I: Getting the Right Variance with the Wrong Model. *Journal of Econometrics*, 52, pp. 61-90.
- Officer, R. (1973). The Variability of the Market Factor of the New York Stock Exchange. *Journal of Business*, 46, pp. 434-453.
- Papke, L., Wooldridge, J. (1996). Econometric Methods for Fractional Response Variables with an Application to 401(K) Plan Participation rates. *Journal of Applied Econometrics*, 11, pp. 619-632.
- Pasquariello, P. (2007). Imperfect Competition, Information Heterogeneity, and Financial Contagion. *Review of Financial Studies*, 20, pp. 391-426.
- Pavlova, A., Rigobon, R. (2007). Asset Prices and Exchange Rates. *Review of Financial Studies*, 20, pp. 1139-1181.
- Picard D. (1985). Testing and Estimating Change-Points in Time Series. *Advances in Applied Probability*, 176, pp. 841-867.
- Pindyck, R., Rotemberg, J. (1990). The Excess Comovement of Commodity Prices. *Economic Journal*, 100, pp. 1173-1189.
- Pindyck, R., Rotemberg, J. (1993). The Comovement of Stock Prices. *Quarterly Journal of Economics*, 108, pp. 1073-1104.
- Quandt, R. (1958). The Estimation of the Parameters of a Linear Regression System Obeying Two Separate Regimes. *Journal of the American Statistical Association*, 53, pp. 873-880.
- Quandt, R. (1960). Tests of the Hypothesis That a Linear Regression System Obeys Two Separate Regimes. *Journal of the American Statistical Association*, 55, pp. 324-330.
- Sias, R. (2004). Institutional Herding. *Review of Financial Studies*, 17, pp. 165-206.
- Veldkamp, L. (2006). Information Markets and the Comovement of Asset Prices. *Review of Economic Studies*, 73, pp. 823-845.
- Wermers, R. (1999). Mutual Fund Herding and the Impact on Stock Prices. *Journal of Finance*, 54, pp. 581-622.
- Zhou, G. (1993). Asset-pricing Tests under Alternative Distributions. *Journal of Finance*, 48, pp. 1927-1942.

Table 1. Descriptive statistics

This table reports summary statistics for the time series of monthly returns for the composite home price index ( $r_{mt}$ ) and the 14 individual metropolitan market indices ( $r_{kt}$ ), defined in Section 4, over the interval February, 1987 - October, 2006 (237 observations).  $N$  is the number of sub-category constituents for each index.  $\mu$  is the mean and  $\sigma$  is the standard deviation of each series.  $Skew$  is the coefficient of skewness, while  $Kurt$  is the excess kurtosis; their standard errors, in their asymptotic normal distributions, are computed as  $(\frac{6}{T})^{\frac{1}{2}}$  and  $(\frac{24}{T})^{\frac{1}{2}}$ , respectively.  $\widehat{\rho}_1$  is the first-order autocorrelation.  $LB(5)$  is the Ljung-Box test of randomness for up to the fifth-order autocorrelation, asymptotically distributed as  $\chi^2 [5]$  under the null hypothesis that the series is white noise. A “ \* ”, “ • ”, or “ † ” indicate significance at the 10%, 5%, or 1% level, respectively.

Index	N	$\mu$	$\sigma$	Skew	Kurt	$\widehat{\rho}_1$	$LB(5)$
Composite	237	0.537% <sup>†</sup>	0.66%	0.11	-0.40	1.230 <sup>†</sup>	427.91 <sup>†</sup>
Metropolitan Markets							
Los Angeles	237	0.645% <sup>†</sup>	1.03%	0.25	-0.08	0.961 <sup>†</sup>	279.04 <sup>†</sup>
San Diego	237	0.631% <sup>†</sup>	0.97%	0.73 <sup>†</sup>	1.82 <sup>†</sup>	0.594 <sup>†</sup>	120.29 <sup>†</sup>
San Francisco	237	0.646% <sup>†</sup>	1.03%	0.54 <sup>†</sup>	0.38	1.139 <sup>†</sup>	347.60 <sup>†</sup>
Denver	237	0.429% <sup>†</sup>	0.54%	0.01	0.33	0.689 <sup>†</sup>	205.65 <sup>†</sup>
Washington	237	0.563% <sup>†</sup>	0.80%	0.53 <sup>†</sup>	0.35	0.898 <sup>†</sup>	219.22 <sup>†</sup>
Miami	237	0.592% <sup>†</sup>	0.72%	0.61 <sup>†</sup>	1.30 <sup>†</sup>	0.469 <sup>†</sup>	139.34 <sup>†</sup>
Tampa	237	0.467% <sup>†</sup>	0.70%	0.88 <sup>†</sup>	1.29 <sup>†</sup>	0.501 <sup>†</sup>	149.40 <sup>†</sup>
Chicago	237	0.484% <sup>†</sup>	0.62%	0.34 <sup>•</sup>	1.55 <sup>†</sup>	0.556 <sup>†</sup>	129.29 <sup>†</sup>
Boston	237	0.388% <sup>†</sup>	0.79%	-0.07	0.05	0.874 <sup>†</sup>	207.52 <sup>†</sup>
Charlotte	237	0.298% <sup>†</sup>	0.43%	0.27 <sup>*</sup>	1.13 <sup>†</sup>	0.475 <sup>†</sup>	132.81 <sup>†</sup>
Las Vegas	237	0.531% <sup>†</sup>	1.00%	2.15 <sup>†</sup>	9.57 <sup>†</sup>	0.654 <sup>†</sup>	120.67 <sup>†</sup>
New York	237	0.444% <sup>†</sup>	0.69%	0.06	-0.40	1.057 <sup>†</sup>	347.22 <sup>†</sup>
Cleveland	237	0.348% <sup>†</sup>	0.49%	0.18	0.03	0.485 <sup>†</sup>	105.03 <sup>†</sup>
Portland	237	0.627% <sup>†</sup>	0.62%	0.74 <sup>†</sup>	1.33 <sup>†</sup>	0.593 <sup>†</sup>	157.89 <sup>†</sup>

