

Research Article

Tetsushi Horie and Yohei Yamamoto*

Identifying Common and Idiosyncratic Explosive Behaviors in the Large Dimensional Factor Model with an Application to U.S. State-Level House Prices

<https://doi.org/10.1515/jem-2022-0017>

Received May 18, 2022; accepted February 26, 2023; published online March 14, 2023

Abstract: This study applies the date-stamping methodologies for explosive behaviors proposed in the seminal work of Phillips, P. C. B., and J. Yu. (2011. “Dating the Timeline of Financial Bubbles during the Subprime Crisis.” *Quantitative Economics* 2 (3): 455–91), Phillips, P. C. B., S. Shi, and J. Yu. (2015a. “Testing for Multiple Bubbles: Historical Episodes of Exuberance and Collapse in the S&P 500.” *International Economic Review* 56 (4): 1043–78), and Phillips, P. C. B., S. Shi, and J. Yu. (2015b. “Testing for Multiple Bubbles: Limit Theory of Real Time Detectors.” *International Economic Review* 56 (4): 1079–134) to a large dimensional factor model. To this end, we compare two methods of identifying common and idiosyncratic components: the Panel Analysis of Nonstationarity in Idiosyncratic and Common Components (PANIC) method by Bai, J., and S. Ng. (2004. “A Panic Attack on Unit Roots and Cointegration.” *Econometrica* 72 (4): 1127–77) and the Cross-Sectional regression (CS) method by Yamamoto, Y., and T. Horie. (2022. “A Cross-Sectional Method for Right-Tailed PANIC Tests under a Moderately Local to Unity Framework.” *Econometric Theory* (forthcoming)). We show that, when the explosive behavior lies only in the common component, the origination and termination dates are precisely estimated by either method. However, when the explosive behaviors exist in idiosyncratic components, only the CS method can detect them. We apply our method to the U.S. state-level real house price indices. We find that the 2000s boom was driven by not only the national bubble factors but also local components, while the 2010s onward expansion is dominated by the effect of national components.

Keywords: explosive behaviors; date-stamping; large dimensional factor model; state-level house price index

JEL Classification: C33; G12; F31

1 Introduction

Testing for speculative bubbles in asset prices is a long-standing problem for which numerous econometric techniques have been developed. The most recent studies include the seminal work of Phillips et al. (2011, 2015a,

Yamamoto acknowledges the financial support from MEXT Grants-in-Aid for Scientific Research No 20H00073. We are grateful to the editor Zhongjun Qu, an associate editor and two anonymous referees for their constructive suggestions. We also thank the seminar participants at SETA2022 for their useful comments. All remaining errors are our own.

*Corresponding author: Yohei Yamamoto, Graduate School of Economics, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo 186-8601, Japan, E-mail: yohei.yamamoto@r.hit-u.ac.jp

Tetsushi Horie, Department of Economics, Toyo University, 5-28-20 Hakusan, Bunkyo, Tokyo 112-8606, Japan, E-mail: horie128@toyo.jp

2015b) which linked speculative bubbles to explosive behaviors of asset prices.¹ Their strategy is to fit a univariate autoregressive model and sequentially test whether the root is greater than unity. The origination and termination dates of bubbles are also identified as the intersections of tests and critical values. Motivated by these studies, we explicitly account for the speculative bubbles that prevailed in global and individual markets during the period of exuberance. The goal of this study is to provide a tool to formally analyze whether these bubbles are an economy-wide phenomenon or market-specific events.

A standard practice in the empirical finance literature is to assume a common factor structure in the panel data of asset prices. A short list of major works includes Fama and French (1993), Litterman and Sheinkman (1991), and Ang and Piazzesi (2003), who examined stock and bond prices. There is a recent upsurge of the literature which applies a factor model to U.S. house prices and discusses whether bubbles in the 2000s are a national phenomenon or a collection of local behaviors, conjecturing that bubbles in the former may be attributed to monetary policies, financial integration and macroeconomic consequences of the COVID-19 etc., while the latter more to the supply and demand in the local markets. For example, Del Negro and Otrok (2007) used the state-level HPI from 1986 to 2005 and found that they were driven by local components, although the increase in 2001–2005 was considered a national phenomenon. Hu and Oxley (2018) used the state-level HPI from 1975 to 2014 and concluded that the bubble in the 2000s was widespread but not a national phenomenon. More recently, Landier et al. (2017) and Choi (2019) documented the sharp increase of comovement in HPI. Krugman (2022) noted that the surge in house prices after 2021 was not accompanied by a construction boom, thus suggesting a possibility of nation-wide insufficient construction caused by malfunctioning in the supply-chain of building materials. Despite these prolific applications, when it comes to econometric methodology, to our knowledge, there is no existing study that can successfully incorporate common factor structure and explosive behaviors to decompose the bubbles in the common and idiosyncratic components.

In this study, we adopt the framework of principal component estimation of the large dimensional common factor model developed by Stock and Watson (2002), Bai (2003), and Bai and Ng (2002, 2004). Specifically, Bai and Ng (2004) proposed the Panel Analysis of Nonstationarity in Idiosyncratic and Common Components (PANIC) framework to test for the unit root against stationarity in the common and idiosyncratic components of the factor model. They showed that the standard augmented Dickey–Fuller (ADF) tests applied to the estimated common and idiosyncratic components have good size and power as long as the left-tailed versions are considered. However, the bubble detection requires a test which certainly rejects when the process has a root greater than unity, that is, the right-tailed unit root tests. In this regard, Yamamoto and Horie (2022) recently discovered that the ADF tests in the PANIC framework have very different size and power properties when the right-tailed versions are considered as follows. First, the test for the common component suffers from serious size distortions when idiosyncratic components are explosive. Second, the test for an idiosyncratic component exhibits the nonmonotonic power problem, that is, the power can go down to zero when the idiosyncratic component is moderately or strongly explosive. If one attempts to apply the PANIC framework to the bubble testing, size distortions may result in overdetecting nonexistent bubbles and power losses overlook the bubbles present.

To address these problems, Yamamoto and Horie (2022) proposed a method based on cross-sectional (CS) regressions to disentangle the common and idiosyncratic components with a fixed sample. However, little is known about whether or how the CS method can be applied to a recursive implementation of the date-stamping bubbles. As a recent similar attempt, Chen et al. (2022) investigated the theoretical property of date-stamping methods applied to the common component identified by the principal component using the level data, assuming no explosive behavior in idiosyncratic components. Apparently, their study rules out an important empirical fact about local bubbles which we will focus in this study.

More specifically, we apply the date-stamping methodologies for the origination and termination of bubbles based on the sequential ADF tests proposed in the seminal work of Phillips et al. (2011, 2015a, 2015b) to the common and idiosyncratic components of the large dimensional factor model. To this end, we compare two methods

¹ Phillips et al. (2011) and Phillips and Yu (2011) developed methods for a single bubble, and Phillips et al. (2015a, 2015b) modified these to account for multiple bubbles. For other testing methods, see Gürkaynak (2008) and Homm and Breitung (2012).

of identifying these components: PANIC and CS. Our Monte Carlo simulation shows that when bubbles lie only in the common component, these dates are precisely estimated by either the PANIC or CS method. However, when bubbles exist in idiosyncratic components, the PANIC method loses its power of detection and suffers from spurious bubbles in the common component. More importantly, these problems are largely resolved by using the CS method, although some tendency of overdetection remains when idiosyncratic components are strongly explosive.

We apply our identification method to the U.S. state-level real house price indices (HPI) over the sample period from 1991Q1 to 2021Q4. In this period, the U.S. experienced two unprecedented expansions in the housing market; hence, the 2000s boom and the expansion in the 2010s onward are compared. The CS method identifies five national factors among which two are considered bubble factors and the rest are non-bubble factors. We find that the 2000s boom was led by the two bubble factors while both bubble and non-bubble factors increased from the late 2010s onward. We also find that many states experienced local bubbles in the 2000s, while much fewer states did so in the 2010s onward. Hence, we conclude that the 2000s boom was driven by not only the national bubble factors but also a set of local bubbles, while expansion since the 2010s is dominated by the effect of national component.

The remainder of this paper is organized as follows. In Section 2, we introduce the model. In Section 3, we explain our identification strategy which includes the date-stamping methodology of Phillips et al.'s (2011, 2015a, 2015b) and two approaches to disentangle the common and idiosyncratic components. In Section 4, we implement Monte Carlo simulations to assess the empirical probability of correctly or incorrectly detecting explosive behaviors. We also look into the accuracy of the origination and termination dates. Section 5 provides the empirical analysis using the U.S. state-level real HPI and Section 6 presents some concluding remarks.

2 Model

We assume that a panel data set of asset prices follow the factor model:

$$X_{i,t} = \mu_i + \lambda_i' F_t + U_{i,t}, \quad \text{for } i = 1, \dots, N \text{ and } t = 1, \dots, T, \quad (1)$$

where $X_{i,t}$ is a scalar of the observed asset price with subscripts i and t indicating cross-section and time, respectively. μ_i is a cross-section specific intercept, F_t and λ_i are $r \times 1$ vectors of the common factors and factor loadings, and $U_{i,t}$ is a scalar idiosyncratic component. The cross-section and time dimensions N and T are large, while the number of factors r is small indicating that the large dimensional panel data is driven by a small number of factors. For the moment, we consider $r = 1$ for notational simplicity. We assume that the factor and the idiosyncratic components follow the first-order autoregressive (AR) processes: $F_t = \alpha F_{t-1} + e_t$ and $U_{i,t} = \rho_i U_{i,t-1} + z_{i,t}$, where α and ρ_i are the AR coefficients and e_t and $z_{i,t}$ are stationary disturbances, respectively. We assume e_t and $z_{i,t}$ for any i are mutually independent for all leads and lags. In what follows, we call F_t the common component.²

The fact that the asset price $X_{i,t}$ follows a random walk is consistent with the efficient market hypothesis (Fama 1970). Therefore, during normal times, both F_t and $U_{i,t}$ follow a random walk ($\alpha = \rho_i = 1$). When asset price $X_{i,t}$ is subject to a speculative bubble, it exhibits explosive behavior, such that $\alpha > 1$ and/or $\rho_i > 1$. Thus, we allow for regimes that switch between the random walk and explosive process in the common and idiosyncratic components. Let the common and idiosyncratic components have m and m_i bubbles. The j th bubble in the common component originates at $T_e^{F,j}$ and terminates at $T_f^{F,j}$ with the AR coefficient α_j , and the j th bubble in the i th idiosyncratic component originates at $T_e^{U_{i,j}}$ and terminates at $T_f^{U_{i,j}}$ with the AR coefficient $\rho_{i,j}$. Let I_A be the indicator function that takes one when an arbitrary condition A is true and zero otherwise. Then, the data generating processes are

² In the literature on the factor model such as Bai and Ng (2002) and Bai (2003), the common component is defined as $\lambda_i' F_t$, however, when we focus on the time-series property and there is no confusion, we use the factors and the common component interchangeably.

$$F_t = (F_{t-1} + e_t)I_{t \in [1, T_e^{F,1}]} + \sum_{j=1}^m \left\{ (\alpha_j F_{t-1} + e_t)I_{t \in [T_e^{F,j}+1, T_f^{F,j}]} + \left(\sum_{k=T_f^{F,j}+1}^t e_k + F_{T_e^{F,j}} + F_j^* \right) I_{t \in [T_f^{F,j}+1, T_e^{F,j+1}]} \right\}, \quad (2)$$

where F_j^* for $j = 1, \dots, m$ are bounded random variables and $T_e^{F,m+1} = T$. Also,

$$U_{i,t} = (U_{i,t-1} + z_{i,t})I_{t \in [1, T_e^{U,i,1}]} + \sum_{j=1}^{m_i} \left\{ (\rho_{i,j} U_{i,t-1} + z_{i,t})I_{t \in [T_e^{U,i,j}+1, T_f^{U,i,j}]} + \left(\sum_{k=T_f^{U,i,j}+1}^t z_{i,k} + U_{i,T_e^{U,i,j}} + U_{i,j}^* \right) I_{t \in [T_f^{U,i,j}+1, T_e^{U,i,j+1}]} \right\}, \quad (3)$$

where $U_{i,j}^*$ for $j = 1, \dots, m_i$ are bounded random variables and $T_e^{U,i,m_i+1} = T$ for all i . Since any bubbles in the common or in the i th idiosyncratic component induce an explosive behavior in X_i , we let $h_j = \max_i T_f^{X_i,j} - \min_i T_e^{X_i,j}$ be the lengths of the j th explosive period. If there is no chance of confusion, we simply denote the length of bubble period by h .

