

# On the Price Comovement of U.S. Residential Real Estate Markets

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

March 26, 2012

<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), Cornell University, School of Hotel Administration (chl62@cornell.edu), and Ross School of Business, University of Michigan (ppasquar@umich.edu), respectively. We benefited from the comments of Sreedhar Bharath, Michael Bordo, Pierre Siklos, Eugene White, and seminar participants at the Norges Bank Workshop on Fundamental and Non-Fundamental Asset Price Dynamics and the 2009 AREUEA meetings in San Francisco. We are grateful to Ryan Polansky for able research assistance. The usual disclaimer applies.

## Abstract

This study addresses the recent performance of the U.S. residential real estate markets. We investigate the comovement among Case-Shiller Home Price Indices for 14 metropolitan areas between 1992 and 2008. We identify the portion of this comovement deemed as *fundamental* (*excessive*), which we define as the covariation that can (cannot) be attributed to common fundamental factors directly influencing real estate prices. We find that the degree of comovement in these markets increased over the sample period, most significantly so in the late 1990s, but that this increase is largely due to systematic sources of risk; the degree of excess comovement is a less important factor. Further analysis indicates that the dynamics of comovement among metropolitan U.S. residential real estate markets within the sample period are mostly attributable to underlying systematic real and financial factors, consistent with a greater fundamental integration of those markets. We discuss the implications of these results for the evolution of U.S. real estate prices over the last two decades and the ongoing credit crisis.

*JEL classification:* E30; R20

*Keywords:* Real Estate; Correlation; Contagion; Housing Bubble; Credit Crisis

# 1 Introduction

The recent, widespread collapse of residential real estate prices in most metropolitan areas across the United States has been the subject of an immense amount of public outcry, policy debate, and academic scrutiny (e.g., see Acharya and Richardson, 2009; Brunnermeier, 2009). In many publications, the housing crisis has been compared to the *dot.com* bubble.<sup>1</sup> For instance, in July of 2007 Robert Shiller observed that housing prices were over-valued throughout the nation 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, relatively little research has been done to measure and explain the level and dynamics of U.S. real estate price *comovement* in correspondence with that crisis.<sup>3</sup> This is however an issue of increasing importance for understanding the nearly contemporaneous, ongoing turmoil in financial and credit markets since, for instance, the rating and valuation of the mortgage-backed securities currently held by many investors and financial institutions are crucially sensitive to whether the underlying portfolios of residential and/or commercial real estate loans were and still are well-diversified (e.g., Sanders, 2008; Gabriel et al., 2009).

This study contributes to fill this gap by performing a comprehensive empirical investigation of the process of price co-formation in the U.S. residential real estate market over the last two decades. We use a dataset of Case-Shiller Home Price Indices measuring the path of residential housing prices in 14 major U.S. markets. We then decompose the observed raw 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 crises, the degree

---

<sup>1</sup>E.g., the popular web-based encyclopedia Wikipedia reports that “[t]he 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.

<sup>3</sup>E.g., at the micro level, Ioannides (2002), Immergluck and Smith (2006), Harding et al. (2008), and Lin et al. (2009) find evidence of non-fundamental spillover effects of distressed, foreclosed, or vacant single-family homes in the U.S. on the value of non-distressed, non-vacant nearby properties between 1989 and 2007. At the macro level, Redfearn (1999) explores the role of industrial similarities in explaining the intermetropolitan correlation of quarterly housing returns between 1975 and 1996; Chiang (2010) reports that the comovement of monthly equity returns of Real Estate Investment Trusts (REITs) within the same property type — i.e., of REITs sensitive to common property-type information shocks — has increased between 1980 and 2004; Yunus and Swanson (2011) find significantly greater cointegration among broad U.S. housing regions since the bursting of the housing bubble in 2006.

of excess correlation is used as a measure of financial contagion. Importantly, our empirical strategy allows for both structural changes in the fundamental interdependence among residential housing price changes (i.e., time-varying factor sensitivities) and dynamics of their moments (i.e., nonstationarity).

Briefly noted, our novel results are the following. First, our analysis indicates that unconditional, observed raw comovement in U.S. residential real estate prices rose significantly between 1992 and 2008, more than doubling by the end of the sample period. Our analysis also suggests that much of that increase stems from fundamental correlation; much less can be attributed to excess correlation. This inference 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 primarily driven by common fundamental fluctuations in mortgage rates, realized and expected inflation, and GDP growth. Secondly, despite economically significant heterogeneity in the relation between raw and excess comovement across those markets, the magnitude and 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 statistically significant structural breaks upward over the sample period. All of these breaks occur in the late 1990s, well before most commentators were calling attention to residential price dynamics in the U.S., 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 sources of risk in real, financial, and real estate markets.<sup>4</sup>

Broadly interpreted, our evidence indicates that the increase in observed raw comovement among prices of metropolitan U.S. residential real estate markets over the last two decades is (economically and statistically) significant, persistent, and consistent with a greater fundamental integration of these markets.<sup>5</sup> As such, that increase may have also played a role in the relatively

---

<sup>4</sup>In contemporaneous studies, Miao et al. (2011) and Zhu et al. (2011) find evidence of significant *volatility interdependence* in the U.S. residential real estate market in the last two decades. Volatility dynamics have been shown to bias the estimation of measures of *return interdependence* (e.g., Forbes and Rigobon, 2002). As mentioned above, our statistical procedure is designed to control for those dynamics and their effects on levels and dynamics of residential real estate *return correlations*.

<sup>5</sup>Accordingly, Saks (2008) finds that national economic conditions can explain a “non-trivial” portion of both new construction and annual changes in house prices between 1981 and 2006. Yunus and Swanson (2011) find significantly greater cointegration among broad U.S. housing regions since the bursting of the housing bubble in 2006. They interpret this evidence as consistent with the notion that financial crises are typically accompanied by greater, possibly “contagious” cross-market linkages. As mentioned above, our analysis indicates that, once controlling for dynamics in volatility and factor loadings (also affected by periods of turmoil), raw (but not excess)

poorer performance of diversified mortgage lenders in the U.S. during the 2007-2008 credit crisis — as reported in some recent studies (e.g., Loutskina and Strahan, 2011; Purnanandam, 2011) — by making their loan portfolios more sensitive to real estate price fluctuations and thus weakening their overall financial positions more significantly during the accompanying economic downturn, i.e., precisely when diversification benefits were most necessary.

The outline 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 price comovement

There is an extensive literature studying the comovement among asset prices in *financial* markets.<sup>6</sup> The objective of this study is to assess the extent and dynamics of fundamental and excess comovement in *real asset* markets. In particular, we focus on the market for single-family residences in the U.S. Following that aforementioned literature (e.g., Bekaert et al., 2009), 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 the long-term valuation of those residences — and contagion as the circumstance of its occurrence. In this section we amend the multi-step methodology of Kallberg and Pasquariello (2008) to estimate the degree of intertemporal fundamental and 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 seasonally adjusted Home Price Indices (CSIs,  $r_{kt}$ ) between January 1987 and December 2008. These indices correspond to 14 individual metropolitan markets: Los Angeles, San Diego, San Francisco, Denver, Washington (DC), Miami, Tampa (FL), Chicago, Boston, Charlotte (NC), Las Vegas, New York, Cleveland (OH), and Portland (OR).<sup>7</sup>

CSI returns measure nominal price changes of individual single-family residences in the U.S. over time. 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

---

U.S. residential real estate return interdependence experienced a significant upward jump much earlier.

<sup>6</sup>E.g., King and Wadhvani (1990), Pindyck and Rotemberg (1990, 1993), Karolyi and Stulz (1996), Fleming et al. (1998), Barberis et al. (2005), Bekaert, Harvey, and Ng (2005), Kallberg et al. (2005), and Bekaert et al. (2009).

<sup>7</sup>Virtually identical inference ensues from employing non-seasonally adjusted CSI returns.

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. Since their inception, CSIs have become the most closely watched measure of U.S. home prices. Nonetheless, in unreported analysis we find the inference that follows to be robust to replacing CSI returns with monthly returns of regional House Price Indexes (HPIs) from the Office of Federal Housing Enterprise Oversight (OFHEO), available only from January 1991.

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 them.<sup>8</sup> Not surprisingly — given the growth experienced by the U.S. real estate market in the past two decades (despite its considerable downturn in the latter part of the sample) — all mean monthly returns are positive, statistically significant, and greater than the average monthly nationwide inflation (0.23%, 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 (2008) 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  $[t - N + 1, t]$ . In the spirit of Bekaert et al. (2009), we assume that, for each real asset  $k = 1, \dots, K$ , the return  $r_{kt}$  is characterized by the following linear factor structure:

$$r_{kt} = \alpha_{kt} + f_t \beta_{kt} + e_{kt}, \tag{1}$$

where  $f_t$  is a  $N \times N_f$  matrix of systematic factors  $f_{it}$  affecting all real 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>8</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  $k = 1, \dots, K$ . Of course, the model of Eq. (2) — like any explicit attempt to explain the commonality among residential real estate returns  $r_{kt}$  — may omit additional factors affecting all metropolitan areas in our sample, and therefore be misspecified. However, its estimation helps us to interpret our evidence, discussed in Section 3, on the level and dynamics of CSI return comovement. Its estimation also indicates that the chosen specification, albeit parsimonious, performs very satisfactorily in the data.