Table 2. Descriptive statistics: BASE and OLS comovement

This table reports summary statistics for the monthly time series of excess square correlation of OLS residuals ( $\widehat{\rho}_t^{OLS*}$  in Eq. (6)), and of benchmark square correlation of gross returns ( $\widehat{\rho}_t^{BASE*}$ ), for the nationwide index ( $\widehat{\rho}_t^{OLS*}$  and  $\widehat{\rho}_t^{BASE*}$ ) and the 14 individual metropolitan market indices ( $\widehat{\rho}_{kt}^{OLS*}$  and  $\widehat{\rho}_{kt}^{BASE*}$ ), defined in Section 4, over the interval January, 1992 - October, 2006 (178 observations). OLS residuals are estimated according to the procedure described in Section 3 from the following specification of the fundamental model of Eq. (2):

$$r_{kt} = \alpha_{kt} + \beta_{mt}^k r_{mt} + \beta_{SPt}^k r_{SPt} + \beta_{CPIt}^k CPI_t + \beta_{MTGt}^k MTG_t + \beta_{SLPt}^k SLP_t + \beta_{UNE_t}^k UNE_t + \beta_{POP_t}^k POP_t + \beta_{INC_t}^k INC_t + \beta_{GDP_t}^k GDP_t + e_{kt},$$

where  $r_{mt}$  is the monthly return for the composite home price index,  $r_{SPt}$  is the monthly return for the S&P500 index,  $CPI_t$  is the monthly percentage change in the CPI index (all items excluding shelter),  $MTG_t$  is the 30-year conventional mortgage rate,  $SLP_t$  is the monthly difference between 10-year and 3-month U.S. Treasury constant-maturity rates,  $UNE_t$  is the monthly change in the civilian unemployment rate,  $POP_t$  is the monthly percentage change in total U.S. population,  $INC_t$  is the monthly percentage change in disposable personal income, and  $GDP_t$  is the interpolated monthly percentage change in nominal GDP, for each  $k = 1, \dots, K$ .  $B\rho$  is the correlation between each pair  $\widehat{\rho}_t^{OLS*}$  and  $\widehat{\rho}_t^{BASE*}$ , while  $R\rho$  is the mean ratio between each pair  $\widehat{\rho}_t^{OLS*}$  and  $\widehat{\rho}_t^{BASE*}$ .  $T\rho$  is the mean percentage of the corresponding conditional correlations  $\widehat{\rho}_{knt}$  significant at the 10% level using the  $t$ -ratio test  $\widehat{t}_{knt} = \widehat{\rho}_{knt} \left[ \frac{1 - (\widehat{\rho}_{knt})^2}{N-2} \right]^{\frac{1}{2}} \sim t[N-2]$ , for  $N = 100$ , i.e., either the mean of the corresponding metropolitan-wide ratios  $\frac{1}{K-1} \sum_{\substack{n=1 \\ n \neq k}}^K I_{knt}$ , where, as in Section 2.4,  $I_{knt} = 1$  if  $2[1 - \Pr\{|\widehat{t}_{knt}| \leq t_{\frac{\alpha}{2}}[N-2]\}] \leq \alpha$ , with  $\alpha = 0.10$ , and zero otherwise, or the mean of its market-wide averages.  $F\rho^{*2}$  is the mean percentage of the corresponding square unconditional correlations  $(\widehat{\rho}_{knt}^*)^2$  significant at the 10% level using the  $t$ -square ratio test  $\widehat{t}_{knt}^2 = (\widehat{\rho}_{knt}^*)^2 [1 - (\widehat{\rho}_{knt}^*)^2]^{-1} (N-2) \sim F[1, N-2]$ , i.e., either the mean of the corresponding metropolitan-wide ratios  $\frac{1}{K-1} \sum_{\substack{n=1 \\ n \neq k}}^K I_{knt}^2$ , where  $I_{knt}^2 = 1$  if  $[1 - \Pr\{\widehat{t}_{knt}^2 \leq F_{\alpha}[1, N-2]\}] \leq \alpha$ , with  $\alpha = 0.10$ , and zero otherwise, or the mean of its market-wide averages. A “\*”, “•”, or “†” indicate that the  $t$ -square statistic  $\widehat{\rho}_t^* [1 - \widehat{\rho}_t^*]^{-1} (N-2)$  for the test of the null hypothesis  $H_0 : \widehat{\rho}_t^* = 0$  against the alternative hypothesis  $H_1 : \widehat{\rho}_t^* > 0$  is significant at the significance at the 10%, 5%, or 1% level, respectively.

Table 2 (Continued).

Index	$\widehat{\rho}_t^{OLS*}$			$\widehat{\rho}_t^{BASE*}$						
	$\mu$	$\sigma$	$F\rho^{*2}$	$\mu$	$\sigma$	$F\rho^{*2}$				
Nationwide	0.112 <sup>†</sup>	0.027	-0.238	28%	33%	43%	0.396 <sup>†</sup>	0.159	81%	83%
Metropolitan Markets										
Los Angeles	0.133 <sup>†</sup>	0.073	0.065	29%	40%	50%	0.455 <sup>†</sup>	0.206	79%	81%
San Diego	0.104 <sup>†</sup>	0.063	-0.049	24%	35%	44%	0.430 <sup>†</sup>	0.205	78%	81%
San Francisco	0.142 <sup>†</sup>	0.074	0.225	35%	39%	49%	0.401 <sup>†</sup>	0.186	82%	86%
Denver	0.099 <sup>†</sup>	0.070	0.038	21%	31%	43%	0.465 <sup>†</sup>	0.209	86%	85%
Washington	0.113 <sup>†</sup>	0.075	-0.114	26%	34%	43%	0.441 <sup>†</sup>	0.187	82%	86%
Miami	0.116 <sup>†</sup>	0.101	-0.032	32%	31%	43%	0.361 <sup>†</sup>	0.187	76%	78%
Tampa	0.096 <sup>†</sup>	0.092	-0.189	30%	26%	31%	0.327 <sup>†</sup>	0.200	72%	72%
Chicago	0.098 <sup>†</sup>	0.070	-0.020	22%	29%	43%	0.450 <sup>†</sup>	0.183	84%	86%
Boston	0.096 <sup>†</sup>	0.090	-0.002	20%	27%	35%	0.470 <sup>†</sup>	0.182	89%	89%
Charlotte	0.086 <sup>†</sup>	0.039	-0.011	31%	34%	42%	0.277 <sup>†</sup>	0.134	81%	81%
Las Vegas	0.128 <sup>†</sup>	0.091	0.105	47%	32%	42%	0.275 <sup>†</sup>	0.182	71%	75%
New York	0.131 <sup>†</sup>	0.045	-0.074	28%	41%	52%	0.462 <sup>†</sup>	0.165	89%	90%
Cleveland	0.102 <sup>†</sup>	0.070	0.024	27%	32%	41%	0.374 <sup>†</sup>	0.116	89%	91%
Portland	0.131 <sup>†</sup>	0.144	0.035	36%	31%	44%	0.363 <sup>†</sup>	0.157	82%	83%