When $r > 1$, each factor may have a different AR coefficient and one can consider the common factors series by series to investigate whether the estimated individual factor is explosive. If one is interested in the entire factor space, it would be adequate to realistically assume that any linear combination of explosive factors is explosive. This contrasts with the case of investigating the stationarity of the common factor space when individual common factors may be integrated of order one, but under the cointegration relationship, such individual factors do not necessarily imply a random walk for the common factor space.

We consider the AR coefficients in the explosive period in the forms of

$$\alpha_j = 1 + \frac{c_j}{k_h} \text{ and } \rho_{i,j} = 1 + \frac{c_{i,j}}{k_h} \quad (4)$$

where c_j and $c_{i,j} \geq 0$ for all i and j are the localizing coefficients. We follow Phillips and Magdalinos (2007) to assume $k_h \rightarrow \infty$ and $k_h/h \rightarrow 0$ as $h \rightarrow \infty$ to justify the use of the methods of Phillips et al. (2011, 2015a, 2015b) and Yamamoto and Horie (2022), although further theoretical investigations based on this formulation are beyond the scope of this study thus not pursued.

3 Identification Strategy

3.1 Date-Stamping Methodology

When a single bubble is concerned in a univariate time series Y_t for $t = 1, \dots, T$, Phillips et al. (2011) proposed the following methodology for date-stamping the origination and termination of the bubble. Let the ADF test using the subsample $t = 1, \dots, T_0$, where $1 < T_0 < T$ be $ADF_Y^{[1, T_0]}$; that is, a t -test statistic for the coefficient θ in the regression $Y_t - Y_{t-1} = \mu_Y + \theta Y_{t-1} + \text{error}$ for the null hypothesis of $H_0: \theta = 0$ versus an alternative hypothesis of $H_1: \theta > 0$. We construct a sequence of ADF test statistics, starting from sample $t \in [1, T_0]$ with T_0 being the minimal amount of data, and extending forward $[1, T_0 + 1]$, $[1, T_0 + 2]$, \dots , $[1, T]$. We identify the bubble when the test statistic exceeds the critical value (cv_t) which slowly diverges to infinity as $T \rightarrow \infty$ for certain periods, so that the origination date is estimated by:

$$\hat{T}_e^Y = \min_{t \in [T_0, T]} \left\{ t: ADF_Y^{[1, t]} > cv_t \right\}, \quad (5)$$

and the termination date of the bubble is estimated by

$$\hat{T}_f^Y = \min_{t \in [\hat{T}_e^Y + \lfloor \delta \log(T) \rfloor, T]} \left\{ t: ADF_Y^{[1, t]} < cv_t \right\}, \quad (6)$$

where $\lfloor \delta \log(T) \rfloor$ is the minimum interval between the origination and termination dates with $\delta > 0$ an arbitrary constant.

If multiple bubbles are involved, Phillips et al. (2015a, 2015b) proposed constructing the sequence of ADF tests backwardly. This device prevents the well-known power loss in detecting bubbles after the first one due to the inclusion of earlier bubbles in the sample. In this case, define the backward supremum ADF (BSADF) test as

$$BSADF_Y^t = \max_{s \in [1, t-T_0+1]} ADF_Y^{[s, t]} \quad \text{for } t = T_0, \dots, T.$$

For the purpose of date-stamping, let $\hat{T}_e^{Y, j}$ and $\hat{T}_f^{Y, j}$ be the estimator for the origination and termination dates of the j th bubble. When we consider $j = 1, 2$, these dates are obtained by

$$\hat{T}_e^{Y, 1} = \min_{t \in [T_0, T]} \{t: BSADF_Y^t > cv_t\}, \quad (7)$$

$$\hat{T}_f^{Y, 1} = \min_{t \in [\hat{T}_e^1 + \lfloor \delta \log(T) \rfloor, T]} \{t: BSADF_Y^t < cv_t\}, \quad (8)$$

$$\hat{T}_e^{Y, 2} = \min_{t \in [\hat{T}_f^1 + \lfloor \delta \log(T) \rfloor, T]} \{t: BSADF_Y^t > cv_t\}, \quad (9)$$

$$\hat{T}_f^{Y, 2} = \min_{t \in [\hat{T}_e^2 + \lfloor \delta \log(T) \rfloor, T]} \{t: BSADF_Y^t < cv_t\}. \quad (10)$$

In practice, we follow Phillips et al. (2011, 2015a, 2015b) and set the critical values cv_t which slowly diverge at a rate of double logarithms. We apply these methods to the estimated common and idiosyncratic components identified by the following methods.

3.2 Identifying the Common and Idiosyncratic Components

3.2.1 The PANIC Method

The first approach of identifying the common and idiosyncratic components is the PANIC method proposed by Bai and Ng (2004). The main concept is simple. We use the principal component method of the first-differenced data ($x_{i,t} = X_{i,t} - X_{i,t-1}$) for the entire sample to estimate the first-differenced common component ($f_t = F_t - F_{t-1}$) and the idiosyncratic components ($u_{i,t} = U_{i,t} - U_{i,t-1}$). Then, the levels of the common and idiosyncratic components are recovered by accumulating the differences. The algorithm is described as follows.

1. Take first-differences of the observed data $x_{i,t} = X_{i,t} - X_{i,t-1}$ for $t = 2, \dots, T$.
2. Obtain the principal component estimate of the common component (\hat{f}_t , $t = 2, \dots, T$) out of $x_{i,t}$ as follows. Let x be a $(T-1) \times N$ matrix, with the (t, i) th element being $x_{i,t+1}$ and $\hat{f} = [\hat{f}_2, \dots, \hat{f}_T]'$ be a $(T-1) \times r$ matrix of an estimate for $f = [f_2, \dots, f_T]'$. Then, \hat{f} is the $\sqrt{T-1}$ times the eigenvectors of xx' corresponding to the r largest eigenvalues normalized by $\hat{f}'\hat{f}/(T-1) = I_r$. The factor loadings and the first-differenced idiosyncratic components are estimated by $\hat{\lambda} = x'\hat{f}/(T-1)$, where $\hat{\lambda} = [\hat{\lambda}_1, \dots, \hat{\lambda}_N]'$ is a $N \times r$ matrix, and $\hat{u}_{i,t} = x_{i,t} - \hat{\lambda}_i'\hat{f}_t$ is for $i = 1, \dots, N$ and $t = 2, \dots, T$, respectively.
3. The levels of the common and idiosyncratic components are obtained by $\hat{F}_t = \sum_{s=2}^t \hat{f}_s$ and $\hat{U}_{i,t} = \sum_{s=2}^t \hat{u}_{i,s}$, for $i = 1, \dots, N$ and $t = 2, \dots, T$, respectively.

Bai and Ng (2004) considered the left-tailed ADF test for the common and idiosyncratic components for the null hypothesis of unit root against an alternative hypothesis of stationarity. They showed that tests for either component have the same asymptotic distribution as $N, T \rightarrow \infty$ and good size and power in finite samples. However, the property of the right-tailed version of the ADF test was not investigated. More recently, Chen et al. (2022) investigated the theoretical property of date-stamping methodology of explosive behavior in the common factors using the principal components of the level data for the entire sample.³ However, they ruled out explosive

³ Precisely speaking, they propose estimating the factor loadings first as the principal components and subsequently estimating factors as the coefficients by the least squares method.

behaviors in the idiosyncratic components in their model. It is interesting to look into the property of their estimators in the presence of explosive behaviors in idiosyncratic components, however, this is beyond the scope of this study.

3.2.2 The CS Method

Yamamoto and Horie (2022) thoroughly investigated the right-tailed version of the ADF tests under the PANIC framework. They showed that the test for the common component suffers from serious size distortions when idiosyncratic components are explosive. They also showed that the test for idiosyncratic components exhibits the nonmonotonic power problem such that the power can go down to zero when the idiosyncratic component is moderately or strongly explosive, because they may be misidentified as the common component. To address these problems, they proposed a method based on cross-sectional regressions to disentangle the common and idiosyncratic components. The algorithm is described as follows.

1. Divide the entire sample period ($t = 1, \dots, T$) into two: the training sample ($t = 1, \dots, T_1$), and the testing sample ($t = T_1 + 1, \dots, T$). The training sample must be selected such that no explosive series or only weakly explosive series are included.
2. Use the first-differenced data $x_{i,t}$ in the training sample to estimate the first-differences of the common component and factor loadings by the principal component method. The number of factors r can be estimated with this sample by using, say, the information criteria proposed by Bai and Ng (2002). Denote the first-differenced common component, the level estimates, and the factor loadings by \hat{f}_t and $\hat{F}_t = \sum_{s=2}^t \hat{f}_s$ for $t = 2, \dots, T_1$ and $\hat{\lambda}_i^*$ for $i = 1, \dots, N$. The first differences of idiosyncratic components are obtained as regression residuals $\hat{u}_{i,t} = x_{i,t} - \hat{\lambda}_i^{*'} \hat{f}_t$ and the level estimates are obtained by $\hat{U}_{i,t} = \sum_{s=2}^t \hat{u}_{i,s}$ for $t = 2, \dots, T_1$ in the training sample.
3. In the testing sample for $t = T_1 + 1, \dots, T$, the common factor is estimated by the cross-sectional regression with $X_{i,t}$ and $\hat{\lambda}_i^*$ as the regressand and regressors for $i = 1, \dots, N$

$$\hat{F}_t = \left(\sum_{i=1}^N \hat{\lambda}_i^* \hat{\lambda}_i^{*'} \right)^{-1} \left(\sum_{i=1}^N \hat{\lambda}_i^* X_{i,t} \right),$$

and idiosyncratic components are estimated by $\hat{U}_{i,t} = X_{i,t} - \hat{\lambda}_i^{*'} \hat{F}_t$. Their first-differences are $\hat{f}_t = \hat{F}_t - \hat{F}_{t-1}$ and $\hat{u}_{i,t} = \hat{U}_{i,t} - \hat{U}_{i,t-1}$, respectively, for $t = T_1 + 1, \dots, T$.