Housing price changes in individual metropolitan areas may be affected by nationwide fluctuations 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 (e.g., Painter and Redfearn, 2002, and references therein).  $SLP_t$ , 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 Federal Reserve Board of Governors) — 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 seasonally adjusted 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 fluctuations to impact the transaction prices of individual single-family residences (e.g., Potepan, 1996; Glaeser et al., 2005; Glaeser and Gyourko, 2006; Saks, 2008; Yunus and Swanson, 2011; Zhu et al., 2011):  $UNE_t$ , the monthly percentage change in the civilian unemployment rate (seasonally adjusted, from the Bureau of Labor Statistics);  $POP_t$ , the monthly percentage change in total U.S. population;

$INC_t$ , the monthly percentage change in seasonally adjusted disposable personal income; and  $GDP_t$ , the interpolated monthly percentage change in nominal Gross Domestic Product (GDP, seasonally adjusted), 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: For instance, 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 fundamental and 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  $\hat{e}_{kt}$ , where

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

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  $F_t = [\iota, f_t]$  is a  $N \times M$  matrix of systematic factors affecting  $r_{kt}$  (in which  $\iota$  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, p. 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 and raw 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 a consistent estimate of the unknown matrix  $\Sigma_t$ ,  $\widehat{\Sigma}_t^{OLS}$ , whose individual elements are (e.g., Greene, 1997, p. 676)

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

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

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

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  $\widehat{\rho}_{knt}$  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 Eqs. (1) to (3), 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. Failure to account for statistically insignificant  $\widehat{\rho}_{knt}$  may bias our analysis of the significance and extent of excess CSI return comovement. Hence, we conservatively set to zero all statistically insignificant conditional correlations for each market  $k = 1, \dots, K$  (according to the  $t$ -ratio test  $\widehat{t}_{knt}^{OLS} = \widehat{\rho}_{knt}^{OLS} \left[ \frac{1 - (\widehat{\rho}_{knt}^{OLS})^2}{N-2} \right]^{-\frac{1}{2}} \sim t[N-2]$ ) among the  $K-1$  possible  $\widehat{\rho}_{knt}$ , and measure excess comovement by computing the following means of excess *square* unconditional 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  $2 \left[1 - \Pr \left\{ |\widehat{t}_{knt}^{OLS}| \leq t_{\frac{\alpha}{2}} [N-2] \right\}\right] \leq \alpha$  and  $I_{knt}^{OLS} = 0$  otherwise. 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>9</sup> We follow the same procedure to obtain benchmark measures of *raw* comovement for each market  $k$  (labeled  $\widehat{\rho}_{kt}^{BASE*}$ ) as means of square unconditional correlations between any two raw CSI return series  $r_{kt}$  and  $r_{nt}$  (labeled  $\widehat{\rho}_{knt}^{*BASE}$ ) — i.e., by replacing the corresponding OLS residuals  $\widehat{e}_{kt}^{OLS}$  and  $\widehat{e}_{nt}^{OLS}$  with  $r_{kt}$  and  $r_{nt}$  in Eqs. (7) to (10):

$$\widehat{\rho}_{kt}^{BASE*} = \frac{1}{K-1} \sum_{\substack{n=1 \\ n \neq k}}^K (\widehat{\rho}_{knt}^{BASE*})^2 I_{knt}^{BASE}, \quad (11)$$

where  $\widehat{\rho}_{knt}^{*BASE} = \frac{\widehat{\rho}_{knt}^{BASE}}{\{1 + \widehat{\delta}_{kt}^{BASE} [1 - (\widehat{\rho}_{knt}^{BASE})^2]\}^{\frac{1}{2}}}$ ,  $\widehat{\rho}_{knt}^{BASE} = \frac{\widehat{\sigma}_{knt}^{BASE}}{[\widehat{\sigma}_{kk}^{BASE} \widehat{\sigma}_{nn}^{BASE}]^{\frac{1}{2}}}$ ,  $\widehat{\delta}_{kt}^{BASE} = \frac{\widehat{\sigma}_{kk}^{BASE}}{(\widehat{\sigma}_{kk}^{BASE})_{LT}} - 1$ ,  $\widehat{\sigma}_{knt}^{BASE} = \frac{(r_{kt})' r_{nt}}{N}$  in  $\widehat{\Sigma}_t^{BASE}$  for  $r_{kt}$ ,  $I_{knt}^{BASE} = 1$  if  $2 [1 - \Pr\{|\widehat{t}_{knt}^{BASE}| \leq t_{\frac{\alpha}{2}} [N-2]\}] \leq \alpha$  and  $I_{knt}^{BASE} = 0$  otherwise, and  $\widehat{t}_{knt}^{BASE} = \widehat{\rho}_{knt}^{BASE} \left[ \frac{1 - (\widehat{\rho}_{knt}^{BASE})^2}{N-2} \right]^{-\frac{1}{2}} \sim t [N-2]$ .<sup>10</sup> We also compute nationwide measures of excess and raw comovement among housing price changes as means of means of the aforementioned excess and raw square unconditional correlation coefficients, respectively, across all of the 14 metropolitan areas, i.e.,

$$\widehat{\rho}_t^{OLS*} = \frac{1}{K} \sum_{k=1}^K \widehat{\rho}_{kt}^{OLS*} \quad (12)$$

and

$$\widehat{\rho}_t^{BASE*} = \frac{1}{K} \sum_{k=1}^K \widehat{\rho}_{kt}^{BASE*}. \quad (13)$$

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 moments. Ignoring time-varying factor loadings and nonstationarity in Eq. (4) may bias the inference on non-fundamental and raw comovement from Eqs. (10) to (13). Parametric ARCH and stochastic volatility models, as well as their generalizations to multivariate settings, are frequently employed to describe such dynamics.<sup>11</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; Goetzmann et al., 2005). In addition, 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.

<sup>9</sup>Section 3.2 analyzes the robustness of our inference to this and other features of our empirical specification.

<sup>10</sup>First adjusting CSI returns by their sample-wide, or short and long-term means when computing  $\widehat{\rho}_{knt}^{*BASE}$  yields qualitatively similar inference.

<sup>11</sup>E.g., see Campbell et al. (1997) for a review of the literature on parametric models of changing volatility.



January 1987 to December 2008. We use the resulting estimated residuals and corresponding raw returns to compute conditional measures of excess and raw CSI return comovement ( $\widehat{\rho}_{knt}^{OLS}$  of Eq. (8) and  $\widehat{\rho}_{knt}^{BASE}$ ) for each pair of metropolitan markets  $k$  and  $n$ ; yet, in each month  $t$  we retain only the  $\widehat{\rho}_{knt}^{OLS}$  and  $\widehat{\rho}_{knt}^{BASE}$  that are statistically significant at the 10% level (e.g.,  $\alpha = 0.10$  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 returns and return residuals over a five-year interval (i.e.,  $g = 2$  in Eq. (14)). Therefore, the initial  $t = gN$  corresponds to January 1992.

We plot the resulting time series of nationwide excess and raw unconditional CSI return comovement —  $\{\widehat{\rho}_t^{OLS*}\}_{t=gN}^T$  and  $\{\widehat{\rho}_t^{BASE*}\}_{t=gN}^T$  of Eqs. (12) and (13), respectively — in Figure 2a between January 1992 and December 2008, as well as their conditional equivalents  $\{\widehat{\rho}_t^{OLS}\}_{t=gN}^T$  and  $\{\widehat{\rho}_t^{BASE}\}_{t=gN}^T$  computed using conditional correlations among those returns (i.e.,  $\widehat{\rho}_{knt}^{BASE}$  and  $\widehat{\rho}_{knt}^{OLS}$  of Eq. (8)). 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 nonstationary CSI return means and 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 and raw comovement 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 is high — e.g., the average  $\widehat{\rho}_t^{BASE*} = 0.417$  in Table 2 — 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 more than doubles in the second half of the sample period (0.530 on average for the interval 1998-2008) as compared to its mean in 1997 (0.197). 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 both the period of stable or protracted declining housing prices that followed the brief recession of 1990-1991 (e.g., Los Angeles in Figure 3a) and the continued slump in the U.S. real estate market between 2005 and 2008 (e.g., Boston in Figure 3i).

Further analysis reveals that an economically significant portion of the observed raw 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 37% over our sample period and is never lower than 32% in its latter part, between 2002 and 2008.<sup>12</sup> More interestingly, the average explanatory power of the systematic factors listed in Section 2.2 nearly steadily increases toward 80% 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 factors  $f_{it}$  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,  $\overline{R}_{fit}^2$ , 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. (14). The second one,  $\overline{r}_{fit}^2$ , addresses possible multicollinearity bias in  $\overline{R}_{fit}^2$ : 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>13</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 1997 may be attributed to the increasing importance of common real estate fluctuations ( $r_{mt}$  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 1g and 1i) for the price changes of individual single-family residences in the U.S.<sup>14</sup>

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,

<sup>12</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  $S_{knt} = N^{-1} r'_{kt} r_{nt} - \overline{r}_{kt} \overline{r}_{nt}$  — averages 95% and is always greater than 79% over the entire sample, 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>13</sup>Thus, the ensuing  $\overline{r}_{fit}^2$  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).

<sup>14</sup>Consistently, Zhu et al. (2011) provide evidence of a significant relationship between U.S. residential housing returns and both population growth and (lagged) lagged personal income.

nationwide excess square correlation  $\widehat{\rho}_t^{OLS*}$  averages about 0.107 — equivalent to an average absolute return residual correlation of  $\left| \sqrt{\widehat{\rho}_t^{OLS*}} \right| = 0.328$  — 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 around 26% 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  $1 - R\rho \approx 74\%$  — 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 high in the early to mid-1990s, when it accounts for between 30% and 80% of the raw comovement measure  $\widehat{\rho}_t^{BASE*}$  (the ratio  $\widehat{\rho}_t^{OLS*} / \widehat{\rho}_t^{BASE*}$  in Figure 2b). 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 slowly increasing again by the end of the sample period, accompanying the sharp nationwide decline of residential housing prices between 2006 and 2008.