Table 3. Break tests

This table reports OLS estimates (and their  $t$  statistics below) for all coefficients in Eq. (15), when the latter experiences a statistically significant break (at the 5% level) in month  $\hat{\tau}$  for either the nationwide or each metropolitan market-wide comovement index ( $\hat{\rho}_t^{BASE*}$  and  $\hat{\rho}_{kt}^{BASE*}$ ) over the interval January, 1992 - October, 2006 (178 observations).  $F(\hat{\tau})$  is the corresponding Wald statistic (Eq. (17)), while  $\hat{\tau}^-$  and  $\hat{\tau}^+$  are the lower and upper bounds of the 95% confidence interval around  $\hat{\tau}$  (Eq. (18)).  $R_a^2$  is the adjusted  $R^2$ . A “ \* ”, “ • ”, or “ † ” indicate significance at the 10%, 5%, or 1% level, respectively.

Index	$F(\hat{\tau})$	$\hat{\tau}^-$	$\hat{\tau}$	$\hat{\tau}^+$	$R_a^2$	$\hat{a}$	$\hat{b}$	$\widehat{\Delta a}$	$\widehat{\Delta b}$
Nationwide	1337.32 <sup>†</sup>	Sep-98	Sep-98	Sep-98	88.95%	0.224 <sup>†</sup> 8.32	0.050 0.23	0.324 <sup>†</sup> 9.07	-0.204 -0.67
Metropolitan Markets									
Los Angeles	837.43 <sup>†</sup>	May-98	Jun-98	Jun-98	82.57%	0.196 <sup>†</sup> 11.61	0.336 <sup>†</sup> 3.41	0.335 <sup>†</sup> 10.45	0.334 1.48
San Diego	1769.35 <sup>†</sup>	Sep-98	Sep-98	Sep-98	90.91%	0.074 <sup>†</sup> 5.31	1.233 <sup>†</sup> 12.33	0.550 <sup>†</sup> 29.38	-1.514 <sup>†</sup> -9.72
San Francisco	167.90 <sup>†</sup>	Dec-03	Jan-04	Jan-04	50.85%	0.305 <sup>†</sup> 13.39	0.241 1.59	0.478 <sup>†</sup> 6.97	-0.854 <sup>•</sup> -2.34
Denver	742.00 <sup>†</sup>	Jun-97	Jul-97	Jul-97	80.70%	0.161 <sup>†</sup> 8.73	0.602 <sup>†</sup> 4.35	0.453 <sup>†</sup> 18.49	-0.672 <sup>†</sup> -3.39
Washington	1152.89 <sup>†</sup>	Jun-99	Jun-99	Jun-99	86.84%	0.273 <sup>†</sup> 19.46	-0.045 -0.45	0.345 <sup>†</sup> 18.63	0.018 0.13
Miami	459.04 <sup>†</sup>	Aug-98	Sep-98	Sep-98	72.07%	0.171 <sup>†</sup> 10.69	0.120 1.38	0.294 <sup>†</sup> 12.28	0.250 1.48
Tampa	792.11 <sup>†</sup>	May-99	Jun-99	Jun-99	82.33%	0.167 <sup>†</sup> 12.89	-0.190 <sup>•</sup> -2.25	0.331 <sup>†</sup> 17.72	0.302 <sup>•</sup> 2.01
Chicago	557.08 <sup>†</sup>	Apr-98	May-98	May-98	75.80%	0.244 <sup>†</sup> 13.63	0.223 <sup>*</sup> 1.74	0.280 <sup>†</sup> 11.62	0.486 <sup>•</sup> 2.40
Boston	876.94 <sup>†</sup>	May-98	Jun-98	Jun-98	83.16%	0.290 <sup>†</sup> 20.43	-0.110 -0.90	0.328 <sup>†</sup> 18.67	0.078 0.55
Charlotte	290.97 <sup>†</sup>	Jul-98	Aug-98	Aug-98	61.94%	0.160 <sup>†</sup> 8.29	-0.001 -0.01	0.147 <sup>†</sup> 4.47	0.768 <sup>•</sup> 2.15
Las Vegas	843.88 <sup>†</sup>	Apr-00	May-00	May-00	82.80%	0.066 <sup>†</sup> 5.53	0.499 <sup>†</sup> 7.34	0.423 <sup>†</sup> 17.68	-0.783 <sup>†</sup> -4.51
New York	957.50 <sup>†</sup>	Feb-99	Mar-99	Mar-99	84.44%	0.407 <sup>†</sup> 16.33	-0.752 <sup>†</sup> -4.20	0.166 <sup>†</sup> 5.32	1.016 <sup>†</sup> 4.52
Cleveland	447.25 <sup>†</sup>	Apr-98	May-98	May-98	71.53%	0.294 <sup>†</sup> 24.08	-0.344 <sup>†</sup> -3.27	0.152 <sup>†</sup> 9.24	0.444 <sup>†</sup> 3.26
Portland	559.65 <sup>†</sup>	May-99	Jun-99	Jun-99	75.90%	0.199 <sup>†</sup> 18.93	0.189 <sup>†</sup> 4.33	0.284 <sup>†</sup> 15.23	-0.065 -0.51