In practice, this method requires us to select the sample demarcation point ($t = T_1$) in Step 1. Ideally, the training sample should be the longest possible, as far as the factor loading estimates $\hat{\lambda}_i^*$ are not contaminated by the inclusion of explosive series in the sample. Hence, we wish to select the sample $[1, T_1]$ in which no or only weakly explosive series are included in the data $X_{i,t}$. A simple scheme is to apply the aforementioned date-stamping method to the cross-sectional average $\bar{X}_t = N^{-1} \sum_{i=1}^N X_{i,t}$ for $t = 1, \dots, T$ and use the estimated origination date as T_1 . For later use, this is called CS1. This criterion would be justified by the fact that, when some series start to be explosive, the average (\bar{X}_t) also exhibits explosive behavior. A more direct criterion is given by the date-stamping method applied to the estimated first common component via the full-sample PANIC method, called CS2. This is because the factor loading estimate will not be contaminated as long as the factor estimate is not contaminated. The first factor estimate will capture any explosive behavior in sample. We will use both criteria in the Monte Carlo study and the latter criterion as a preferred one in the empirical section.

4 Monte Carlo Study

In this section, we conduct a Monte Carlo study to investigate the empirical probability of detecting explosive behaviors by the proposed procedures. We also assess the accuracy of the origination and termination dates when the explosive behavior is present and detected. Throughout this section, we generate data using models

(1), (2), (3), and (4), where $r = 1$ and $\mu_i = 0$ for all i . A case of multiple bubbles follows the benchmark case of single bubbles.

We first investigate the case of single bubbles with $m = m_i = 1$ for all i . To simplify notation, we denote $c_1 = c$ and $c_{i,1} = c_i$ for all i . We set the following two experiments. In the first experiment, we generate data with a single bubble only in the common component, such that $c > 0$ and $c_i = 0$ for all i . In the second experiment, the data contains single bubbles in all idiosyncratic components but not in the common component so that $c = 0$ and $c_i > 0$ for all i . In both experiments, we identify the common and idiosyncratic components and implement the date-stamping method (5) and (6) for both components. To report the results of idiosyncratic components, we particularly select the first cross-section unit. This choice loses no generality as the model is symmetric for cross-section units.

Each Monte Carlo replication takes the following steps. If the sequence of ADF tests exceeds the critical value for more than four consecutive periods, the first date when the test exceeds the critical value is recorded as the origination date: \hat{T}_e^F for the common component and \hat{T}_e^{Ui} for the idiosyncratic components. This means that we choose δ such that $\lfloor \delta \log(T) \rfloor = 4$. If the origination is detected, the first date when the test falls below the critical value for four consecutive periods after \hat{T}_e^F is regarded as the termination dates \hat{T}_f^F for the common component and \hat{T}_f^{Ui} for the idiosyncratic components. The critical values are set at $cv_t = \log(\log(t))/1.5$ following Phillips et al.'s (2011) recommendation.⁴ The scale constant 1.5 is chosen such that the false detection rate becomes approximately 5% in our setting by a separate simulation. Based on 5000 replications, we compute the empirical probability that the bubble is not detected falsely when there is one present in the model. We also compute the empirical probability that the bubble is incorrectly detected when there is none in the model. Although the former corresponds to Type I error and the latter corresponds to Type II error in hypothesis testing, we collectively call them the error rates for our purpose. Given that the bubble is correctly detected, we compute $\left(\frac{\hat{T}_l^F - T_l^F}{T}\right)$ and $\left(\frac{\hat{T}_l^{Ui} - T_l^{Ui}}{T}\right)$ for $l = e, f$ in that replication. After the 5000 replications, their average across replications is regarded as the bias. We also compute $\left(\frac{\hat{T}_l^F - T_l^F}{T}\right)^2$ and $\left(\frac{\hat{T}_l^{Ui} - T_l^{Ui}}{T}\right)^2$ for $l = e, f$ in each replication when the bubble is correctly detected and their averages across replications are called the mean squared error (MSE). We present the case of $N = 100$, $T = 200$, $T_e^F = T_e^{Ui} = 40$, and $T_f^F = T_f^{Ui} = 80$ so that the duration of the bubble is $h = 40$, although reasonable variations regarding the model setting do not affect our qualitative results. We also computed the same set of results with $N = 50$ and $N = 200$ to assess the effects of cross-sectional dimension. The minimal amount of data is set at $T_0 = \lfloor r_0 T \rfloor$, where $r_0 = 0.01 + 1.8/\sqrt{T}$, following Phillips et al. (2015a). Because we are interested in how the results change as the AR coefficient varies, a set of values for c and $c_i = [0.2, 0.4, \dots, 2.0]$ is considered in the AR coefficients (4) with $k_h = h^\kappa$ and $\kappa = 0.85$. We compared the error rates for the common and idiosyncratic components, the bias and the MSE for the PANIC and CS methods. We consider two CS methods with different selecting criteria for the sample demarcation point as we explained in the previous section. We call it CS1 when we use \bar{X}_t and CS2 when we use \hat{F}_t to select the end of the training sample (T_1). We also compute the error rates, the bias, and the MSE when the true common and idiosyncratic components are used and these are labeled “Observed”. This serves as a reference because it is free from the effects of identification between the common and idiosyncratic components.

Figure 1a shows the results pertaining to the origination date for the case of explosive behavior in the common component but not in idiosyncratic components. The error rate for the common component is higher when c is smaller, however, it steadily declines as c increases. This corresponds to the standard power curve of the ADF test against the explosive alternative hypothesis. The error rate for idiosyncratic components remains low in any methods, however, we observe some tendency for overdetection of CS1 and CS2 when the common component is strongly explosive. This reflects the size distortions of the CS based test for idiosyncratic components

⁴ Phillips et al. (2011) recommended $cv_t = \log(\log(t))/100$ as it gives values close to the 5% critical value of the right-tailed ADF test. However, these values are considerably lower than the supremum ADF statistics reported in Phillips et al. (2015a); hence, we adjusted the constant such that the sequence of ADF tests exceeds for more than four consecutive periods approximately 5% of all replications of random walk sequence.

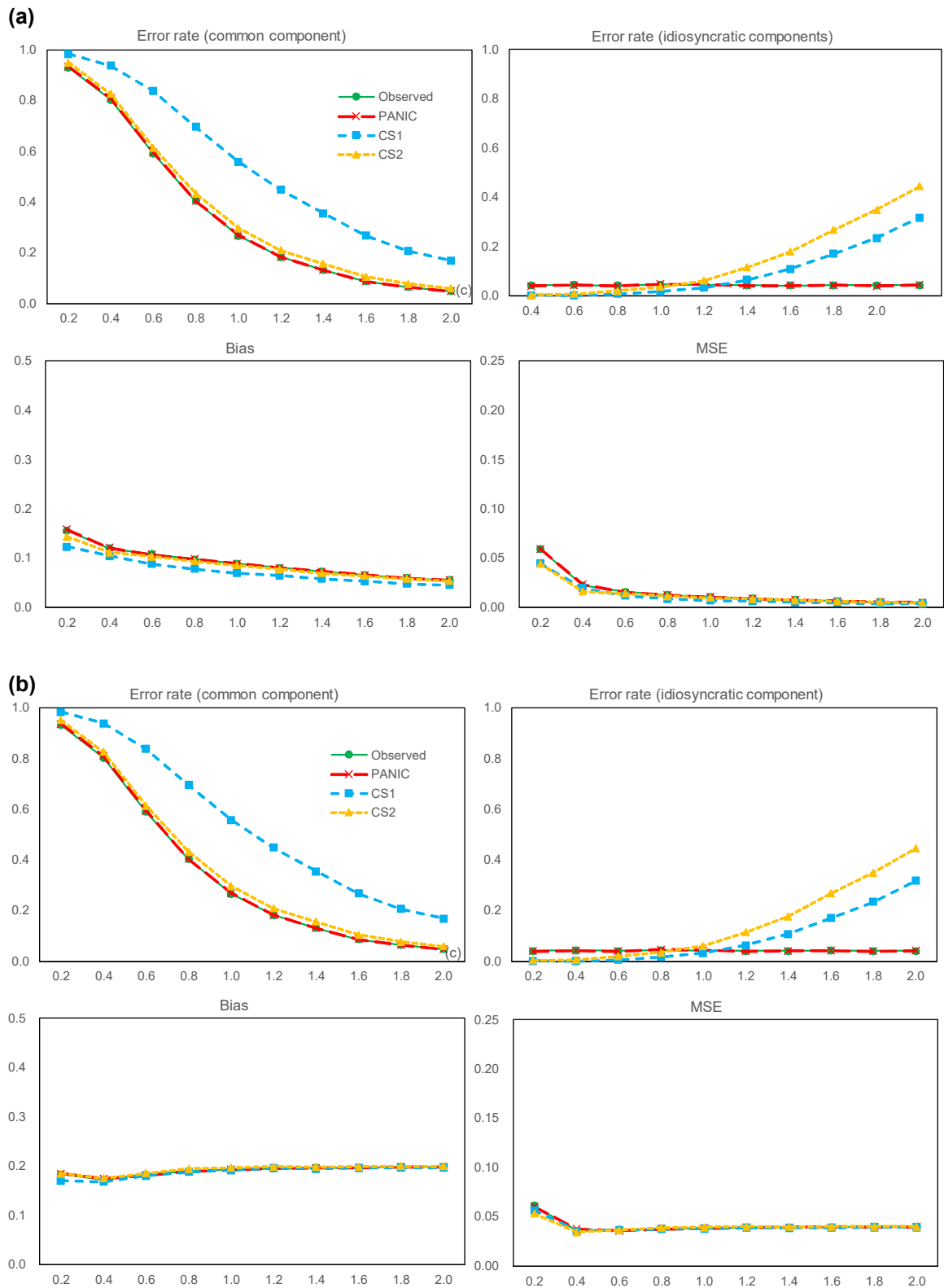


Figure 1: Error rates, Bias, and MSE when a single bubble is in the common component. (a) Origination date and (b) termination date.

documented by Yamamoto and Horie (2022). There is no large difference in the biases and the MSEs across the different methods in this case. Overall, PANIC shows very similar properties to the Observed, which proves the usefulness of PANIC in this case. However, CS1 and CS2 do not differ much from them. Figure 1b shows the results of the termination date. The results are very similar to those of the origination date, although the bias and the MSE do not decline as c becomes large, and this is regarded as an intrinsic delay caused by the sequential ADF test implemented forwardly.