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 among the lowest (0.330 in Table 2), but the corresponding ratio between  $\widehat{\rho}_{kt}^{OLS*}$  and  $\widehat{\rho}_{kt}^{BASE*}$  is the highest ( $R\rho = 45\%$  in Table 2). Both raw and excess price comovement are high in California, Denver, Miami, 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 briefly decline and comove significantly less with the rest of the country than in the 1990s; those prices’ subsequent increase (then decrease) is instead accompanied by sharply higher (then lower) raw and excess comovement of their dynamics with those observed nationwide. Consistently,  $\beta\rho$  for San Francisco is the highest (0.523) in Table 2. We observe a similar trend in both  $\widehat{\rho}_{kt}^{BASE*}$  and  $\widehat{\rho}_{kt}^{OLS*}$  in proximity to the sharp increase in residential housing prices in Las Vegas, between 2003 and 2005, and of their ensuing collapse, between 2005 and 2008 (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 observed raw and excess comovement among U.S. residential housing prices may bias 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 *absolute* (rather than square) CSI return and return residual correlations. Specifically, we redefine  $\hat{\rho}_{kt}^{OLS*} = \frac{1}{K-1} \sum_{\substack{n=1 \\ n \neq k}}^K |\hat{\rho}_{knt}^{OLS*}| I_{knt}^{OLS}$  and  $\hat{\rho}_{kt}^{BASE*} = \frac{1}{K-1} \sum_{\substack{n=1 \\ n \neq k}}^K |\hat{\rho}_{knt}^{BASE*}| I_{knt}^{BASE}$ , and perform all subsequent steps listed in Section 2 accordingly. This definition is as effective as that of Eq. (10) or (11) at preventing excess or raw return 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 returns and return residuals,  $\hat{\rho}_{knt}^{BASE}$  and  $\hat{\rho}_{knt}^{OLS}$ , may bias our empirical tests. We preliminarily addressed this issue in Section 2.4 by imposing that all statistically insignificant conditional correlation coefficients (at the 10% level,  $\alpha = 0.10$ ) be equal to zero when computing each aggregate measure  $\hat{\rho}_{kt}^{BASE*}$  and  $\hat{\rho}_{kt}^{OLS*}$ . The resulting percentages of statistically significant conditional correlations from either  $\hat{\Sigma}_t^{BASE}$  or  $\hat{\Sigma}_t^{OLS}$  (in columns  $T\rho$  of Table 2) are relatively high — averaging either 81% or 29% — and fluctuate either between 40% and 90% or between 20% and 40% for most of the sample period. As importantly, the ensuing unconditional coefficients of determinations  $(\hat{\rho}_{knt}^{BASE*})^2$  and  $(\hat{\rho}_{knt}^{OLS*})^2$ , in columns  $F\rho^{*2}$  of Table 2, are also statistically significant (according to the  $t$ -square ratio test  $\hat{t}_{knt}^2$ ) and with almost identical frequencies across metropolitan areas and over time. Furthermore, the Lagrange multiplier (LM) statistic of Breusch and Pagan (1980) —  $\hat{\lambda}_t = N \sum_{k=2}^K \sum_{n=1}^{k-1} (\hat{\rho}_{knt})^2 \sim \chi^2 \left[ \frac{K(K-1)}{2} \right]$  — leads us to reject the null hypothesis that either  $\hat{\Sigma}_t^{BASE}$  or  $\hat{\Sigma}_t^{OLS}$  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  $\hat{\rho}_{knt}^{OLS}$

and  $\widehat{\rho}_{knt}^{BASE}$  in the excess and raw comovement measures  $\widehat{\rho}_{kt}^{OLS*}$  (Eq. (10)) and  $\widehat{\rho}_{kt}^{BASE*}$  (Eq. (11)), respectively.

### 3.2.3 Conditional comovement

Our inference in Section 3.1 is also 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 since the aforementioned evidence about  $\widehat{\rho}_{knt}^{BASE*}$  in Table 2 suggests that unconditional correlations of raw CSI returns  $r_{kt}$  are large and statistically significant.<sup>15</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 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>16</sup>

<sup>15</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>16</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 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 nationwide excess comovement measure  $\widehat{\rho}_t^{OLS*}$  or each metropolitan-wide measure  $\widehat{\rho}_{kt}^{OLS*}$  (between 25% and 35%, in column  $T\rho$  of Table 2).

### 3.3 Regime shifts in CSI 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 this 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 this 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,

$$\widehat{\rho}_t^{BASE*} = a + b\widehat{\rho}_t^{OLS*} + \varepsilon_t, \quad (15)$$

as well as for each metropolitan market  $k$ ,

$$\widehat{\rho}_{kt}^{BASE*} = a_k + b_k\widehat{\rho}_{kt}^{OLS*} + \varepsilon_{kt}, \quad (16)$$

where  $k = 1, \dots, 14$ . We then test for breaks in the parameters of Eqs. (15) and (16). 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. (15) and (16) 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

$$y_t = X_t'\phi + d_t(\tau) X_t'S'\Delta\phi + \varepsilon_t, \quad (17)$$

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  $\phi$  that are allowed to change. Hence, Eq. (17) 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

$$y_t = Z_t(\tau)'\Phi(\tau) + \varepsilon_t, \quad (18)$$

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. (17), based on Wald statistics (e.g., Quandt, 1958, 1960), considers the maximum of the following  $F$  process:

$$F(\tau) = T \left[ R\widehat{\Phi}(\tau) \right]' \left\{ R \left[ T^{-1} \sum_{t=1}^T Z_t(\tau) \widehat{\sigma}^2 Z_t(\tau)' \right]^{-1} R' \right\}^{-1} \left[ R\widehat{\Phi}(\tau) \right], \quad (19)$$

where  $R = [0, S]$  is such that  $R\Phi(\tau) = S\Delta\phi$ , and  $\widehat{\Phi}(\tau)$  and  $\widehat{\sigma}^2$  are OLS estimators for  $\Phi(\tau)$  and  $\text{var}(\varepsilon)$ , respectively, under the alternative hypothesis of one break at date  $\tau$ . The estimated break date,  $\widehat{\tau} = \arg \max F(\tau)$ , is statistically significant if  $F(\widehat{\tau})$  is greater than its critical value at the chosen level of significance.<sup>17</sup>

If  $\widehat{\tau}$  is statistically significant, a confidence interval around it is typically obtained by assuming that residuals in Eq. (17) are normally distributed.<sup>18</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  $[\widehat{\tau}^-, \widehat{\tau}^+]$  for the true break date:

$$\widehat{\tau}^\pm = \widehat{\tau} \pm c_{\frac{1}{2}\pi} \left\{ (S\Delta\phi)' S \left[ (\widehat{\sigma}^2 T)^{-1} \sum_{t=1}^T X_t X_t' \right] S' (S\Delta\phi) \right\}^{-1}, \quad (20)$$

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. (19) and its associated confidence interval around  $\widehat{\tau}$  (Eq. (20)) 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. (17).

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. (15)) and for each metropolitan market in our sample (Eq. (16)). All of the estimated break dates cluster in the late 1990s and in no circumstance (with the sole exception of Cleveland) is the corresponding estimated confidence interval wider than one month. The estimated breaks are economically significant as well: Most absolute shifts in either  $\widehat{a}$ ,  $\widehat{a}_k$ ,  $\widehat{b}$ ,

<sup>17</sup>To compute these critical values, Bai et al. (1998) suggest to approximate the limiting distribution of  $F(\widehat{\tau})$  with partial sums of normal random variables for each possible rank of the selection matrix  $S$  in Eq. (17). Bekaert, Harvey, and Lumsdaine (2002) report a table with critical values for up to 68 parameters allowed to break.

<sup>18</sup>Bai et al. (1998) review the available literature on the topic.

or  $\widehat{b}_k$  are for more 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 231% higher between September 1998 and December 2008 ( $\widehat{a} + \widehat{\Delta a} = 0.542$ ) than in the previous six years ( $\widehat{a} = 0.235$ ); among the metropolitan markets in our sample, nominal price changes of individual single-family residences in San Diego, San Francisco, Miami, Las Vegas, Boston, and New York 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 over the sample period:  $\widehat{b} \approx 0$  and  $\widehat{\Delta b} \approx 0$  in Table 3 for the model for  $\widehat{\rho}_t^{BASE*}$  (Eq. (15), 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 more heterogeneous. Specifically, Table 3 reports that only for Chicago, Tampa, and Charlotte does excess comovement play a more important role for the dynamics of  $\widehat{\rho}_{kt}^{BASE*}$  following their corresponding estimated break dates ( $\widehat{b}_k \approx 0$  and  $\widehat{\Delta b}_k > 0$ ). Vice versa, the observed increase in raw CSI return correlations in Boston, New York, and Cleveland cannot be attributed to  $\widehat{\rho}_{kt}^{OLS*}$  ( $\widehat{b}_k \approx 0$  and  $\widehat{\Delta b}_k \lesssim 0$ ). Interestingly, excess comovement — albeit being positively related to  $\widehat{\rho}_{kt}^{BASE*}$  during the 1990s — does not appear to explain the subsequent 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 latter part of the sample period (see Table 1):  $\widehat{b}_k > 0$  and  $\widehat{\Delta b}_k \lesssim 0$  in California (Los Angeles, San Diego, San Francisco), Colorado (Denver), Nevada (Las Vegas), and Oregon (Portland). Excess unconditional correlation  $\widehat{\rho}_{kt}^{OLS*}$  instead either remains or becomes more relevant for  $\widehat{\rho}_{kt}^{BASE*}$  in the Mid-Atlantic and South-East:  $\widehat{b}_k \gtrsim 0$  and/or  $\widehat{b}_k + \widehat{\Delta b}_k \gtrsim 0$  for the comovement of CSI price changes in Washington, Miami, Tampa, and Charlotte with the rest of the U.S. Overall, the evidence in Table 3 indicates that the observed increase in the raw 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 ( $\widehat{\rho}_t^{BASE*}$ ) is both statistically and economically significant; (ii) observed raw CSI comovement experienced a sharp increase over the last two decades; (iii) the extent and dynamics of observed raw CSI comovement appear to be primarily

related to common fundamental sources of risk in U.S. real estate, financial, and real markets, especially to mortgage rates, realized and expected inflation, and GDP growth; (iv) comovement in excess of those systematic factors ( $\widehat{\rho}_t^{OLS*}$ ) is nonetheless non-trivial; (v) there is economically significant heterogeneity in the extent and dynamics of raw and excess comovement across U.S. metropolitan markets ( $\widehat{\rho}_{kt}^{BASE*}$  and  $\widehat{\rho}_{kt}^{OLS*}$ ); and (vi) 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  $\widehat{\rho}_t^{BASE*}$  (and  $\widehat{\rho}_{kt}^{BASE*}$ ) and  $\widehat{\rho}_t^{OLS*}$  (and  $\widehat{\rho}_{kt}^{OLS*}$ ) 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.