Table 4. Further properties of BASE and OLS comovement

This table reports OLS estimates (and their  $t$  statistics below) for the coefficients of the regressions of either  $\hat{\rho}_t^{BASE*}$  and each  $\hat{\rho}_{kt}^{BASE*}$  (Panel A) or  $\hat{\rho}_t^{OLS*}$  and each  $\hat{\rho}_{kt}^{OLS*}$  (Panel B) on the following explanatory variables, over the interval January, 1992 - October, 2006 (178 observations):  $d_t^R$ , a dummy variable equal to one if the U.S. economy was in a NBER recession in month  $t$  and zero otherwise;  $d_t^+$  ( $d_t^-$ ), a dummy variable equal to one if  $sign(r_{mt}) = sign(r_{mt-1}) = sign(r_{mt-2}) = +$  ( $-$ ) and zero otherwise;  $d_{SPt}^+$  ( $d_{SPt}^-$ ), a dummy variable equal to one if  $sign(r_{SPt}) = sign(r_{SPt-1}) = sign(r_{SPt-2}) = +$  ( $-$ ) and zero otherwise;  $d_{kt}^+$  ( $d_{kt}^-$ ), a dummy variable equal to one if  $sign(r_{kt}) = sign(r_{kt-1}) = sign(r_{kt-2}) = +$  ( $-$ ) and zero otherwise.  $R_a^2$  is the adjusted  $R^2$ . Statistical significance is evaluated using Newey-West standard errors. A “\*”, “•”, or “†” indicate significance at the 10%, 5%, or 1% level, respectively.

Index	Panel A: $\hat{\rho}_t^{BASE*}$ and $\hat{\rho}_{kt}^{BASE*}$								$R_a^2$	
	Constant	$d_t^R$	$d_t^+$	$d_t^-$	$d_{SPt}^+$	$d_{SPt}^-$	$d_{kt}^+$	$d_{kt}^-$		
Nationwide	0.264 <sup>†</sup> 8.78	0.071 <sup>†</sup> 3.48	0.191 <sup>†</sup> 5.69	-0.042 -1.36	0.019 0.63	0.060 <sup>†</sup> 2.83				41.39%
Metropolitan Markets										
Los Angeles	0.334 <sup>†</sup> 6.61	0.148 <sup>†</sup> 7.38	-0.023 -0.38	0.054 0.96	-0.021 -0.62	0.062 <sup>•</sup> 2.11	0.251 <sup>†</sup> 3.91	-0.134 <sup>*</sup> -1.93		54.37%
San Diego	0.247 <sup>†</sup> 5.00	0.096 <sup>†</sup> 3.85	0.124 <sup>•</sup> 2.46	-0.037 -0.80	0.000 0.00	0.071 <sup>•</sup> 2.55	0.157 <sup>†</sup> 3.06	0.038 0.78		42.04%
San Francisco	0.262 <sup>†</sup> 7.63	-0.021 -0.44	0.158 <sup>†</sup> 3.81	-0.003 -0.07	0.010 0.25	-0.072 -1.36	0.061 1.42	0.002 0.06		23.67%
Denver	0.348 <sup>†</sup> 7.69	0.056 <sup>•</sup> 2.29	0.268 <sup>†</sup> 6.67	-0.083 <sup>*</sup> -1.87	0.037 1.18	0.029 0.88	-0.086 <sup>†</sup> -2.97	0.092 1.44		51.87%
Washington	0.230 <sup>†</sup> 7.23	0.028 1.37	0.111 <sup>†</sup> 3.10	0.014 0.40	0.017 0.57	0.031 1.25	0.210 <sup>†</sup> 6.14	0.034 0.89		50.23%
Miami	0.182 <sup>†</sup> 3.73	0.016 0.69	0.180 <sup>†</sup> 3.59	0.007 0.15	0.073 <sup>*</sup> 1.89	0.044 <sup>•</sup> 2.07	0.054 1.08	-0.072 -1.53		26.85%
Tampa	0.140 <sup>†</sup> 3.41	0.030 1.38	0.131 <sup>†</sup> 2.93	-0.014 -0.34	-0.009 -0.23	0.062 <sup>•</sup> 2.55	0.152 <sup>†</sup> 3.51	-0.003 -0.05		40.48%
Chicago	0.289 <sup>†</sup> 6.91	0.047 <sup>•</sup> 2.14	0.208 <sup>†</sup> 4.69	-0.063 -1.35	0.050 1.49	0.058 <sup>•</sup> 1.99	0.020 0.51	0.046 0.97		39.56%
Boston	0.330 <sup>†</sup> 9.28	0.121 <sup>†</sup> 4.50	0.263 <sup>†</sup> 6.90	-0.094 <sup>•</sup> -2.27	0.013 0.46	0.090 <sup>†</sup> 3.02	-0.072 <sup>•</sup> -2.43	0.046 1.30		47.14%
Charlotte	0.212 <sup>†</sup> 6.64	-0.014 -0.55	0.116 <sup>†</sup> 3.66	-0.046 -1.48	-0.013 -0.51	0.064 1.49	-0.006 -0.24	-0.042 -1.49		24.99%
Las Vegas	0.085 <sup>•</sup> 2.38	0.126 <sup>†</sup> 4.53	0.141 <sup>†</sup> 3.36	-0.009 -0.20	0.054 1.65	0.086 <sup>†</sup> 2.78	0.133 <sup>†</sup> 4.04	0.088 1.46		40.69%
New York	0.345 <sup>†</sup> 8.81	0.084 <sup>†</sup> 3.75	0.181 <sup>†</sup> 3.49	-0.052 -1.31	0.014 0.44	0.067 <sup>•</sup> 2.58	-0.011 -0.20	0.014 0.28		32.12%
Cleveland	0.313 <sup>†</sup> 13.00	0.095 <sup>†</sup> 5.52	0.115 <sup>†</sup> 4.75	-0.051 <sup>•</sup> -2.06	-0.014 -0.76	0.070 <sup>†</sup> 3.08	-0.025 -1.31	0.026 1.20		41.25%
Portland	0.219 <sup>†</sup> 4.46	0.104 <sup>†</sup> 4.52	0.148 <sup>†</sup> 4.33	-0.009 -0.24	-0.007 -0.24	0.056 <sup>*</sup> 1.69	0.052 1.32	0.200 <sup>†</sup> 3.80		25.81%

Table 4 (*Continued*).