Figure 2a in turn presents the results of the origination date for the case when idiosyncratic components have explosive behaviors but the common component does not. Remarkably, the error rate for the common component when using PANIC rapidly increases to one. This corresponds to the size distortion of the ADF test for the common component when idiosyncratic components are explosive. This occurs because PANIC misidentifies the common and idiosyncratic components and suffers from spurious explosive behavior in the former. More interestingly, the error rate for idiosyncratic components when using PANIC remains very close to one. This is due to the nonmonotonic power problem of the ADF test applied to idiosyncratic components as discussed by Yamamoto and Horie (2022). These imply that the nonexistent explosive behavior in the common component is falsely detected and the explosive behaviors in idiosyncratic components are hardly detected by PANIC. In contrast, the error rates of idiosyncratic components when using CS1 or CS2 show similar patterns to that of the Observed. The error rates of the common component when using CS1 and CS2 are low, although we see some tendency for over detection when c_i becomes large. The bias and the MSE when using PANIC are high across all values of c_i , however, when using CS1 or CS2, the values are very low and similar to those of the Observed. In addition, Figure 2b presents the results of the termination date for the same case and they are very similar to those of the origination date. Let us give a remark on how the sample is demarcated by CS1 and CS2. In these Monte Carlo results, the bias and MSE are very similar for both methods, however, the latter tends to provide a lower error rate than the former. This is because the goal of these methods is to detect the date until which we can obtain a clean estimate for the factor loadings. From this point of view, CS1 uses the cross-sectional average of $X_{i,t}$ and has little to do with the factor loading estimate, while CS2 uses the first principal component and has a more direct links to the factor loading estimate. Hence, we recommend CS2 for empirical practice. Of interest are the effects of the cross-sectional dimension N , because the factor estimate may become more precise as the cross-sectional regression when N increases. To that effect, in Figure 3a and b, we present results of the origination dates in the cases of $N = 50, 100$ and 200 when a single bubble is in the common component and when single bubbles are in idiosyncratic components, respectively. In Figure 3b, we observe some improvement in the error rate of the common components as N increases, but the results remain unchanged otherwise. This is consistent with Yamamoto and Horie (2022) in that the size of the test can slightly improve as both N and T increase, although the effect is not discernible. The results of the termination date are very similar to those of the origination date; hence, they are not reported. Figure 4 shows results of different N when we use PANIC. As expected, none of the error rates, bias, and MSE improve as N increases for both common and idiosyncratic components. This is because, although a larger N is expected to reduce factor estimation errors, the problem with PANIC is not factor estimation errors but rather misidentification between the common and idiosyncratic components.

Next, we consider the origination and termination dates when two bubbles exist. We use models with $m = m_i = 2$, where $N = 100$, $T = 250$, $T_e^{F,1} = T_e^{U,1} = 80$, $T_f^{F,1} = T_f^{U,1} = 120$, $T_e^{F,2} = T_e^{U,2} = 160$, and $T_f^{F,2} = T_f^{U,2} = 200$. We again let $c_j = c$ and $c_{i,j} = c_i$ for all i and $j = 1, 2$ and conduct two experiments where: (i) the common components are explosive ($c > 0$ and $c_i = 0$ for all i) and (ii) the idiosyncratic components are explosive ($c = 0$ and $c_i > 0$ for all i). After applying the date-stamping procedure based on the BSADF test described in Section 3.1, the error rates, bias, and MSE of the first bubble for the common components and the idiosyncratic components are computed. Those for the second bubble are also computed, but the results are qualitatively the same as those for the first one; thus, they are not reported. Figure 5a and b show the results of the origination and termination dates, respectively, when two bubbles are present in the common component. Figure 6a and b present these results when two bubbles exist in idiosyncratic components. The patterns are very similar to the case of single bubbles presented in Figures 1 and 2 except that the positive bias (delay) of the termination date becomes

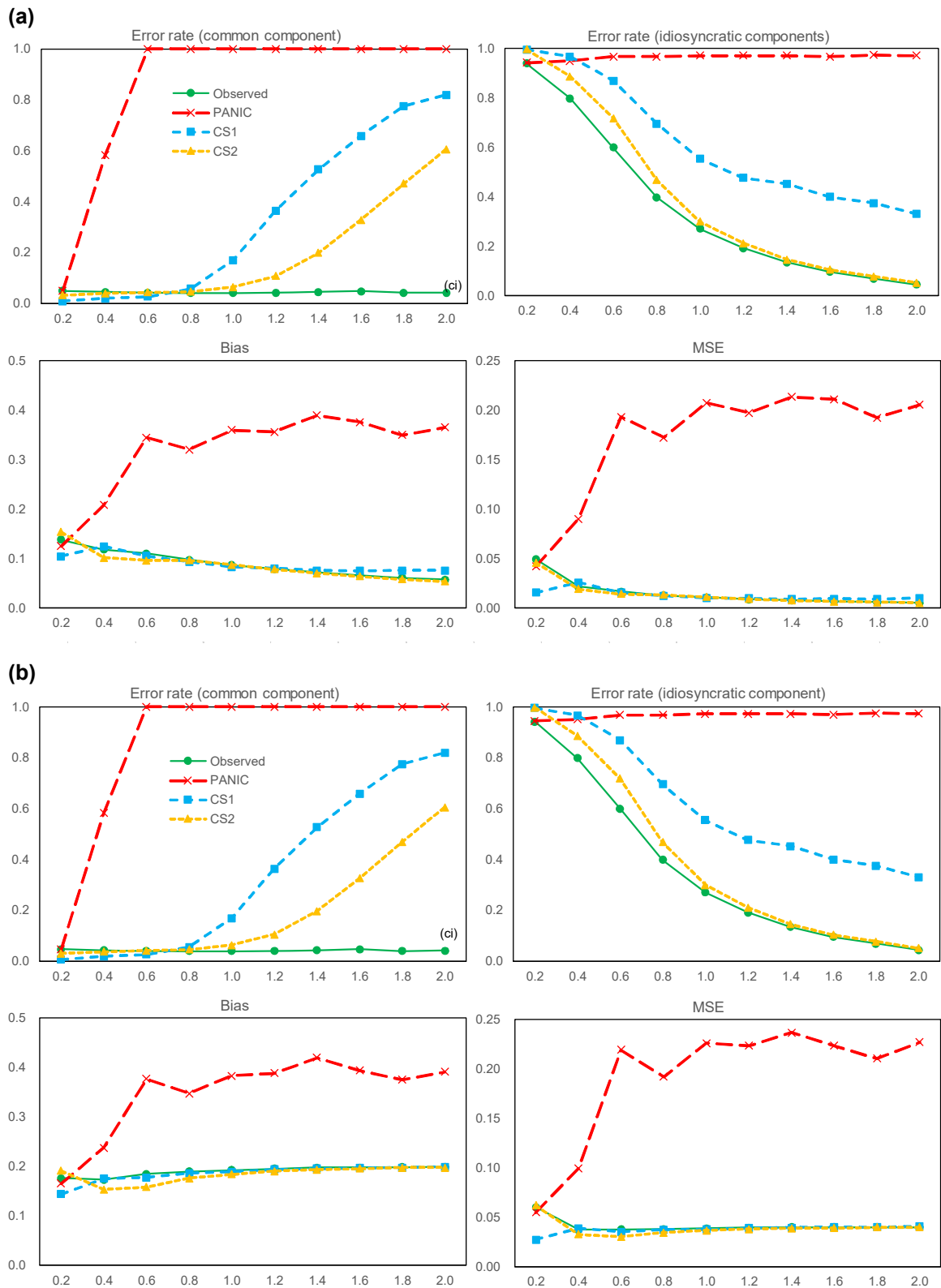


Figure 2: Error rates, Bias, and MSE when single bubbles are in the idiosyncratic components. (a) Origination date and (b) termination date.

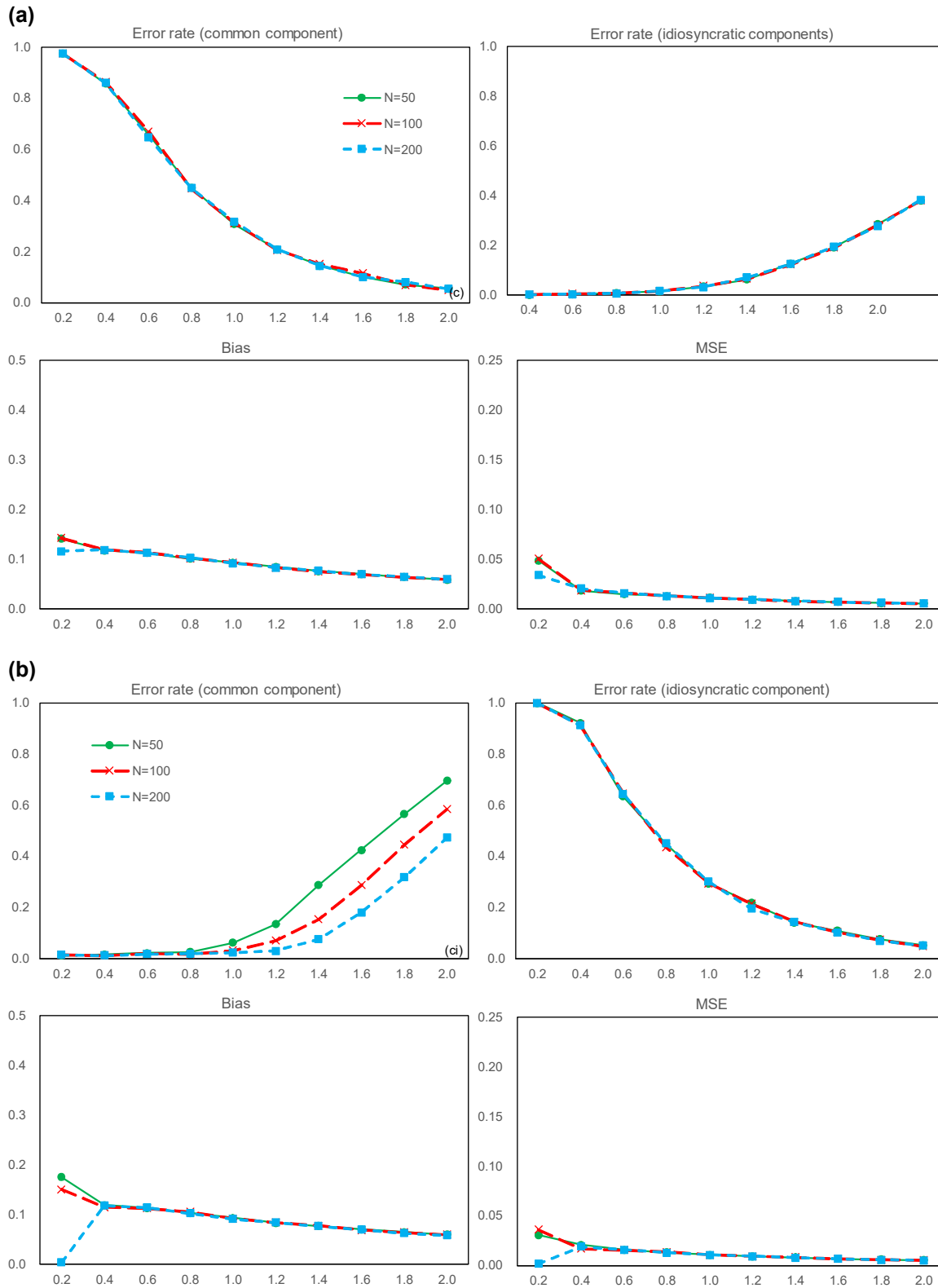


Figure 3: Error rates, Bias, and MSE of the origination date The CS2 method with $N = [50, 100, 200]$. (a) Single bubble is in the common component and (b) single bubbles are in idiosyncratic components.