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>19</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>19</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>20</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 1992-2008 (Panel A of Table 4). For instance, the adjusted  $R^2$  is about 18% for  $\hat{\rho}_t^{BASE*}$  and between 12% (San Francisco) and 56% (Denver) 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 47% higher than its conditional mean ( $\frac{0.134}{0.284}$ ) when  $d_t^R = 1$  and 23% higher when  $d_{SPt}^- = 1$  ( $\frac{0.066}{0.284}$ ) at the 1% significance level. Among the metropolitan areas in our sample, raw comovement of CSI returns for Las Vegas, Tampa, and Washington is the most sensitive to fluctuations in the U.S. business cycle and stock market downturns. Vice versa, raw comovement of residential real estate prices with the rest of the U.S. is unrelated to prolonged stock market increases ( $d_{SPt}^+ = 1$ ). Furthermore, both  $\hat{\rho}_t^{BASE*}$  and  $\hat{\rho}_{kt}^{BASE*}$  are almost uniformly 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 56% to nationwide raw comovement ( $\frac{0.160}{0.284}$ ), and between 25% ( $\frac{0.038}{0.153}$  for Washington) and as much as 105% ( $\frac{0.274}{0.262}$  for Boston) to raw metropolitan comovement. According to Panel A of Table 4, square correlation of raw CSI returns is positively, albeit more weakly, related to downward

---

<sup>20</sup>King et al. (1994) and Carrieri et al. (2006), among others, estimate linear regressions whose dependent variables are correlation coefficients.

momentum in  $r_{mt}$  as well, especially in California, Miami, Tampa, and Las Vegas and especially in the latter part of the sample (e.g., see Figures 3a, 3c, 3f, 3g, and 3k). The estimated impact of prolonged upswings and downswings in  $r_{kt}$  on the corresponding measure  $\hat{\rho}_{kt}^{BASE*}$  is of similar significance and absolute and relative magnitude.

A less clear, more heterogeneous picture emerges from the analysis of excess CSI return comovement conditional upon the state of the U.S. economy, in Panel B of Table 4. For example, the corresponding adjusted  $R^2$  (Column  $R_a^2$ ) are 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 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 lower in correspondence with prolonged upswings in residential housing prices nationwide — on average by about 17% ( $\frac{-0.016}{0.117}$  when  $d_t^+ = 1$ ); yet, the local excess comovement measures  $\hat{\rho}_{kt}^{OLS*}$  are generally insensitive to fluctuations in both  $r_{mt}$  and  $r_{kt}$ . Lastly, and again contrary to what is 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 — between 21% ( $\frac{-0.027}{0.126}$  for Cleveland) and 107% ( $\frac{-0.059}{0.055}$  for San Francisco) when  $d_t^R = 1$  — but is generally insensitive to stock market fluctuations, especially for some urban areas in the West (Los Angeles, San Francisco, and Denver) and the East Coast (Boston, New York, Miami, and Tampa).

In short, the evidence in Table 4 suggests that both raw and excess square CSI return correlation between 1992 and 2008 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 sources of risk in real, financial, and real estate markets.

## 4 Conclusions

This paper studies the process of price co-formation in the U.S. residential real estate market over the last two decades. Using housing price index data for 14 of the largest U.S. metropolitan areas, we find that observed price comovement among these markets increased between 1992 and 2008, and that this increase stems mostly from fluctuations in fundamental factors driving real estate returns. 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., 2009).

Our empirical conclusions arise from a distinct, novel 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, especially over the latter part of our data. We then use this model to decompose the observed raw 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). We find that raw unconditional comovement among regional residential real estate markets increases significantly over our sample period, but so does the explanatory power of the common factors driving residential real estate prices. Thus, excess comovement explains a declining portion of the observed comovement of those prices in the sample.

We also analyze the possible determinants of the extent and dynamics of raw and excess comovement among U.S. metropolitan real estate markets. First, a structural break analysis indicates that raw comovement among price changes of single-home residences in all but one of the 14 metropolitan areas in our sample experiences a (statistically and economically) significant upward regime shift in the late 1990s; yet none of these breaks can be attributed to an increase in excess comovement. Second, we show that both raw and excess 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 the debate about the recent U.S. residential real estate crisis. It shows that a parsimonious fundamental model of residential real estate returns can explain much 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 crisis. Our evidence of increasing integration of regional residential real estate markets in the U.S. may also *(i)* help interpret the poor performance of diversified mortgage lenders in the U.S. (especially relative to more concentrated ones) over that same period (e.g., Loutskina and Strahan, 2011; Purnanandam, 2011); and *(ii)* suggest that future shocks to any of those regional markets are more likely to have nationwide, prolonged effects than they did in the past.

## References

- Acharya, V., Richardson, M. eds. (2009). *Restoring Financial Stability: How to repair a Failed System*. Wiley, New York.
- Bai, J., Lumsdaine, R., Stock, J. (1998). Testing for and Dating Common Breaks in Multivariate Time Series. *Review of Economic Studies*, 65, 395-432.
- Barberis, N., Shleifer, A., Wurgler, J. (2005). Comovement. *Journal of Financial Economics*, 75, 283-317.
- Bekaert, G., Harvey, C., Lumsdaine, R. (2002). Dating the Integration of World Equity Markets. *Journal of Financial Economics*, 65, 203-247.
- Bekaert, G, Harvey, C., Ng, A. (2005). Market Integration and Contagion. *Journal of Business*, 78, 39-69.
- Bekaert, G., Hodrick, R. , Zhang, X. (2009). International Stock Return Comovements. *Journal of Finance*, 64, 2591-2626..
- 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, 239-254.
- Brunnermeier, M. (2009). Deciphering the 2007-08 Liquidity and Credit Crunch. *Journal of Economic Perspectives*, 23, 77-100.
- 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, 1-43.
- Carrieri, F., Errunza, V., Sarkissian, S. (2006). Economic Integration, Industrial Structure, and International Portfolio Diversification, Working Paper, McGill University.
- Chiang, K. (2010). On the Comovement of REIT Prices. *Journal of Real Estate Research*, 32, 187-200.

- 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, 1177-1199.
- De Bondt, W., Thaler, R. (1985). Does the Stock Market Overreact? *Journal of Finance*, 40, 793-805.
- De Bondt, W., Thaler, R. (1987). Further Evidence on Investor Overreaction and Stock Market Seasonality. *Journal of Finance*, 42, 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, 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, 2223-2261.
- Gabriel, S., Quigley, J., Rosenthal, L. (2009). The Mortgage Meltdown, the Economy, and Public Policy. *B.E. Journal of Economic Analysis & Policy*, 9, 1-16.
- Glaeser, E., Gyourko, J. (2006). Housing Dynamics. NBER Working Paper No. 12787.
- Glaeser, E., Gyourko, J., Saks, R. (2005). Why Have House Prices Gone Up? *American Economic Review*, 95, 329-333.
- Goetzmann, W., Li, L., Rouwenhorst, G. (2005). Long-Term Global Market Correlations. *Journal of Business*, 78, 1-38.
- Greene, W. (1997). *Econometric Analysis*. Prentice Hall, New Jersey.
- Harding, J., Rosenblatt, E., Yao, V. (2008). The Contagion Effect of Foreclosed Properties. Working Paper, University of Connecticut.
- Immergluck, D., Smith, G. (2006). The External Costs of Foreclosure: The Impact of Single-Family Mortgage Foreclosures on Property Values. *Housing Policy Debate*, 17, 57-79.
- Ioannides, Y. (2002). Interactive Property Valuations. *Journal of Urban Economics*, 53, 145-170.

- Jegadeesh, N., Titman, S. (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *Journal of Finance*, 48, 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, 169-211.
- Kallberg, J., Pasquariello, P. (2008). Time Series and Cross-Sectional Excess Comovement in Stock Indexes. *Journal of Empirical Finance*, 15, 481-502.
- Karolyi, A., Stulz, R. (1996). Why Do Markets Move Together? An Investigation of U.S.-Japan Stock Return Co-movements. *Journal of Finance*, 51, 951-986.
- King, M., Wadhvani, S. (1990). Transmission of Volatility between Stock Markets. *Review of Financial Studies*, 3, 5-33.
- King, M., Sentana, E., Wadhvani, S. (1994). Volatility and Links between National Stock Markets. *Econometrica*, 62, 901-934.
- Lin, Z., Rosenblatt, E., Yao, V. (2009). Spillover Effects of Foreclosure on Neighborhood Property Values. *Journal of Real Estate Finance and Economics*, 38, 387-407.
- 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.
- Loutskina, E., Strahan, P. (2011). Informed and Uninformed Investment in Housing: The Downside of Diversification. *Review of Financial Studies*, 24, 1447-1480.
- 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, 381-387.
- Miao, H., Ramchander, S., Simpson, M. (2011). Return and Volatility Transmission in U.S. Housing Markets. *Real Estate Economics*, 2011, 701-741.
- Nelson, C. (1992). Filtering and Forecasting with Misspecified ARCH Models I: Getting the Right Variance with the Wrong Model. *Journal of Econometrics*, 52, 61-90.
- Officer, R. (1973). The Variability of the Market Factor of the New York Stock Exchange. *Journal of Business*, 46, 434-453.