Index	Panel B: $\hat{\rho}_t^{OLS*}$ and $\hat{\rho}_{kt}^{OLS*}$								$R_a^2$	
	Constant	$d_t^R$	$d_t^+$	$d_t^-$	$d_{SPt}^+$	$d_{SPt}^-$	$d_{kt}^+$	$d_{kt}^-$		
Nationwide	0.114 <sup>†</sup> 19.83	-0.031 <sup>†</sup> -2.79	-0.007 -1.17	0.021 <sup>•</sup> 2.46	0.007 1.49	0.002 0.20				22.57%
Metropolitan Markets										
Los Angeles	0.162 <sup>†</sup> 5.66	-0.011 <sup>†</sup> -0.78	-0.060 -2.72	0.062 1.38	-0.016 -0.97	0.000 -0.04	0.026 <sup>*</sup> 1.96	-0.050 -1.26		7.89%
San Diego	0.110 <sup>†</sup> 5.32	-0.055 <sup>†</sup> -4.90	0.005 0.28	0.010 0.37	0.008 0.58	-0.040 <sup>†</sup> -3.25	-0.018 -1.11	0.016 0.57		9.53%
San Francisco	0.098 <sup>†</sup> 7.70	-0.058 <sup>†</sup> -4.68	0.054 <sup>†</sup> 3.71	0.019 0.84	0.051 <sup>†</sup> 2.99	-0.002 -0.14	-0.017 -1.20	0.020 1.48		10.53%
Denver	0.104 <sup>†</sup> 6.21	-0.058 <sup>†</sup> -7.28	0.015 1.27	0.018 0.82	0.019 1.57	-0.005 -0.40	-0.024 -1.56	-0.017 -1.01		5.00%
Washington	0.101 <sup>†</sup> 5.56	-0.040 <sup>•</sup> -2.37	-0.009 -0.50	-0.011 -0.51	0.019 1.13	0.020 0.50	0.014 0.80	0.057 <sup>†</sup> 2.79		5.19%
Miami	0.155 <sup>†</sup> 5.22	-0.059 <sup>†</sup> -7.10	0.011 0.49	0.001 0.01	0.000 0.02	-0.005 -0.29	-0.066 <sup>•</sup> -2.51	0.130 0.98		14.29%
Tampa	0.085 <sup>†</sup> 5.30	-0.034 <sup>†</sup> -3.19	0.009 0.61	0.050 1.34	0.033 1.45	-0.032 <sup>•</sup> -2.31	-0.014 -1.08	0.041 0.75		8.37%
Chicago	0.102 <sup>†</sup> 6.84	-0.032 <sup>*</sup> -1.93	0.006 0.32	0.007 0.39	0.011 0.77	-0.018 -1.12	-0.013 -0.71	-0.058 <sup>†</sup> -2.83		-0.56%
Boston	0.094 <sup>†</sup> 12.45	0.035 0.54	0.007 0.36	0.031 1.09	-0.013 -1.07	0.120 <sup>†</sup> 3.28	-0.013 -0.58	-0.035 -1.50		6.33%
Charlotte	0.102 <sup>†</sup> 10.19	-0.016 <sup>*</sup> -1.97	-0.022 <sup>•</sup> -2.40	0.006 0.40	-0.006 -0.82	0.023 <sup>†</sup> 2.96	0.001 0.21	-0.017 -0.93		7.38%
Las Vegas	0.119 <sup>†</sup> 6.32	-0.048 <sup>†</sup> -2.91	0.017 0.72	0.018 0.53	0.016 0.81	-0.024 -1.43	-0.011 -0.53	-0.008 -0.16		-0.65%
New York	0.145 <sup>†</sup> 10.28	-0.028 <sup>†</sup> -4.34	0.016 0.88	-0.026 <sup>*</sup> -1.84	-0.003 -0.32	-0.015 <sup>*</sup> -1.83	-0.032 -1.62	0.049 <sup>•</sup> 2.59		9.28%
Cleveland	0.101 <sup>†</sup> 6.47	-0.046 <sup>†</sup> -4.64	0.014 1.22	0.037 <sup>*</sup> 1.89	-0.008 -0.60	-0.022 -1.59	-0.014 -0.97	-0.034 <sup>•</sup> -2.06		3.18%
Portland	0.166 <sup>†</sup> 3.07	0.002 0.05	-0.070 -1.48	0.024 0.42	-0.031 -0.95	0.027 0.68	0.019 0.91	-0.012 -0.36		3.79%

Figure 1. Systematic shocks to the U.S. economy

Figures 1a to 1i plot (on the left axis) the time series of  $r_{mt}$  (the monthly return for the composite home price index),  $r_{SPt}$  (the monthly return for the S&P500 index),  $CPI_t$  (the monthly percentage change in the CPI index, all items excluding shelter),  $MTG_t$  (the 30-year conventional mortgage rate),  $SLP_t$  (the monthly difference between 10-year and 3-month U.S. Treasury constant-maturity rates),  $UNE_t$  (the monthly change in the civilian unemployment rate),  $POP_t$  (the monthly percentage change in total U.S. population),  $INC_t$  (the monthly percentage change in disposable personal income), and  $GDP_t$  (the interpolated monthly percentage change in GDP). These figures also display (on the right axis) both  $\bar{r}_{fit}^2$ , the average of the square partial correlations  $r_{r_{kt},f_{it}}^2 = \frac{t_{r_{kt},f_{it}}^2}{t_{r_{kt},f_{it}}^2 + N - K}$ , where  $f_{it}$  is the corresponding systematic shock and  $t_{r_{kt},f_{it}}^2$  is the square of the  $t$ -ratio for testing the hypothesis that the coefficient on  $f_{it}$  is zero in Eq. (2), and  $\bar{R}_{fit}^2$ , the average of the  $R^2$  of the regression of *exclusively* factor  $f_{it}$  (and a constant term) on each CSI return series  $r_{kt}$  separately over the interval  $[t - N + 1, t]$ , for  $k = 1, \dots, K$ .