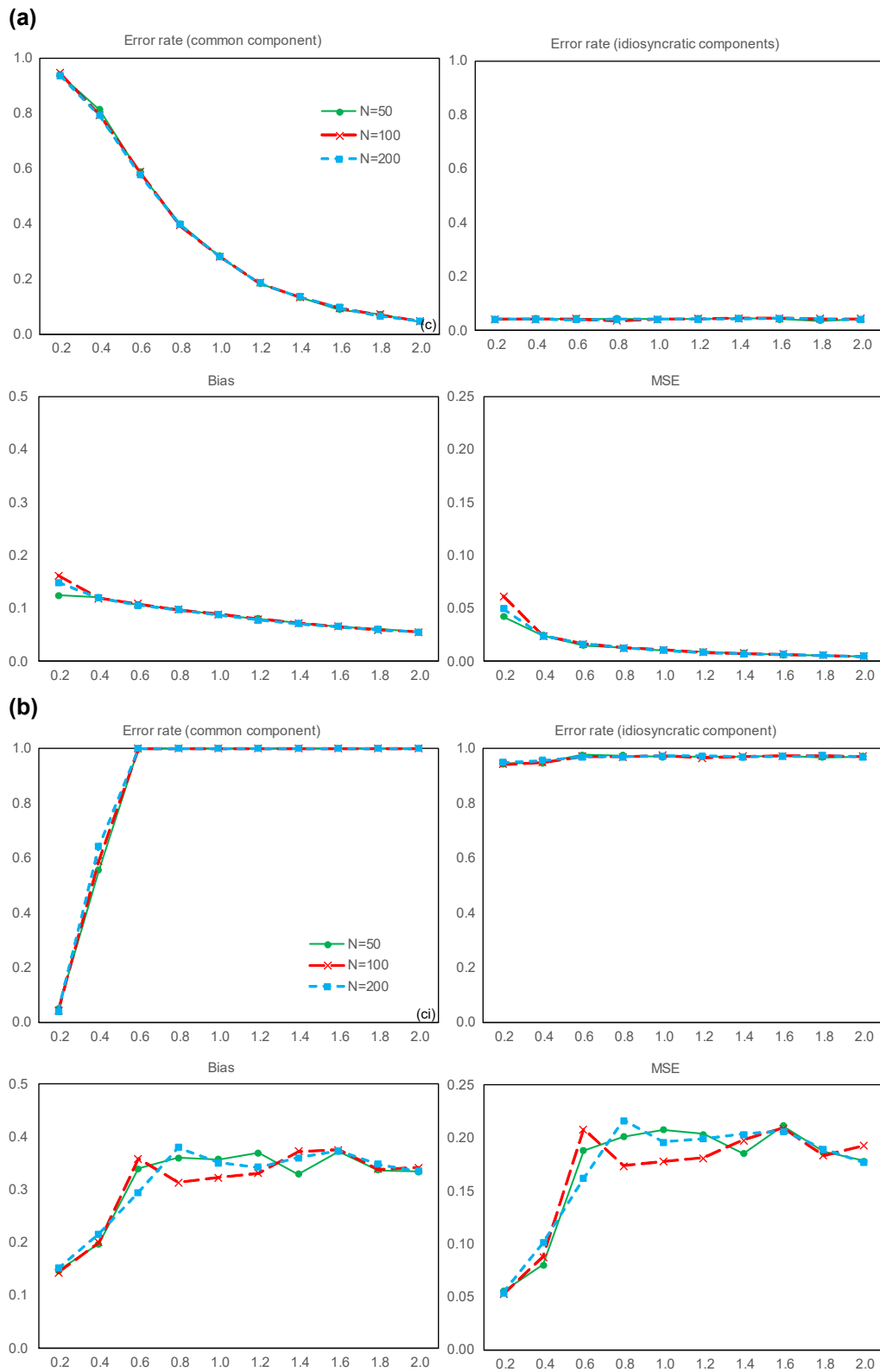


Figure 4: Error rates, Bias, and MSE of the origination date The PANIC method with $N = [50, 100, 200]$. (a) Single bubble is in the common component and (b) single bubbles are in idiosyncratic components.

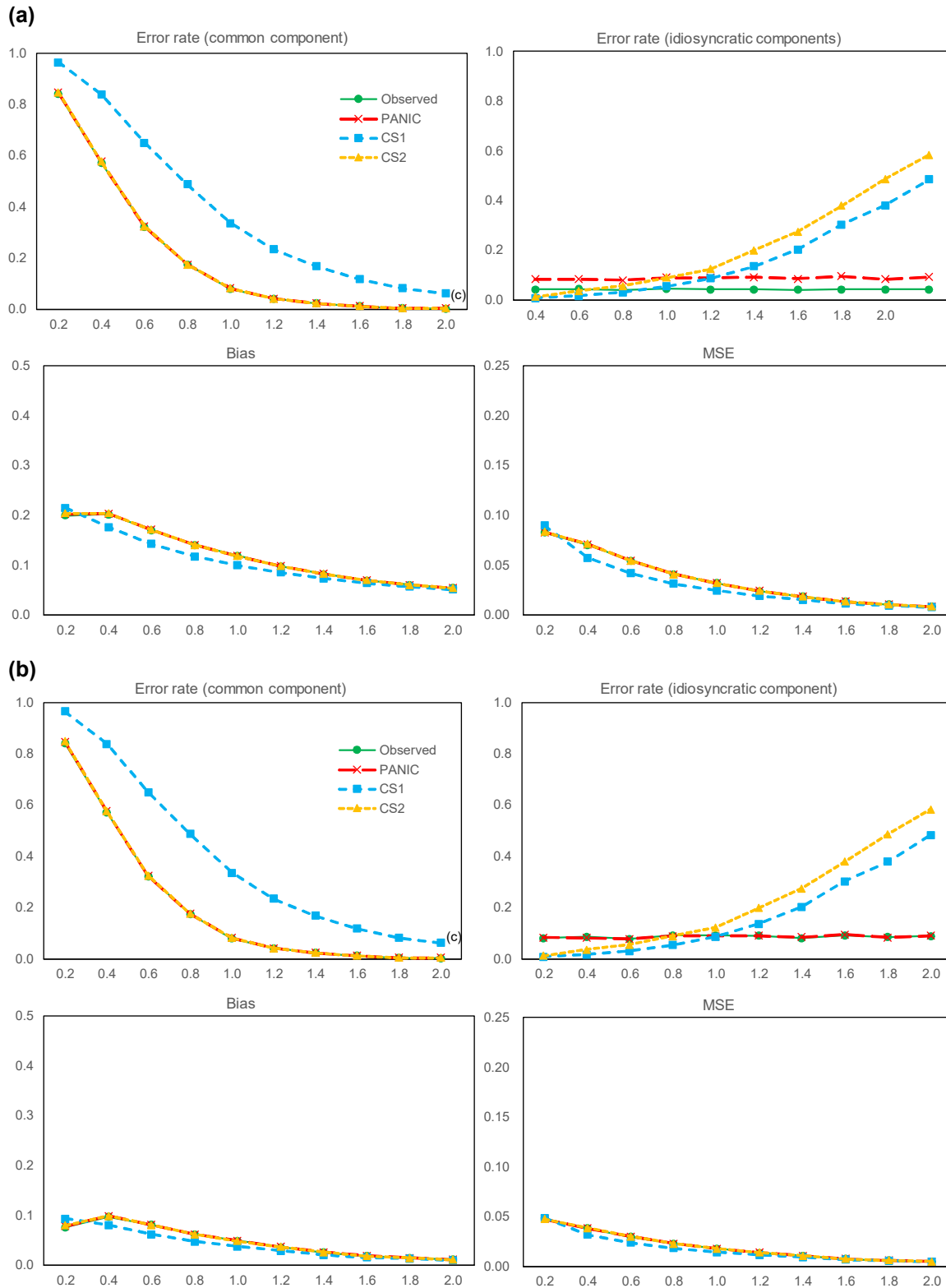


Figure 5: Error rates, Bias, and MSE when two bubbles are in the common component. (a) Origination date and (b) termination date.

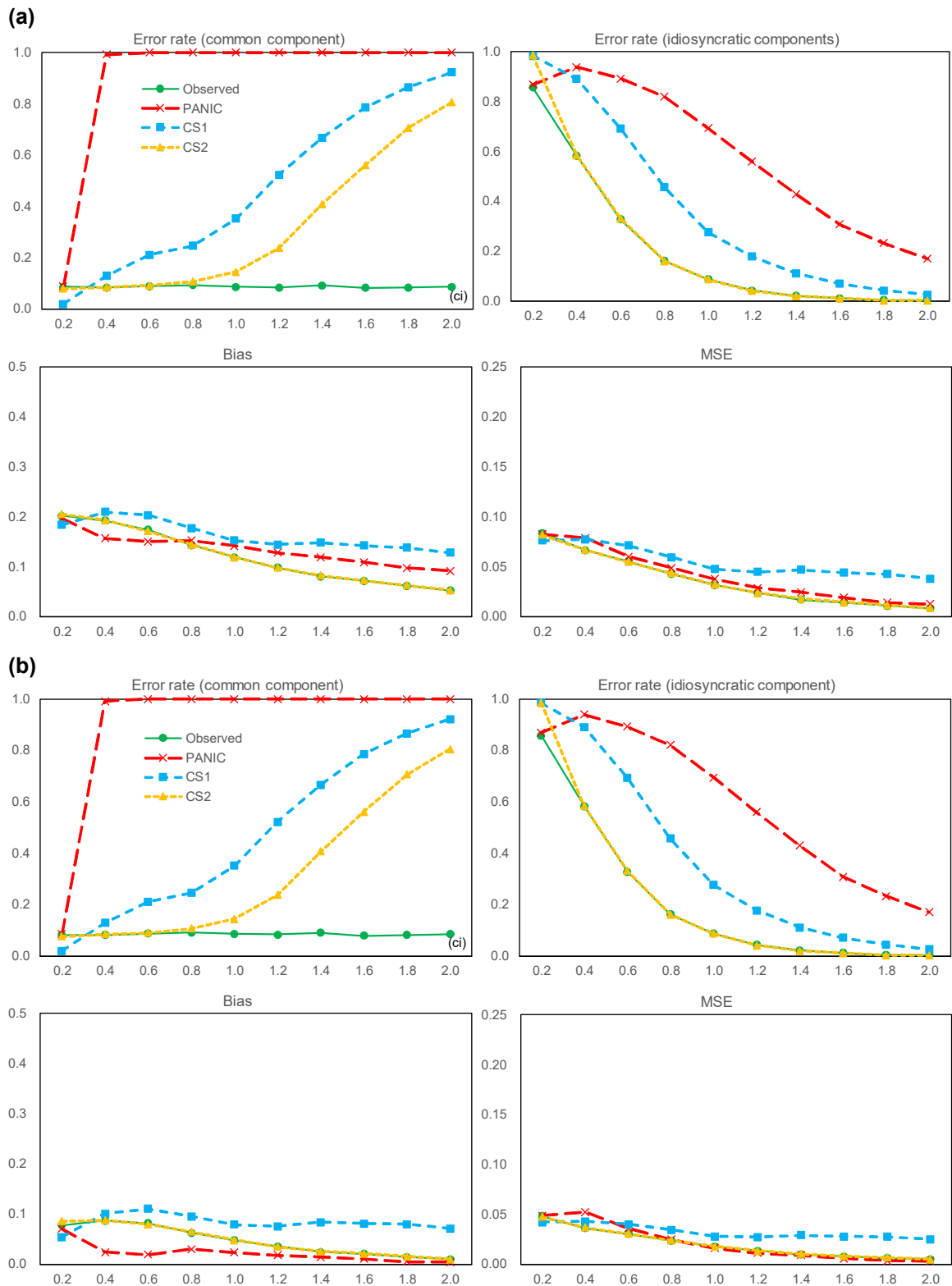


Figure 6: Error rates, Bias, and MSE when two bubbles are in the idiosyncratic components. (a) Origination date and (b) termination date.

smaller when the BSADF test is used. We also produced the results with different N s, but the differences from $N = 100$ are similar to those presented in Figure 3 and thus the figures are suppressed.