- Painter, G., Redfearn, C. (2002). The Role of Interest Rate in Influencing Long-Run Homeownership Rates. *Journal of Real Estate Finance and Economics*, 25, 243-267.
- 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, 619-632.
- Pavlova, A., Rigobon, R. (2007). Asset Prices and Exchange Rates. *Review of Financial Studies*, 20, 1139-1181.
- Picard D. (1985). Testing and Estimating Change-Points in Time Series. *Advances in Applied Probability*, 17, 841-867.
- Pindyck, R., Rotemberg, J. (1990). The Excess Comovement of Commodity Prices. *Economic Journal*, 100, 1173-1189.
- Pindyck, R., Rotemberg, J. (1993). The Comovement of Stock Prices. *Quarterly Journal of Economics*, 108, 1073-1104.
- Potepan, M. (1996). Explaining Intermetropolitan Variation in Housing Prices, Rents and Land Prices. *Real Estate Economics*, 24, 219-245.
- Purnanandam, A. (2011). Originate-to-Distribute Model and the Sub-Prime Mortgage Crisis. *Review of Financial Studies*, 24, 1881-1915.
- Quandt, R. (1958). The Estimation of the Parameters of a Linear Regression System Obeying Two Separate Regimes. *Journal of the American Statistical Association*, 53, 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, 324-330.
- Redfearn, C. (1999). The Composition of Metropolitan Employment and the Correlation of Housing Prices Across Metropolitan Areas. Working Paper, UC Berkeley.
- Saks, R. (2008). Resurrecting the National Housing Market: Reassessing the Role of National and Local Shocks in Metropolitan Area Housing Markets. Working Paper, Federal Reserve Board of Governors.
- Sanders, A. (2008). The Subprime Crisis and Its Role in the Financial Crisis. *Journal of Housing Economics*, 17, 254-261.
- Sias, R. (2004). Institutional Herding. *Review of Financial Studies*, 17, 165-206.

Wermers, R. (1999). Mutual Fund Herding and the Impact on Stock Prices. *Journal of Finance*, 54, 581-622.

Yunus, N., Swanson, P. (2011). A Closer Look at the U.S. Housing Market: Modeling Relationships among Regions. *Real Estate Economics*, forthcoming.

Zhou, G. (1993). Asset-Pricing Tests under Alternative Distributions. *Journal of Finance*, 48, 1927-1942.

Zhu, B., Füss, R., Rottke, N. (2011). Spatial Linkages in Returns and Volatilities among U.S. Regional Housing Markets. *Real Estate Economics*, forthcoming.

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 - December 2008 (263 observations).  $\mu$  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.  $\hat{\rho}_1$  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	$\mu$	$\sigma$	Skew	Kurt	$\hat{\rho}_1$	$LB(5)$
Composite	0.359% <sup>†</sup>	0.79%	-0.99 <sup>†</sup>	1.48 <sup>†</sup>	1.030 <sup>†</sup>	323.82 <sup>†</sup>
Metropolitan Markets						
Los Angeles	0.403% <sup>†</sup>	1.20%	-0.60 <sup>†</sup>	0.94 <sup>†</sup>	0.825 <sup>†</sup>	215.02 <sup>†</sup>
San Diego	0.391% <sup>†</sup>	1.15%	-0.37 <sup>•</sup>	1.66 <sup>†</sup>	0.432 <sup>†</sup>	94.15 <sup>†</sup>
San Francisco	0.390% <sup>†</sup>	1.23%	-1.03 <sup>†</sup>	2.88 <sup>†</sup>	0.915 <sup>†</sup>	263.55 <sup>†</sup>
Denver	0.350% <sup>†</sup>	0.53%	-0.42 <sup>†</sup>	0.75 <sup>•</sup>	0.617 <sup>†</sup>	170.71 <sup>†</sup>
Washington	0.383% <sup>†</sup>	0.89%	-0.50 <sup>†</sup>	0.74 <sup>•</sup>	0.718 <sup>†</sup>	160.86 <sup>†</sup>
Miami	0.333% <sup>†</sup>	1.10%	-1.30 <sup>†</sup>	3.67 <sup>†</sup>	0.617 <sup>†</sup>	202.86 <sup>†</sup>
Tampa	0.265% <sup>†</sup>	0.93%	-0.75 <sup>†</sup>	2.89 <sup>†</sup>	0.638 <sup>†</sup>	193.72 <sup>†</sup>
Chicago	0.353% <sup>†</sup>	0.65%	-1.13 <sup>†</sup>	5.09 <sup>†</sup>	0.574 <sup>†</sup>	156.64 <sup>†</sup>
Boston	0.299% <sup>†</sup>	0.67%	-0.36 <sup>•</sup>	0.13	0.664 <sup>†</sup>	216.36 <sup>†</sup>
Charlotte	0.252% <sup>†</sup>	0.41%	-1.16 <sup>†</sup>	4.95 <sup>†</sup>	0.430 <sup>†</sup>	146.94 <sup>†</sup>
Las Vegas	0.260% <sup>†</sup>	1.29%	-0.28 <sup>*</sup>	5.05 <sup>†</sup>	0.627 <sup>†</sup>	127.17 <sup>†</sup>
New York	0.341% <sup>†</sup>	0.67%	-0.20	-0.27	0.745 <sup>†</sup>	209.92 <sup>†</sup>
Cleveland	0.254% <sup>†</sup>	0.48%	-1.21 <sup>†</sup>	6.22 <sup>†</sup>	0.230 <sup>†</sup>	79.46 <sup>†</sup>
Portland	0.513% <sup>†</sup>	0.65%	-0.31 <sup>•</sup>	2.22 <sup>†</sup>	0.519 <sup>†</sup>	148.78 <sup>†</sup>

Table 2. Descriptive statistics: BASE and OLS comovement

This table reports summary statistics for the monthly time series of both benchmark square correlations of raw CSI returns ( $\widehat{\rho}_{knt}^{BASE*}$ ) and excess square correlations of OLS residuals ( $\widehat{\rho}_{knt}^{OLS*}$  of Eq. (12)). In particular, the table reports their nationwide averages ( $\widehat{\rho}_t^{BASE*}$  and  $\widehat{\rho}_t^{OLS*}$  of Eqs. (13) and (12), respectively) as well as their averages for each of the 14 individual metropolitan market indices ( $\widehat{\rho}_{kt}^{BASE*}$  and  $\widehat{\rho}_{kt}^{OLS*}$  of Eqs. (11) and (10), respectively), defined in Section 4, over the interval January 1992 - December 2008 (204 observations). OLS residuals are estimated according to the procedure described in Section 3 from the following specification of the fundamental model of Eq. (2):

$$\begin{aligned} r_{kt} &= \alpha_{kt} + \beta_{mt}^k r_{mt} + \beta_{SPt}^k r_{SPt} + \beta_{CPI}^k CPI_t + \beta_{MTGt}^k MTG_t + \beta_{SLPt}^k SLP_t + \\ &\quad \beta_{UNE_t}^k UNE_t + \beta_{POP_t}^k POP_t + \beta_{INC_t}^k INC_t + \beta_{GDP_t}^k GDP_t + e_{kt}, \end{aligned}$$

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  $2 \left[ 1 - \Pr \left\{ |\widehat{t}_{knt}| \leq t_{\frac{\alpha}{2}} [N-2] \right\} \right] \leq \alpha$ , 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 ( $\widehat{\rho}_{knt}^*$ )<sup>2</sup> significant at the 10% level using the  $t$ -square ratio test  $\widehat{t}_{knt}^2 = (\widehat{\rho}_{knt}^*)^2 \left[ 1 - (\widehat{\rho}_{knt}^*)^2 \right]^{-1} (N-2) \sim F[1, N-2]$ , 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  $\left[ 1 - \Pr \left\{ \widehat{t}_{knt}^2 \leq F_{\alpha} [1, N-2] \right\} \right] \leq \alpha$ , with  $\alpha = 0.10$ , and zero otherwise, or the mean of its market-wide averages. A “\*”, “•”, or “†” indicate that the  $t$ -square statistic  $\widehat{\rho}_t^* \left[ 1 - \widehat{\rho}_t^* \right]^{-1} (N-2)$  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	$\widehat{\rho}_t^{BASE*}$			$\widehat{\rho}_t^{OLS*}$						
	$\mu$	$\sigma$	$T\rho$	$F\rho^{*2}$	$\mu$	$\sigma$	$B\rho$	$R\rho$	$T\rho$	$F\rho^{*2}$
Nationwide	0.417 <sup>†</sup>	0.167	81%	83%	0.107 <sup>†</sup>	0.024	0.140	26%	29%	40%
Metropolitan Markets										
Los Angeles	0.507 <sup>†</sup>	0.181	87%	90%	0.133 <sup>†</sup>	0.063	0.288	26%	35%	49%
San Diego	0.466 <sup>†</sup>	0.217	81%	86%	0.104 <sup>†</sup>	0.064	0.148	22%	32%	46%
San Francisco	0.403 <sup>†</sup>	0.229	77%	81%	0.124 <sup>†</sup>	0.083	0.523	31%	31%	47%
Denver	0.459 <sup>†</sup>	0.201	86%	86%	0.104 <sup>†</sup>	0.073	0.128	23%	30%	40%
Washington	0.424 <sup>†</sup>	0.250	74%	79%	0.104 <sup>†</sup>	0.077	0.052	25%	28%	42%
Miami	0.429 <sup>†</sup>	0.225	78%	81%	0.107 <sup>†</sup>	0.072	-0.015	25%	31%	42%
Tampa	0.408 <sup>†</sup>	0.212	74%	74%	0.100 <sup>†</sup>	0.077	0.058	24%	25%	33%
Chicago	0.451 <sup>†</sup>	0.207	83%	87%	0.103 <sup>†</sup>	0.074	0.116	23%	27%	39%
Boston	0.452 <sup>†</sup>	0.214	84%	86%	0.112 <sup>†</sup>	0.088	-0.098	25%	28%	38%
Charlotte	0.266 <sup>†</sup>	0.151	75%	75%	0.077 <sup>†</sup>	0.040	0.099	29%	31%	36%
Las Vegas	0.330 <sup>†</sup>	0.228	73%	77%	0.148 <sup>†</sup>	0.117	0.309	45%	28%	43%
New York	0.497 <sup>†</sup>	0.222	85%	88%	0.108 <sup>†</sup>	0.061	0.116	22%	30%	43%
Cleveland	0.357 <sup>†</sup>	0.130	87%	87%	0.077 <sup>†</sup>	0.069	-0.065	22%	25%	32%
Portland	0.391 <sup>†</sup>	0.134	85%	86%	0.089 <sup>†</sup>	0.085	0.299	23%	25%	36%