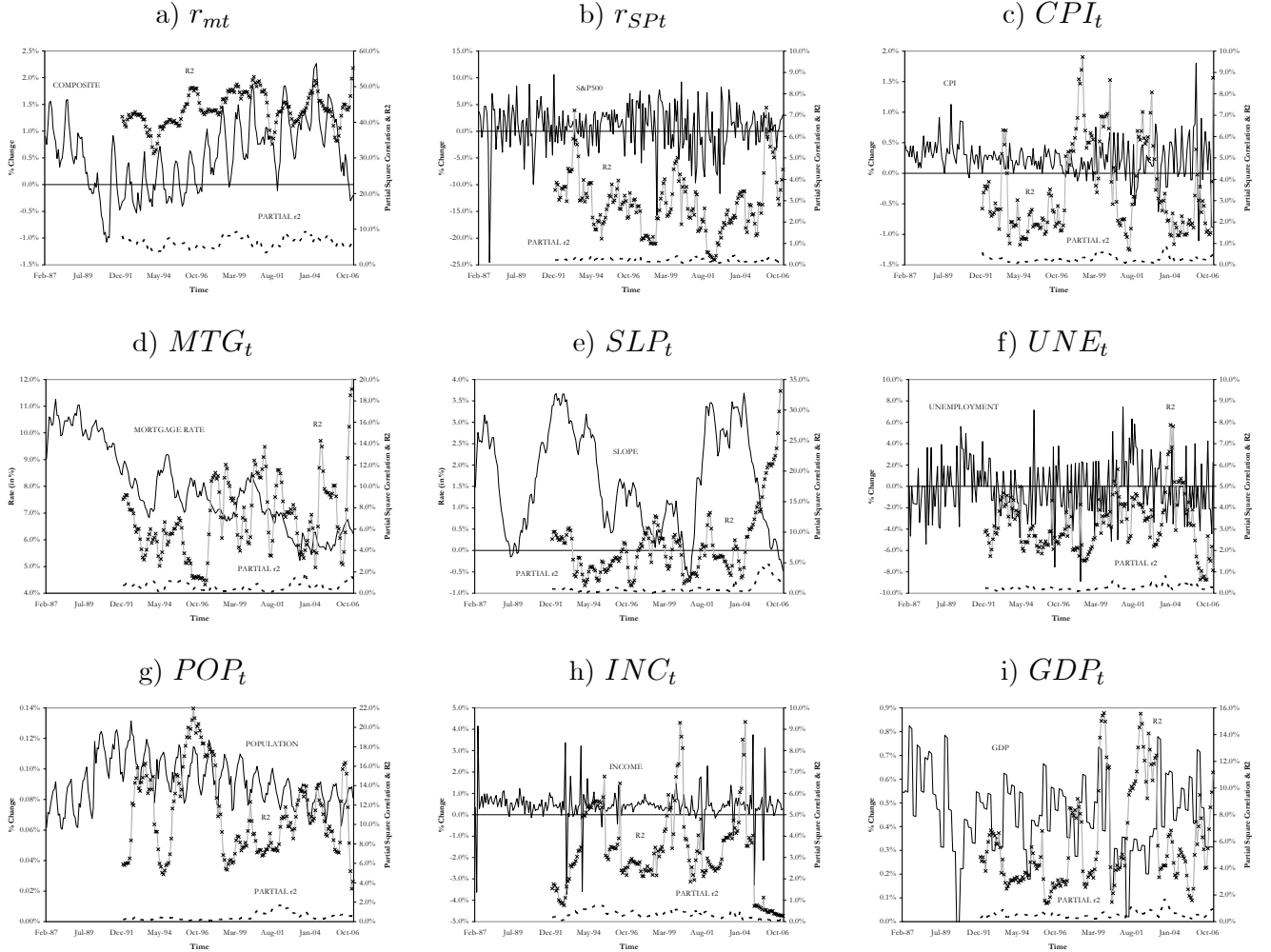
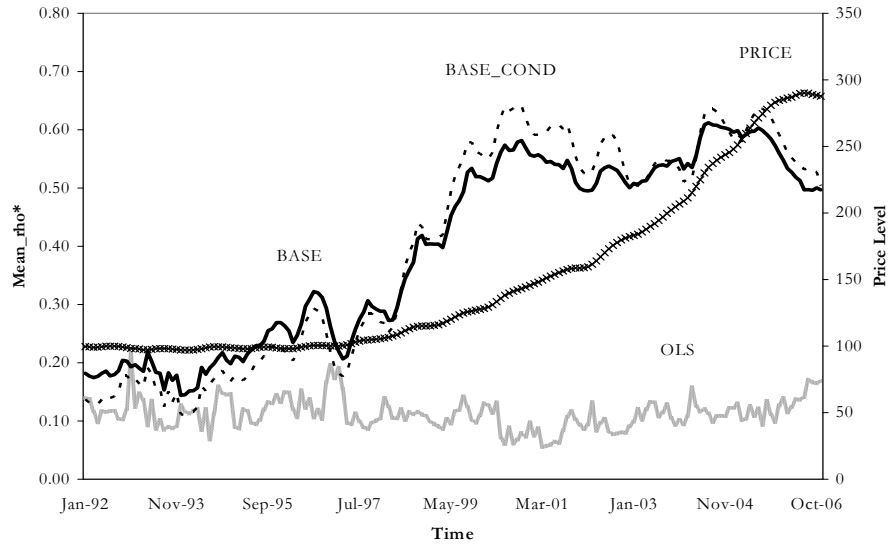