In summary, the properties of the right-tailed unit root tests investigated by Yamamoto and Horie (2022) are passed over to the date-stamping of single and multiple bubbles. When the bubble lies only in the common component, the origination and termination dates are precisely estimated by either PANIC or CS. However, when the bubbles exist in idiosyncratic components, PANIC loses its power of detection and suffers from spurious explosive behavior in the common component. These problems are addressed using CS, although some tendency for overdetection is observed when idiosyncratic components are strongly explosive.

5 Explosive Behaviors in the U.S. State-Level Real House Prices

During the past two decades, the U.S. has experienced unprecedented expansions in its housing markets. Figure 7 shows the national-level real HPI: the nominal HPI provided by the Federal Housing Finance Agency (FHFA) deflated by the Consumer Price Index (CPI).⁵ Both series are seasonally adjusted. The national-level real HPI started to increase in the late 1990s and peaked in 2006, followed by a sharp downturn during the subprime mortgage crisis. After hitting the lowest in 2011, it revamped and continued to increase throughout the 2010s onward. An important question is whether the latter expansion will be followed by a trend similar to the previous boom or whether it is built on a different mechanism. In this section, we provide some geographical evidence by comparing the two expansions through the application of the methods discussed in this study to the state-level real HPI data.

5.1 Related Literature

There is a growing literature discovering the dynamics of HPI by using geographical cross-section data. The central question is whether fluctuations are a national phenomenon or a collection of local behaviors. For example,

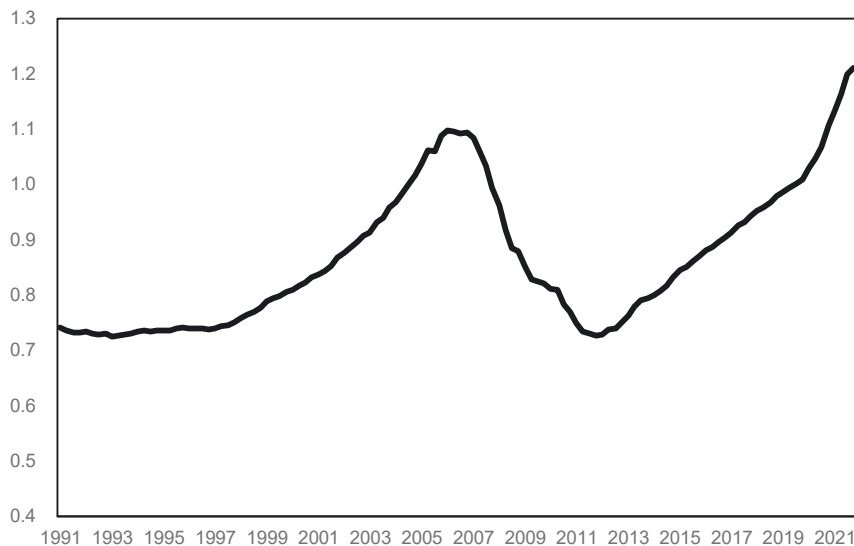


Figure 7: U.S. real HPI: national level.

⁵ We use the CPI of all urban consumers and all items.

Del Negro and Otrok (2007) used the state-level HPI from 1986 to 2005 and found that house prices were driven by local components, although the increase in 2001–2005 was considered a national phenomenon. Moench and Ng (2011) investigated the effects of house prices on consumption using data from four census regions (Northeast, Midwest, West and South). They found that the national and local components are of comparable order in Northeast, Midwest and South, but the local component is much more important in the West. Landvoigt et al. (2015) used finer cross-section data of capital gains and trading volumes in the county of San Diego from 1997 to 2008. They identified that cheaper credit for poor household was a major driver of the 2000s boom. Favara and Imbs (2015) used geographical dispersion of elasticity of housing supply to identify the shock of housing demand induced by banks' credit expansion. Hu and Oxley (2018) used the state-level HPI data from 1975 to 2014 and concluded that the bubble in the 2000s was widespread but was not a national phenomenon. Kuchler et al. (2022) emphasized the importance of local factors in forming housing market expectations. Cohen et al. (2021) identified seven clusters in the metropolitan statistical area (MSA)-level HPI data. These clusters experienced idiosyncratic downturns, although a national cycle existed during the Great Recession. Another important strand of research pertains simply to comovements in regional house prices. For example, Nieuwerburgh and Weill (2010) documented that the dispersion of MSA-level HPI increased since the late 1990s and it is attributed to the cross-sectional distribution of wages. Kallberg et al. (2014) used Case-Shiller home price indices for 14 MSAs from 1992 to 2008 and found that the comovements increased considerably in the late 1990s. Landier et al. (2017) and Choi (2019) discussed that the integration of the U.S. banks contributed to the increase of comovements of HPI.

5.2 Data and Methods

We use the HPI of the U.S. 50 states and the District of Columbia provided by the FHFA. The HPI is deflated by the same CPI as the real national-level index. The use of the MSA-level HPI is also an attractive option; however, it only delivers information on densely populated areas. As we would like to contrast results of the populated and unpopulated areas, the state-level data is preferred. We use the HPI constructed via house values based on repeated sales financed by Freddie Mac and Fannie Mae, augmented by data from an external source that captures trends different from those of enterprise-financed homes. It is known as "expanded-data HPI" and available on FHFA's website.⁶ The data are available from 1991Q1 onward and are seasonally adjusted. We use the sample from 1991Q1 to 2021Q4.

Our main strategy is the CS method described in Section 3. It first identifies the training sample in which no or only weak explosive behaviors exist so as not to contaminate the estimate for factor loadings. To select the training sample, we adopt the criterion of CS2 based on the first common factor identified by the PANIC method. This gives us a training sample from 1991Q1 to 1998Q1. Also, Bai and Ng's (2002) IC_{p2} information criterion picks five factors in the training sample. We then use cross-sectional regressions to estimate the common factors in the testing sample from 1998Q2 to 2021Q4. Once we obtain the common and idiosyncratic component estimates, we consider models (2) and (3) with $m = m_i = 2$ for all i .⁷ The BSADF test is used to date-stamp at most two bubbles. Because we are interested in investigating characteristics of the two explosive subperiods, for every bubble detected, we judge if it is in the 2000s or in the 2010s onward. To address this issue, if the greater portion of the bubble period is in the range prior to 2009Q4, it is considered as a bubble in the 2000s. If the greater portion is after 2010Q1, it is considered as a bubble in the 2010s onward.⁸ Figure 8 illustrates the timeline of our analysis. Our BSADF test is based on a model with a constant term and the number of lags determined by the modified information criterion of Ng and Perron (2001). We set the critical values $cv_t = \log(\log(t))$ for which the scaling constant is chosen by a separate simulation. The origination date is identified when the test statistic exceeds

⁶ <https://www.fhfa.gov/DataTools/Downloads/pages/house-price-index.aspx>.

⁷ We have also tried the case of $m = m_i = 3$ for all i . However, none of the common and idiosyncratic components detected the third bubble.

⁸ In the following analysis, there is only one case that stretches out the two subperiods if we use the CS method (the local bubble in Alabama).

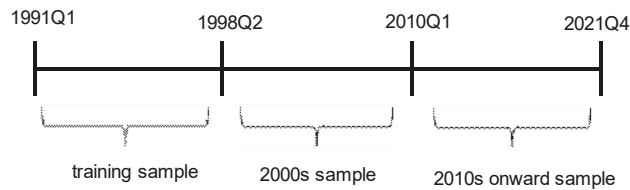


Figure 8: Timeline of the empirical analysis.

the critical value for ten consecutive quarters and, given that the origination is detected, the termination date is identified when the test falls below the critical value for ten consecutive quarters. When we use the PANIC method, the common and idiosyncratic components are obtained using the full sample from 1991Q1 to 2021Q4. The rest of the procedure is the same as the CS method.

5.3 Results

We first assess how the national component behaves during the two expansionary periods. Figure 9 presents the estimated common factors by the CS method from 1991Q1 to 2021Q4. We also compute five common factors by the PANIC method, but they are subject to misidentification as we discussed in the previous section and are not presented.⁹ It is known that the sign of the principal component estimate is undetermined. Therefore, to make sure that the factors have the same sign as the HPI, we flip the sign of the original factor estimate if more than half the factor loading estimates are negative. The second and the fourth common factors in the upper panel exhibit a similar boom to the aggregated index in the 2000s and we call them “bubble factors”. These factors explode in the expansion from the 2010s onward as well. The first, the third, and the fifth factors in the lower panel can be labeled as “non-bubble factors”, as they have little to do with the exuberance of the 2000s. The first factor swings in the 1990s but does not show a surge during the 2000s boom and the third and fifth factors capture milder fluctuations throughout the entire sample. An interesting feature is that for the 2000s boom the first and the fifth factors moved downward while the other factors went upward, while all common factors increased from the 2010s onward. Hence, the two bubble factors lead the 2000s boom, but all five national factors contribute to the expansion from the 2010s onward. Exploring economic implications of these non-bubble factors must be an interesting agenda, however, it is beyond the scope of this study.

Figure 10 captures the geographical pattern of the effects of the common factors by showing the p -values of two-sided t tests for the zero factor loading in each state. These are based on tests obtained in the training sample.¹⁰ Darker blue indicates a smaller p -value so that the loading is significantly different from zero at a lower significance level. The non-bubble factors have significant impacts across all states uniformly, while the bubble factors show strongly significant impacts on the coastal populated states such as California, Florida, Massachusetts and New York but insignificant impacts on less populated inland states such as North Dakota, South Dakota, Montana, and Wyoming. This gives further evidence of the heterogeneity in the 2000s boom as it is driven by the two bubble factors. The homogeneity in the 2010s onward expansion was caused by an increase across all common factors.

Table 1 shows the origination and termination dates of bubbles in the common and idiosyncratic components identified by the method using the BSADF test. If a bubble is detected, the year and quarter are shown. If

⁹ When we use PANIC, the estimated first common factor is strongly correlated with the two bubble factors identified by CS. The intuition behind this is that the first principal component captures the largest variations and thus all or most of the explosive processes. In addition, the second common factor identified by PANIC is closely related to the first identified by CS. The fourth common factor identified by PANIC coincides with the fifth identified by CS. Therefore, the third and the fifth common factors identified by PANIC span a different space from that of CS. These might identify some idiosyncratic components that have a large variation as the common components.

¹⁰ With the CS method, it is important to assume that the factor loadings and the number of factors are constant for the entire sample. However, it does not matter whether the AR coefficient changes smoothly or abruptly. If it changes smoothly to the same level, the ADF test has a lower power and is thus anticipated to make the date-stamping less precise.