Table 3. Break tests

This table reports OLS estimates (and their  $t$  statistics below) for all coefficients in Eq. (17), 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 - December 2008 (204 observations).  $F(\hat{\tau})$  is the corresponding Wald statistic (Eq. (19)), while  $\hat{\tau}^-$  and  $\hat{\tau}^+$  are the lower and upper bounds of the 95% confidence interval around  $\hat{\tau}$  (Eq. (20)).  $R_a^2$  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	1671.08†	Sep-98	Sep-98	Sep-98	89.36%	0.235† 7.41	-0.142 -0.46	0.307† 8.07	0.157 0.44
Metropolitan Markets									
Los Angeles	798.02†	Nov-99	Dec-99	Dec-99	81.34%	0.296† 19.52	0.316† 2.88	0.352† 12.34	-0.242 -1.27
San Diego	1036.02†	Aug-98	Sep-98	Sep-98	83.94%	0.138† 7.13	0.930† 5.08	0.523† 21.35	-1.267† -5.88
San Francisco	217.47†	May-98	Jun-98	Jun-98	64.67%	0.107† 3.73	0.848† 2.96	0.324† 8.68	-0.117 -0.37
Denver	212.77†	Mar-96	Apr-96	Apr-96	51.62%	0.179† 6.82	0.314* 1.92	0.304† 8.11	0.243 0.85
Washington	1924.44†	Aug-99	Aug-99	Aug-99	90.47%	0.237† 15.59	-0.808† -5.89	0.398† 20.68	0.814† 5.08
Miami	1877.46†	Jun-98	Jun-98	Jun-98	90.23%	-0.004 -0.23	1.314† 11.67	0.571† 29.05	-1.090† -7.54
Tampa	1107.32†	Sep-99	Sep-99	Sep-99	84.52%	0.258† 12.33	-0.699† -3.32	0.328† 13.62	0.675† 2.99
Chicago	1014.06†	Jul-98	Aug-98	Aug-98	83.50%	0.229† 13.45	-0.149 -1.02	0.342† 16.17	0.430• 2.47
Boston	283.83†	May-98	Jun-98	Jun-98	58.45%	0.221† 9.08	0.158 1.09	0.376† 11.56	-0.338 -1.50
Charlotte	47.93†	Feb-98	May-98	Jul-98	18.93%	0.216† 5.26	-0.245 -0.57	0.012 0.26	1.413† 2.72
Las Vegas	647.61†	Sep-99	Oct-99	Oct-99	78.33%	0.089† 5.33	0.201* 1.95	0.374† 15.55	0.094 0.71
New York	1138.46†	Dec-98	Dec-98	Dec-98	85.04%	0.240† 8.60	0.100 0.38	0.416† 13.23	-0.017 -0.06
Cleveland	93.42†	Nov-98	Jan-99	Feb-99	31.11%	0.269† 16.18	0.006 0.05	0.136† 5.57	0.194 0.76
Portland	135.31†	Apr-99	May-99	May-99	44.86%	0.265† 20.02	0.380† 3.25	0.187† 9.09	-0.241 -1.43

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 - December 2008 (204 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.

Panel A: $\hat{\rho}_t^{BASE*}$ and $\hat{\rho}_{kt}^{BASE*}$									
Index	Constant	$d_t^R$	$d_t^+$	$d_t^-$	$d_{SPt}^+$	$d_{SPt}^-$	$d_{kt}^+$	$d_{kt}^-$	$R_a^2$
Nationwide	0.284 <sup>†</sup> 6.38	0.134 <sup>†</sup> 4.87	0.160 <sup>†</sup> 3.36	0.048 0.97	-0.009 -0.26	0.066 <sup>†</sup> 3.12			17.57%
Metropolitan Markets									
Los Angeles	0.487 <sup>†</sup> 6.90	0.159 <sup>†</sup> 5.15	-0.158 <sup>†</sup> -2.84	0.111 <sup>•</sup> 2.36	-0.020 -0.66	0.051 <sup>•</sup> 1.98	0.225 <sup>†</sup> 5.54	-0.097* -1.86	38.36%
San Diego	0.344 <sup>†</sup> 4.64	0.123 <sup>†</sup> 3.89	-0.149* -1.89	0.112 1.65	-0.036 -0.87	0.068 <sup>•</sup> 2.57	0.335 <sup>†</sup> 7.54	0.008 0.12	31.26%
San Francisco	0.200 <sup>•</sup> 2.22	0.096* 1.69	0.084 0.89	0.155 <sup>•</sup> 2.29	-0.043 -0.88	-0.015 -0.34	0.160 <sup>†</sup> 2.63	0.119* 1.91	12.02%
Denver	0.344 <sup>†</sup> 6.96	0.019 0.54	0.281 <sup>†</sup> 5.69	-0.097 <sup>•</sup> -2.09	0.033 1.07	0.038* 1.70	-0.091 <sup>†</sup> -3.24	0.057 1.32	55.86%
Washington	0.153 <sup>•</sup> 2.52	0.142 <sup>†</sup> 3.63	0.038 0.52	0.135* 1.84	-0.041 -0.79	0.067 <sup>•</sup> 2.31	0.292 <sup>†</sup> 5.16	0.192 <sup>†</sup> 2.69	28.77%
Miami	0.156 <sup>•</sup> 2.39	0.074 <sup>†</sup> 2.69	0.130* 1.91	-0.099 -1.22	0.069* 1.68	0.057 <sup>†</sup> 2.77	0.198 <sup>†</sup> 4.56	0.443 <sup>†</sup> 5.93	38.19%
Tampa	0.157 <sup>•</sup> 2.49	0.112 <sup>†</sup> 3.72	0.111 1.60	0.044 0.64	-0.018 -0.43	0.052 <sup>•</sup> 2.05	0.187 <sup>†</sup> 4.44	0.292 <sup>†</sup> 4.49	33.04%
Chicago	0.197 <sup>•</sup> 2.41	0.061 <sup>•</sup> 2.46	0.165 <sup>•</sup> 2.52	0.009 0.12	0.051 1.15	0.055* 1.95	0.133 <sup>†</sup> 2.74	0.279 <sup>†</sup> 3.58	22.33%
Boston	0.262 <sup>†</sup> 5.67	0.115 <sup>†</sup> 4.35	0.274 <sup>†</sup> 5.71	-0.078* -1.85	-0.007 -0.17	0.063* 1.94	-0.006 -0.12	0.085* 1.77	39.46%
Charlotte	0.181 <sup>†</sup> 6.99	-0.036 -1.08	0.148 <sup>†</sup> 5.61	-0.036 -1.65	-0.016 -0.66	0.030 0.69	-0.003 -0.13	0.009 0.24	28.27%
Las Vegas	0.145 1.64	0.204 <sup>†</sup> 3.84	0.051 0.60	0.010 0.11	0.000 0.01	0.104 <sup>†</sup> 2.94	0.140 <sup>†</sup> 3.45	0.277 <sup>†</sup> 3.24	32.02%
New York	0.323 <sup>†</sup> 4.51	0.149 <sup>†</sup> 3.02	0.140 1.62	-0.037 -0.46	-0.031 -0.70	0.023 0.61	0.066 0.78	0.285 <sup>†</sup> 3.87	21.23%
Cleveland	0.260 <sup>†</sup> 9.82	0.024 0.64	0.138 <sup>†</sup> 4.80	-0.058 <sup>•</sup> -2.40	-0.015 -0.73	0.076 <sup>†</sup> 2.92	0.020 0.81	0.021 0.78	41.07%
Portland	0.293 <sup>†</sup> 5.25	0.096 <sup>†</sup> 3.05	0.104 <sup>•</sup> 2.34	-0.034 -0.73	-0.015 -0.55	0.021 0.64	0.035 0.94	0.029 0.53	19.72%