Figure 2. Mean excess correlation: OLS procedure

Figure 2a plots (on the left axis) the time series of mean excess square unconditional correlations  $\widehat{\rho}_t^{OLS*}$  of Eq. (11), a benchmark measure of square unconditional correlation,  $\widehat{\rho}_t^{BASE*}$ , computed using raw CSI returns  $r_{kt}$  (instead of estimated residuals  $\widehat{e}_{kt}^{OLS}$ ) in Eqs. (5) to (11), its equivalent,  $\widehat{\rho}_t^{BASE}$ , from their square conditional correlations. Figure 2b plots (in percentage terms) the ratio between  $\widehat{\rho}_t^{OLS*}$  and  $\widehat{\rho}_t^{BASE*}$  (defined in Table 2 as  $R\rho$ ) and  $\overline{R}_{at}^2$ , the mean adjusted  $R^2$  across all metropolitan markets from the OLS estimation of Eq. (2). Both figures plot (on the right axis) the composite CSI index with base 100 in December 1991.

a)  $\widehat{\rho}_t^{OLS*}$ ,  $\widehat{\rho}_t^{BASE*}$ , and  $\widehat{\rho}_t^{BASE}$  (left axis), composite CSI index (right axis)



b)  $\widehat{\rho}_t^{OLS*} / \widehat{\rho}_t^{BASE*}$  and  $\overline{R}_{at}^2$  in % (left axis), composite CSI index (right axis)

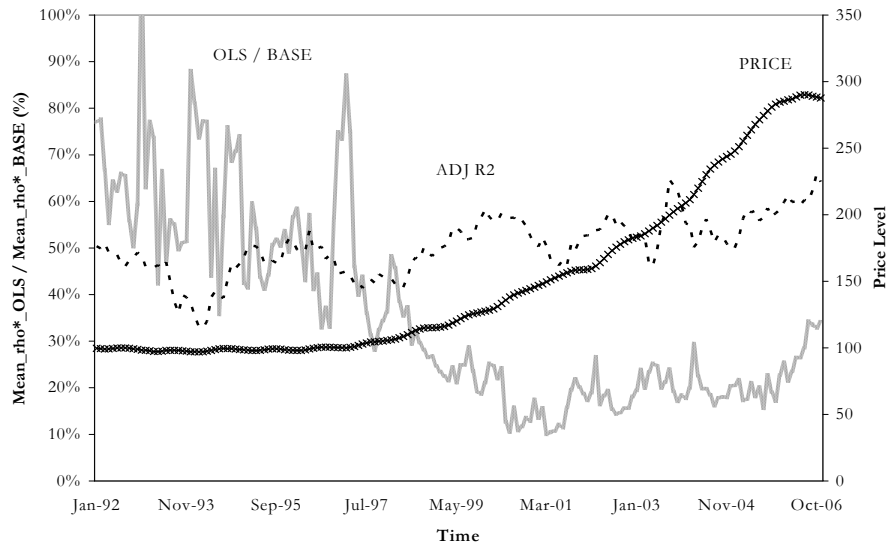


Figure 3. Mean excess correlation for metropolitan markets: OLS procedure

Figures 3a to 3n plot (on the left axis) the time series of mean excess square correlation  $\widehat{\rho}_{kt}^{OLS*}$  of Eq. (10) for each of the 14 metropolitan markets listed in Table 1, a benchmark measure of square correlation,  $\widehat{\rho}_{kt}^{BASE*}$ , computed using raw CSI returns  $r_{kt}$  (instead of estimated residuals  $\widehat{e}_{kt}^{OLS}$ ) in Eqs. (5) to (10), its equivalent,  $\widehat{\rho}_{kt}^{BASE}$ , from their square conditional correlations, and (on the right axis) the corresponding CSI index with base 100 in December 1991.