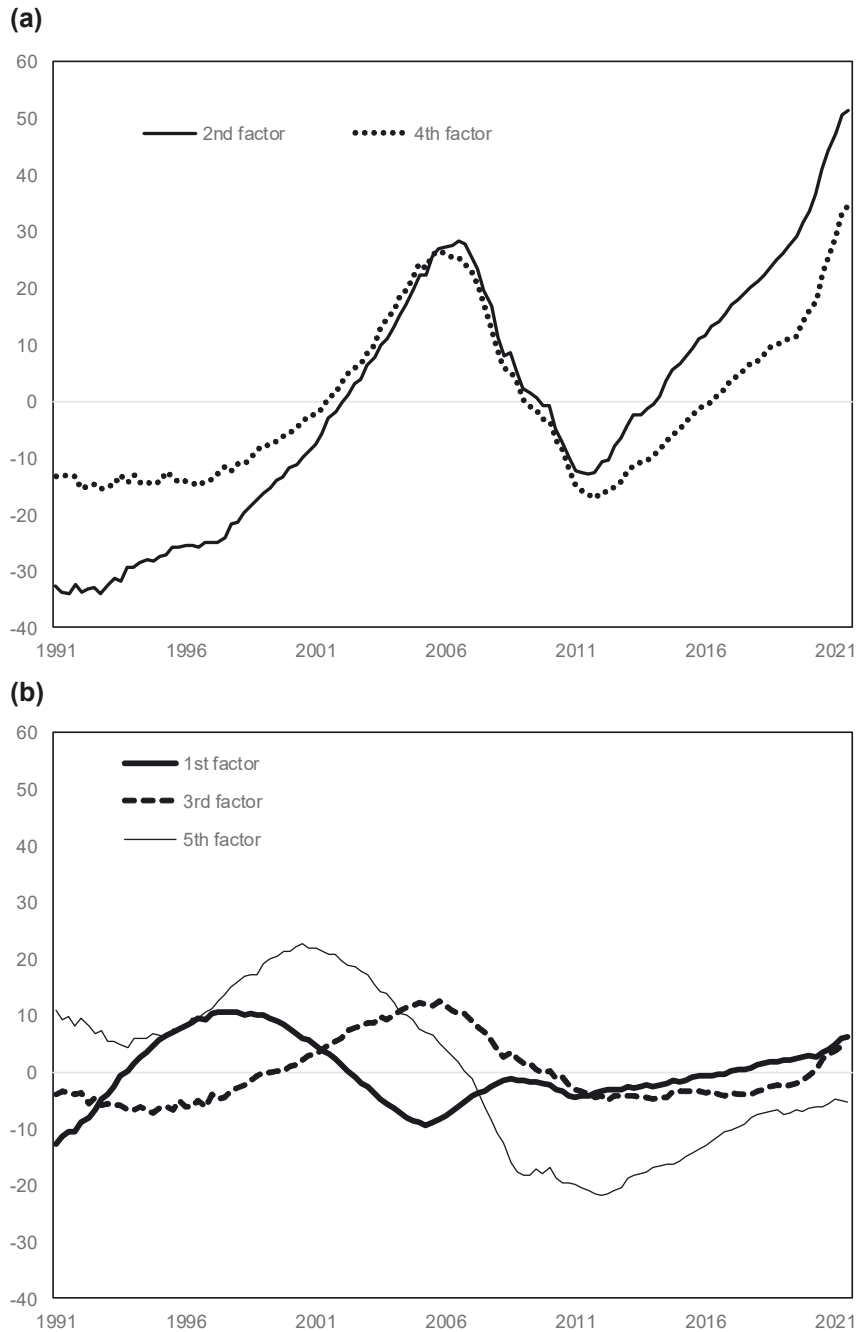
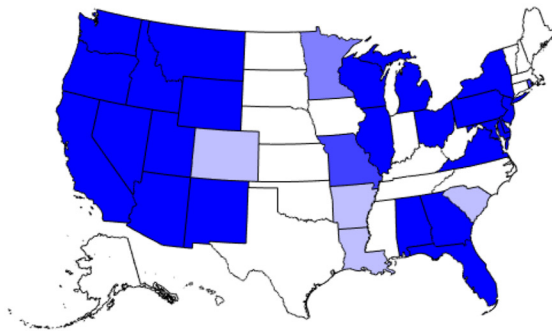


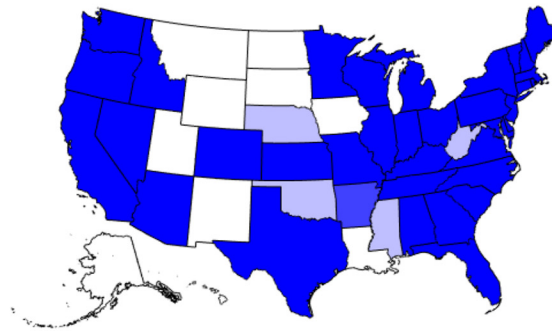
Figure 9: Common components identified by the CS method. (a) Bubble factors and (b) Non-bubble factors.

no bubble is found, we leave it as a blank. The results using the CS method are presented in columns 2 and 3 and those using the PANIC method, in columns 4 and 5. The results for the common component using the CS method agree with our previous investigations. The two bubble factors start to explode in 1998Q4 and collapse in 2006Q2–3. They roughly coincide with the exuberance period suggested by the literature. From the 2010s onward, the first factor starts to explode in 2018Q1, while the others are increasing but not to such an extent. Furthermore, we are interested in local bubbles captured by the idiosyncratic components. The CS method suggests that many states, including highly populated ones such as California, Florida, and Massachusetts, experienced

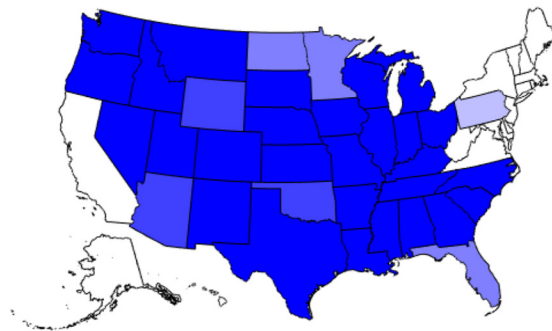
First factor (non-bubble factor)



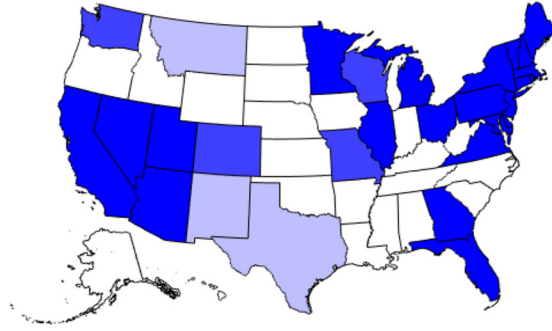
Second factor (bubble factor)



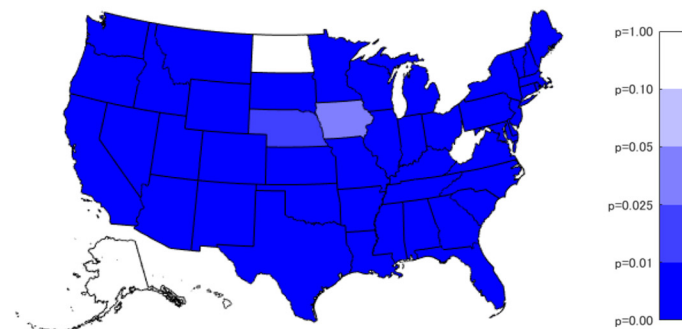
Third factor (non-bubble factor)



Fourth factor (bubble factor)



Fifth factor (non-bubble factor)

**Figure 10:** Statistical significance of factor loadings.

a local bubble in the 2000s. The origination dates are mostly marked in the early 2000s and are consistent with local bubbles discussed in the literature.^{11,12} In contrast, in the 2010s onward, much fewer states exhibit a local bubble.

¹¹ Federal Reserve Board (2005) stated that there appear to be signs of froth in some local markets where home prices seem to have risen to unsustainable levels.

¹² A referee is concerned that the new common factors emerged during this period, but to our knowledge, there is no method to estimate the number of factors in an explosive environment. Hence, we are not in a position to give evidence of such factors.

Table 1: Date-stamping of common and idiosyncratic explosive behaviors.

	CS						PANIC			
	1991Q1–2009Q4			2010Q1–2021Q4			1991Q1–2009Q4		2010Q1–2021Q4	
	Origination	Termination		Origination	Termination		Origination	Termination	Origination	Termination
Common components										
1st factor (non-bubble)				2018	1		1998	3	2006	2
2nd factor (bubble)	1998	4	2006	3					2017	2
3rd factor (non-bubble)										
4th factor (bubble)	1998	4	2006	2					2011	2
5th factor (non-bubble)									2020	1
Idiosyncratic components										
Alaska (AK)					4	2014	3		2008	2
Alabama (AL)				2009					2019	2
Arkansas (AR)	2001	2	2010	2						
Arizona (AZ)	2001	4	2005	4	2019	2			2009	4
California (CA)	2002	2	2006	1					2016	2
Colorado (CO)										
Connecticut (CT)	2005	4	2009	1						
District of Columbia (DC)	2004	4	2008	1						
Delaware (DE)	2002	2	2006	4						
Florida (FL)	2003	4	2006	3						
Georgia (GA)										
Hawaii (HI)									2014	2
Iowa (IA)									2018	2
Idaho (ID)	2002	3	2006	3	2017	3				
Illinois (IL)	2001	3	2004	1					2009	2
Indiana (IN)	2002	3	2009	3	2016	4	1997	3	2001	4
Kansas (KS)	2002	4	2006	3						
Kentucky (KY)	2004	1	2007	3						
Louisiana (LA)					2017	3				
Massachusetts (MA)	1999	2	2003	1						
Maryland (MD)					2017	3			2010	3
Maine (ME)	2000	3	2006	3						
Michigan (MI)	2003	3	2009	2						
Minnesota (MN)	1999	1	2001	3			1999	4	2002	2

Table 1: (continued)

	CS						PANIC			
	1991Q1–2009Q4			2010Q1–2021Q4			1991Q1–2009Q4		2010Q1–2021Q4	
	Origination	Termination		Origination	Termination		Origination	Termination	Origination	Termination
Missouri (MO)										
Mississippi (MS)	2000	2	2005	3						
Montana (MT)	2001	4	2005	3						
North Carolina (NC)										
North Dakota (ND)									2011	4
Nebraska (NE)										2016
New Hampshire (NH)										3
New Jersey (NJ)										
New Mexico (NM)							2001	2	2004	1
Nevada (NV)									2012	4
New York (NY)										
Ohio (OH)	2003	3	2010	4	2017	3	2021	4		
Oklahoma (OK)	2002	2	2006	1	2018	3		1999	2	2001
Oregon (OR)	2001	2	2006	3						
Pennsylvania (PA)										
Rhode Island (RI)										
South Carolina (SC)										
South Dakota (SD)	2004	2	2007	2	2018	2	2021	4		
Tennessee (TN)										
Texas (TX)									2013	2
Utah (UT)										
Virginia (VA)	2001	1	2006	2	2017	4				
Vermont (VT)										
Washington (WA)	2004	2	2006	4				2006	1	2008
Wisconsin (WI)	2001	1	2004	2	2017	1	2021	4		
West Virginia (WV)									2012	2
Wyoming (WY)	2001	3	2008	1					2015	4