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	0.117 <sup>†</sup> 26.50	-0.001 -0.12	-0.016 <sup>†</sup> -3.44	0.000 -0.01	0.000 -0.02	0.001 0.14			6.99%
Metropolitan Markets									
Los Angeles	0.173 <sup>†</sup> 5.64	0.010 0.82	-0.050 <sup>•</sup> -1.99	0.037* 1.75	0.000 0.02	-0.008 -0.68	0.000 0.01	-0.044* -1.72	12.49%
San Diego	0.111 <sup>†</sup> 4.68	0.001 0.05	-0.024 -0.89	0.015 0.53	-0.008 -0.60	-0.010 -0.51	0.000 -0.02	0.037 1.07	20.96%
San Francisco	0.055 <sup>†</sup> 3.41	-0.059 <sup>•</sup> -2.16	0.048 <sup>•</sup> 2.19	0.044* 1.83	-0.015 -1.04	0.020 0.70	0.034* 1.83	0.072 <sup>†</sup> 2.87	8.17%
Denver	0.101 <sup>†</sup> 4.49	-0.011 -1.00	0.018 0.82	-0.021 -1.12	0.009 0.46	0.011 1.07	-0.007 -0.71	-0.003 -0.24	1.84%
Washington	0.096 <sup>†</sup> 6.15	-0.040 <sup>•</sup> -2.44	0.048 <sup>†</sup> 2.73	-0.028* -1.69	-0.006 -0.51	0.042* 1.83	-0.037 <sup>•</sup> -2.04	0.032* 1.96	7.59%
Miami	0.114 <sup>†</sup> 6.36	-0.038 <sup>•</sup> -2.47	0.006 0.33	-0.030 -1.56	0.026 <sup>•</sup> 1.98	-0.016 -1.07	-0.016 -1.07	0.034 1.60	5.39%
Tampa	0.099 <sup>†</sup> 5.99	-0.001 -0.05	0.001 0.07	-0.024 -1.33	-0.012 -1.09	-0.025 -1.31	-0.009 -1.07	0.096 <sup>†</sup> 4.85	13.55%
Chicago	0.130 <sup>†</sup> 3.86	-0.028 <sup>•</sup> -2.03	-0.016 -0.85	-0.023 -1.09	-0.017 -1.18	0.003 0.13	-0.008 -0.38	0.011 0.65	-1.27%
Boston	0.156 <sup>†</sup> 6.71	-0.036 -1.60	-0.022 -0.82	-0.016 -0.67	0.009 0.48	-0.003 -0.16	-0.030 -1.19	-0.028 -1.35	1.94%
Charlotte	0.088 <sup>†</sup> 7.34	-0.013 -0.98	-0.016 -1.50	0.001 0.06	-0.005 -0.71	-0.004 -0.26	0.003 0.32	0.011 0.69	0.66%
Las Vegas	0.142 <sup>†</sup> 5.49	0.091 <sup>†</sup> 2.76	-0.025 -1.00	-0.005 -0.15	0.011 0.59	-0.028 -1.16	0.010 0.51	0.050* 1.71	8.70%
New York	0.111 <sup>†</sup> 10.75	-0.030* -1.67	-0.033 <sup>•</sup> -2.19	-0.004 -0.21	-0.019 <sup>•</sup> -2.22	-0.011 -0.98	0.030 <sup>•</sup> 2.44	0.067 <sup>•</sup> 2.46	8.80%
Cleveland	0.126 <sup>†</sup> 4.72	-0.027 <sup>†</sup> -2.63	-0.051* -1.84	-0.027 -0.98	0.026* 1.66	0.003 0.34	-0.017 -1.38	-0.033* -1.77	9.76%
Portland	0.092 <sup>†</sup> 3.19	0.048 <sup>•</sup> 2.16	-0.026 -0.95	-0.025 -0.79	-0.010 -0.69	0.048 1.64	0.018 1.47	0.028 1.43	5.45%

Figure 1. Systematic sources of risk in 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  $r_{rkt,fit}^2 = \frac{t_{rkt,fit}^2}{t_{rkt,fit}^2 + N - K}$ , where  $f_{it}$  is the corresponding systematic source of risk and  $t_{rkt,fit}^2$  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  $k = 1, \dots, K$ .

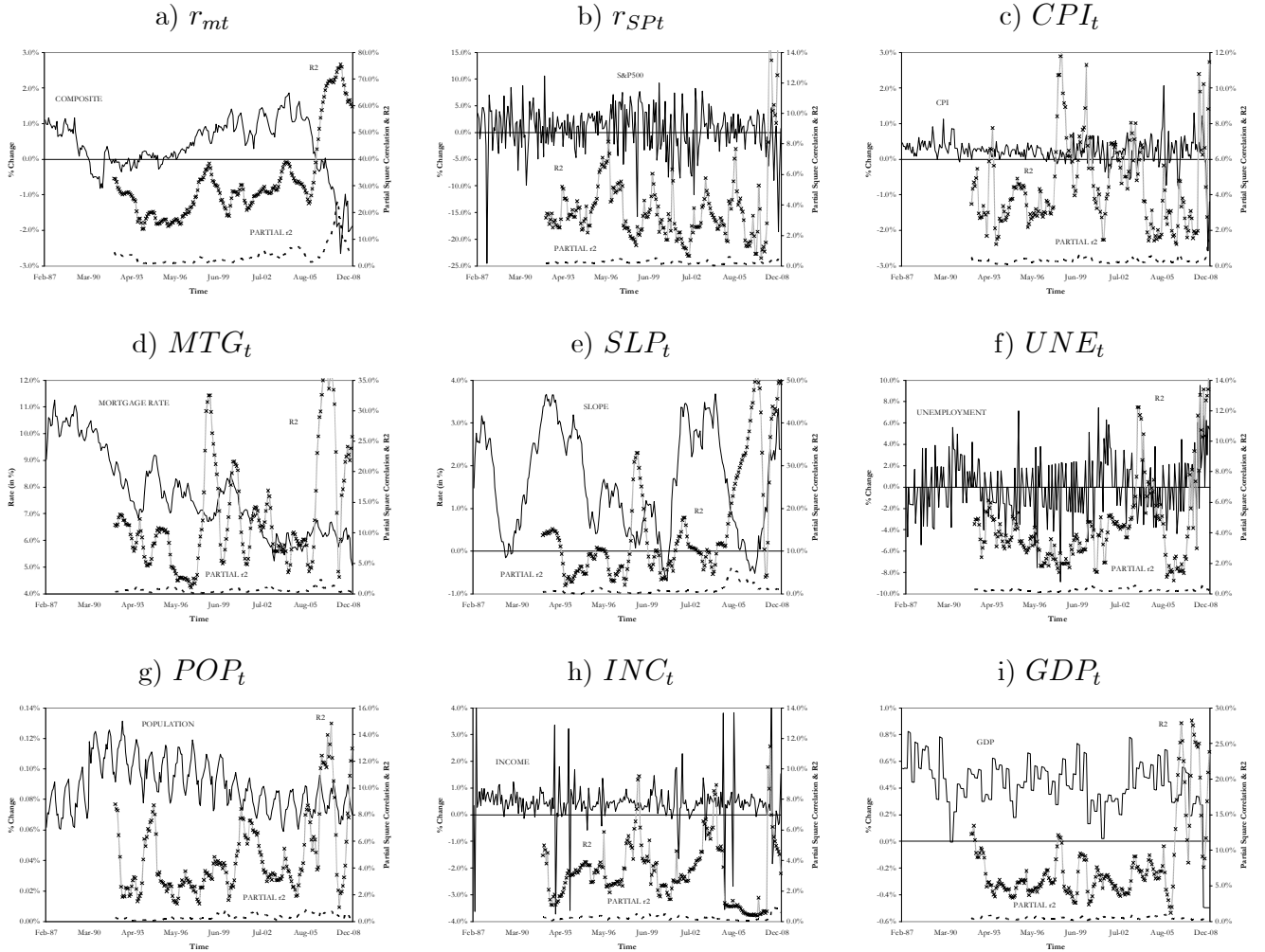
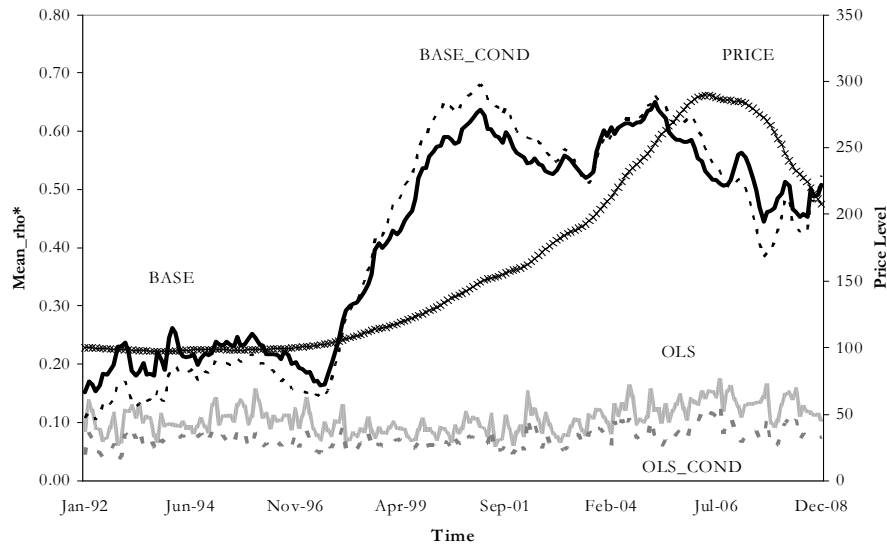


Figure 2. Mean raw and excess correlation: OLS procedure

Figure 2a plots (on the left axis) the time series of mean excess square unconditional correlations  $\widehat{\rho}_t^{OLS*}$  of Eq. (12), a benchmark measure of square unconditional correlation,  $\widehat{\rho}_t^{BASE*}$  of Eq. (11), computed using raw CSI returns  $r_{kt}$  (instead of estimated residuals  $\widehat{e}_{kt}^{OLS}$ ) in Eqs. (5) to (10), their equivalents,  $\widehat{\rho}_t^{OLS}$  and  $\widehat{\rho}_t^{BASE}$ , 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)  $\widehat{\rho}_t^{BASE*}$ ,  $\widehat{\rho}_t^{BASE}$ , and  $\widehat{\rho}_t^{OLS*}$  (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)