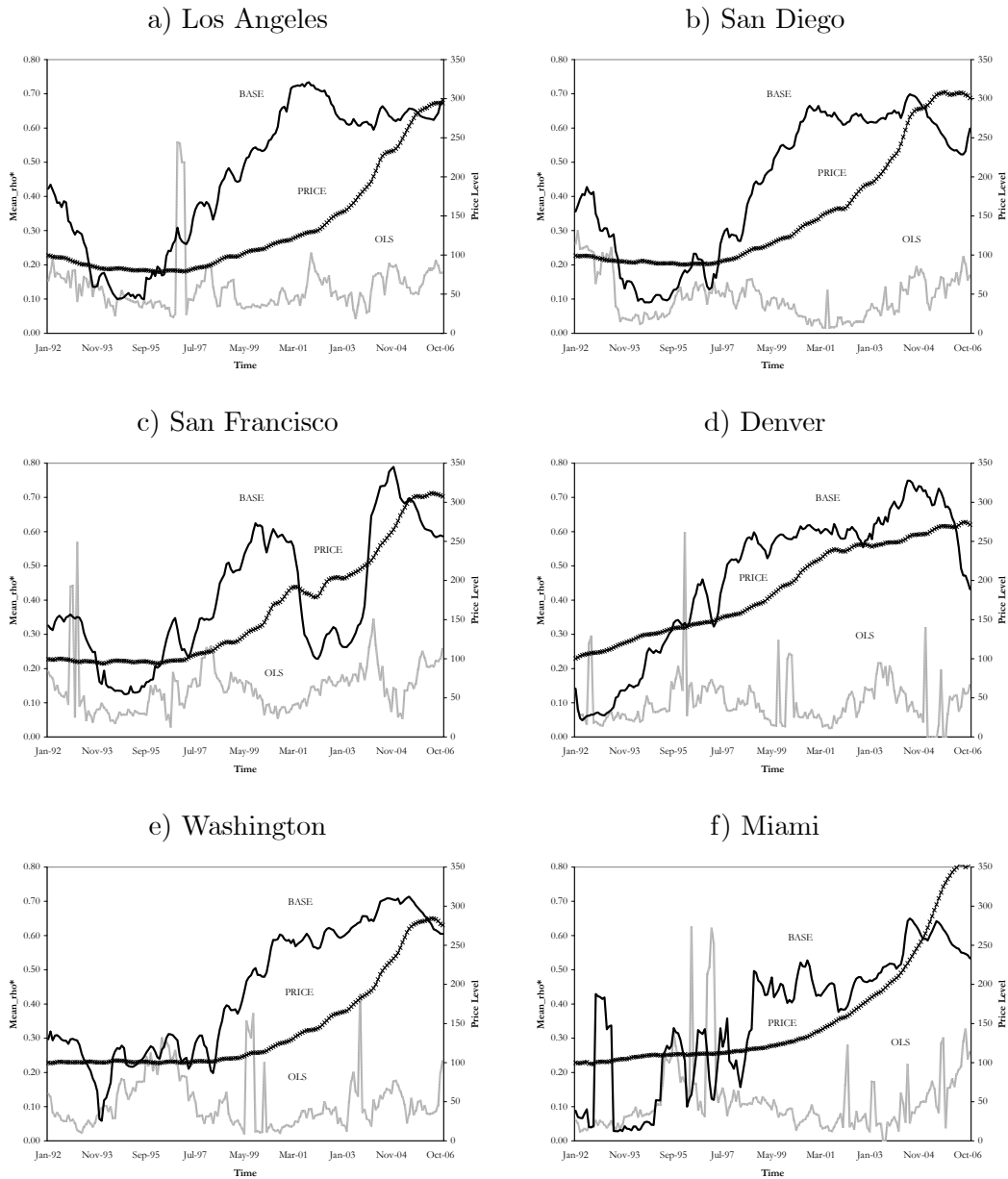
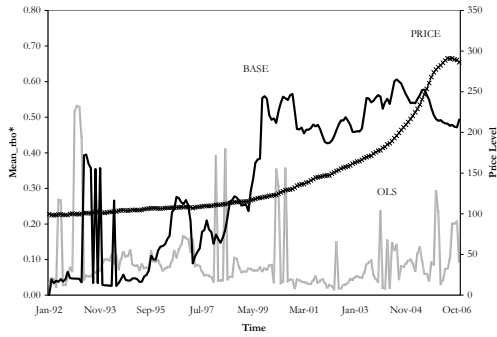
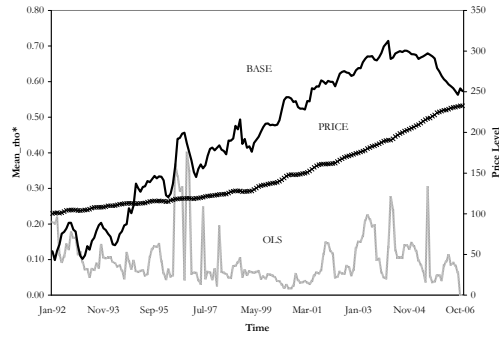


Figure 3 (Continued).

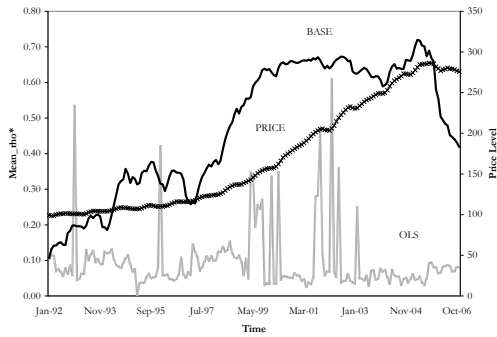
g) Tampa



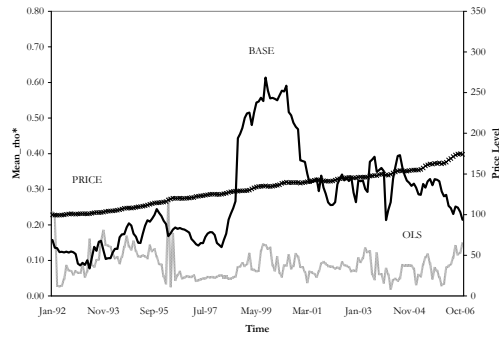
h) Chicago



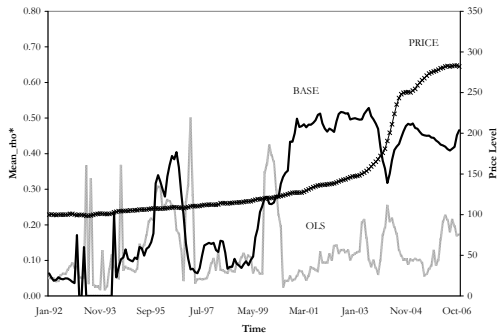
i) Boston



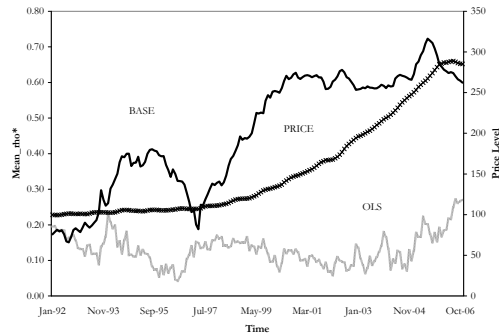
j) Charlotte



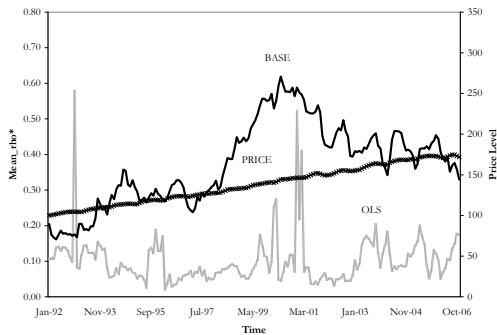
k) Las Vegas



l) New York



m) Cleveland



n) Portland