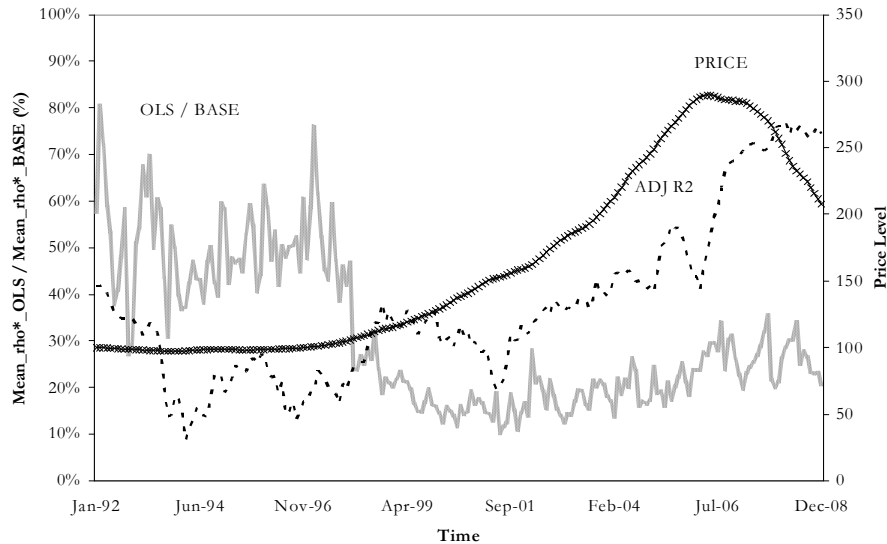


Figure 3. Mean raw and excess correlation for metropolitan markets: OLS procedure

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), and (on the right axis) the corresponding CSI index with base 100 in December 1991.

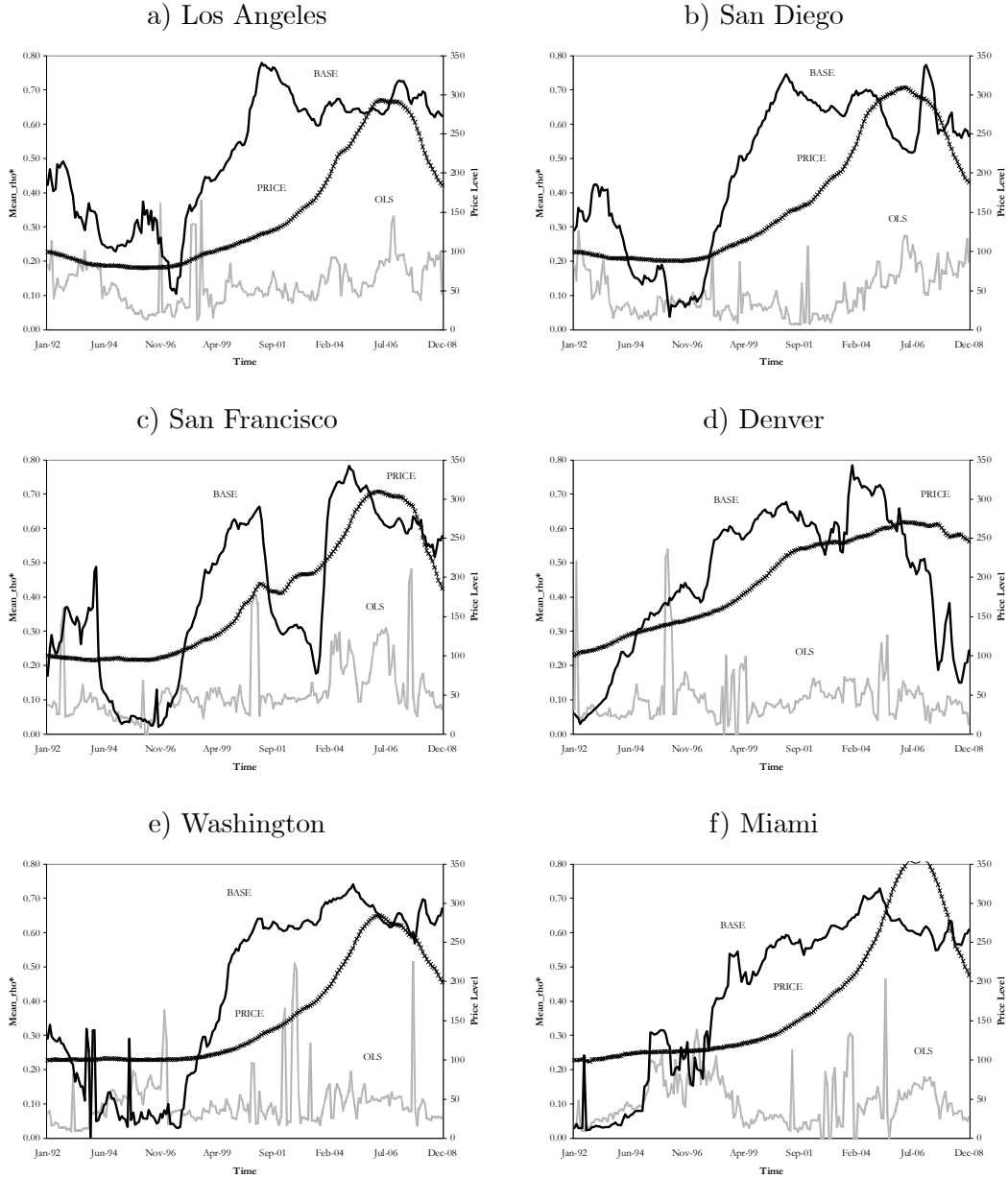
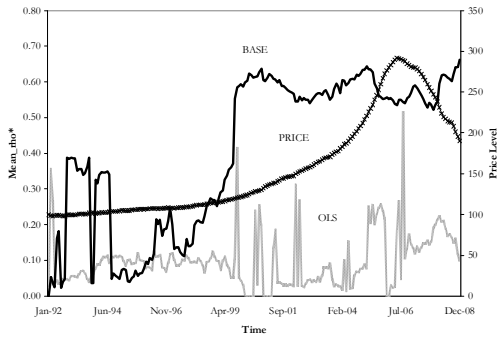
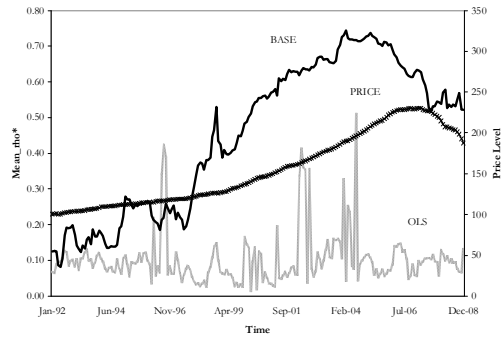


Figure 3 (Continued).

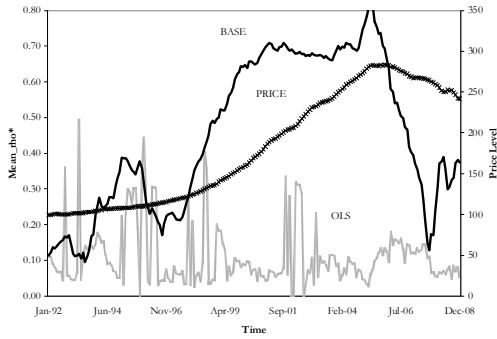
g) Tampa



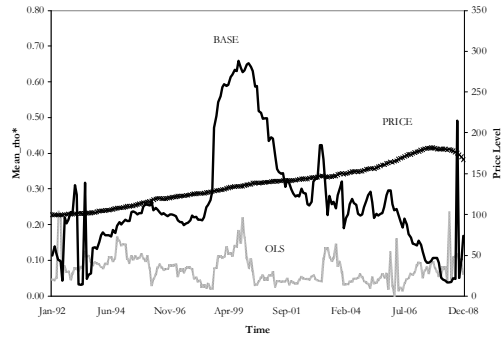
h) Chicago



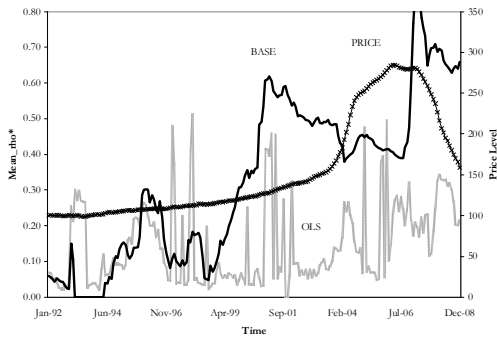
i) Boston



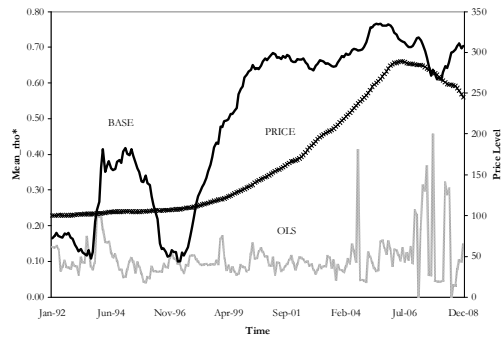
j) Charlotte



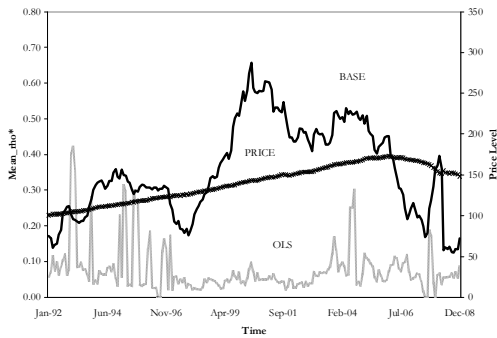
k) Las Vegas



l) New York



m) Cleveland



n) Portland