Table 2 shows bias-corrected estimates for the AR coefficients α and ρ_i and their 95% confidence intervals using the identified bubble sample $[\hat{T}_e + 1, \hat{T}_f]$. If the bubble is ongoing and no termination date is available, we use the sample up to 2021Q4. It is known that the OLS estimate of an AR coefficient has a downward bias if the true value is positive. To correct this bias, we construct an indirect inference estimate proposed by Smith (1993) and Phillips et al. (2011). To explain the methodology, let us denote the sample of estimated common or idiosyncratic components by $\{x_t\}_{t=\hat{T}_e^x+1}^{\hat{T}_f^x}$, the OLS estimate of the AR coefficient by $\hat{\alpha}_{OLS}$ and its residuals by $\{\hat{e}_{\hat{T}_e^x+1}, \dots, \hat{e}_{\hat{T}_f^x}\}$. The indirect inference method first generates mock data $x_t^{*(1)}$ using $x_t^{*(1)} = \alpha_0 x_{t-1}^{*(1)} + e_t^{*(1)}$ for $t = \hat{T}_e^x + 1, \dots, \hat{T}_f^x$ with $x_{\hat{T}_e^x}^{*(1)} = x_{\hat{T}_e^x}$, where $e_t^{*(1)}$ is resampled with replication from the residuals and α_0 being some initial value. Let an AR coefficient estimate of $\{x_t^{*(1)}\}_{t=\hat{T}_e^x+1}^{\hat{T}_f^x}$ be $\tilde{\alpha}^{(1)}(\alpha_0)$ and repeat this H times to obtain $\tilde{\alpha}^{(1)}(\alpha_0), \tilde{\alpha}^{(2)}(\alpha_0), \dots, \tilde{\alpha}^{(H)}(\alpha_0)$. Then, α which minimizes the criteria $(\hat{\alpha}_{OLS} - \frac{1}{H} \sum_{h=1}^H \tilde{\alpha}^{(h)}(\alpha))^2$ is considered a bias-corrected estimate and is denoted by $\hat{\alpha}_H$. We specifically choose $H = 1000$. Given the asymptotic distribution of the AR coefficient estimator under the mildly explosive process derived by Phillips and Magdalinos (2007) and applied in the present context by Yamamoto and Horie (2022), the 95% asymptotic confidence interval is constructed by

$$\hat{\alpha}_H \pm 12.7 \times \frac{(\hat{\alpha}_H)^2 - 1}{(\hat{\alpha}_H)^{\hat{T}_f^x - \hat{T}_e^x}},$$

where 12.7 is the 97.5 percentile of the standard Cauchy distribution.¹³ The results in Table 2 show that the AR coefficient estimate ranges from 1.05 to 1.20 in most of the identified bubble periods and the magnitudes of idiosyncratic bubbles are heterogeneous across states, suggesting the importance of investigating state-specific bubbles. However, since the confidence intervals are generally wide because of the small sample size, the results have to be assessed with caution.

Figure 11 depicts the origination dates of the local bubbles with states in darker red standing for earlier dates, those in lighter red for later dates, and those in white for no local bubbles. The upper panel presents the 2000s sample, showing some similarity to the loadings of bubble factors in Figure 10, suggesting that the 2000s boom was driven by not only the national bubble factors but also a set of local components. However, the lower panel shows that the local components are weak from the 2010s onward and thus the expansion is dominated by the effect of national components. Columns 4 and 5 of Table 1 show the results from the PANIC method. Explosive behaviors in the common component are identified in the 2000s and from the 2010s onward but those in idiosyncratic components are detected in much fewer states than when using the CS method. Importantly, some highly populated states such as Florida and Massachusetts show no local bubbles during the 2000s boom. These omissions may be caused by the power loss of the PANIC method as discovered in the previous section.

Finally, as a robustness check, we separated the testing sample into two subperiods: one from 1998Q2 to 2009Q4 and the other from 2010Q1 to 2021Q4, after we identified the common and idiosyncratic components. Thereafter, we employed the methods for detecting single bubbles (5) and (6) in each subperiod. Our conclusions remain qualitatively unchanged except that no termination dates are found in the second subperiod. This is possibly because of the delay bias of the single bubble method as documented in our Monte Carlo simulation.

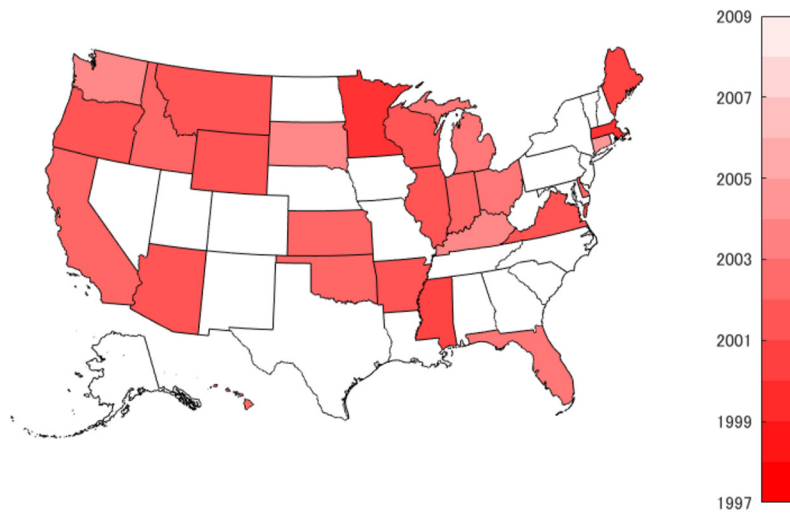
¹³ This result is valid if the factor estimation errors are irrelevant as $N, T \rightarrow \infty$. The conditions are derived in Theorem 2 of Yamamoto and Horie (2022) for the ADF test and are sufficient for the AR coefficient estimate.

Table 2: (continued)

	Dates						AR coefficients	
	1991Q1–2009Q4			2010Q1–2021Q4			1991Q1–2009Q4	2010Q1–2021Q4
	Origination	Termination		Origination	Termination			
Missouri (MO)								
Mississippi (MS)	2000	2	2005	3			1.093 [0.743, 1.443]	
Montana (MT)	2001	4	2005	3			1.149 [0.706, 1.591]	
North Carolina (NC)								
North Dakota (ND)								
Nebraska (NE)								
New Hampshire (NH)								
New Jersey (NJ)								
New Mexico (NM)								
Nevada (NV)								
New York (NY)				2017	3	2021		1.133 [0.753, 1.513]
Ohio (OH)	2003	3	2010	4			1.070 [0.829, 1.311]	1.171 [0.654, 1.688]
Oklahoma (OK)	2002	2	2006	1			1.159 [0.749, 1.570]	
Oregon (OR)	2001	2	2006	3			1.129 [0.886, 1.371]	
Pennsylvania (PA)				2018	2	2021		1.178 [0.754, 1.601]
Rhode Island (RI)								
South Carolina (SC)								
South Dakota (SD)	2004	2	2007	2			1.151 [0.486, 1.815]	
Tennessee (TN)								
Texas (TX)								
Utah (UT)								
Virginia (VA)	2001	1	2006	2	2017	4	1.123 [0.866, 1.381]	1.174 [0.860, 1.488]
Vermont (VT)								
Washington (WA)	2004	2	2006	4			1.227 [0.550, 1.904]	
Wisconsin (WI)	2001	1	2004	2	2017	1	1.170 [0.648, 1.691]	1.124 [0.803, 1.446]
West Virginia (WV)								
Wyoming (WY)	2001	3	2008	1			1.067 [0.761, 1.373]	

The 95% asymptotic confidence interval in brackets. The AR coefficient estimate is constructed by the method of indirect inference and the confidence interval is valid under the moderate deviation framework of Phillips and Magdalinos (2007) and Yamamoto and Horie (2022).

The 2000s sample



The 2010s onward sample

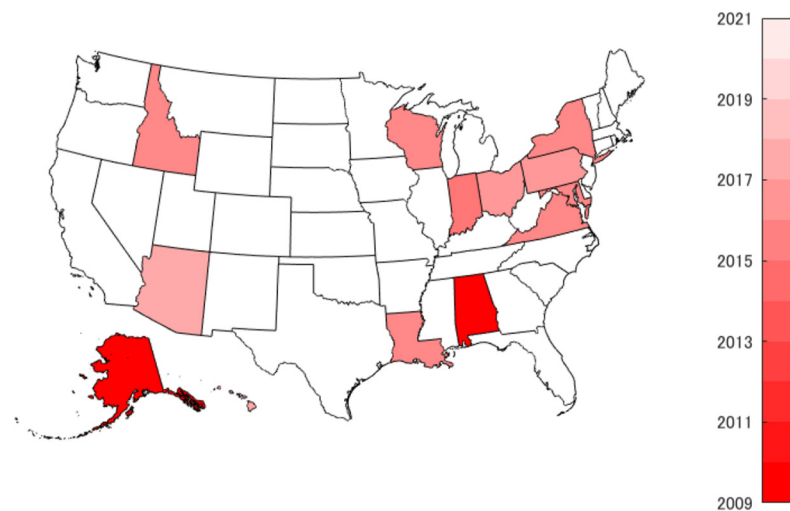


Figure 11: Origination dates of idiosyncratic explosive behaviors.

6 Conclusions

We applied the date-stamping methodologies for the origination and termination of bubbles proposed in the seminal work of Phillips et al. (2011, 2015a, 2015b) to the large dimensional common factor model. To this end, we compared two methods of identifying the common and idiosyncratic components: PANIC and CS. As discovered by Yamamoto and Horie (2022), when the PANIC method is used, the unit root test for the common component may suffer from serious size distortions and that for idiosyncratic components exhibits the nonmonotonic power problem. Our Monte Carlo simulation shows that these features are passed over to the identification methodology. When the bubble lies only in the common component, these dates are precisely estimated by either the PANIC or CS method. However, when the bubbles exist in idiosyncratic components, the PANIC method loses

its power of detection and suffers from spurious explosive behaviors in the common component. These problems are resolved when the CS method is used, although some tendency for overdetection is observed when idiosyncratic components are strongly explosive.

Our empirical application of the CS method to the U.S. state-level real HPI provides new insights into the driving forces of house price dynamics during the recent booms. The five common factors identified by the CS method show clear characteristics: two bubble factors and three non-bubble factors. The geographical patterns of their impacts are uncovered. Also, the results for idiosyncratic components during the boom in the 2000s are consistent with the existence of a set of local bubbles especially in populated states as arguably discussed in the literature of subprime crisis, while in the 2010s onward we newly found the fact that much fewer states exhibit local explosive behaviors. This stark contrast regarding the roles of national and local components between the 2000s and the 2010s onward periods may be due to monetary policies and global supply chain disruptions under the COVID-19 period, for instance, although further detailed analysis is beyond the scope of this study.

We conclude with a few interesting directions to extend the econometric methodologies to invite further theoretical and empirical research. First, one of our crucial assumptions is that the number of factors and the factor loadings are constant for the entire sample. Any attempt to relax this assumption would surely be a useful extension. Second, although the common factor model is an effective and widely accepted way to summarize large dimensional panel data, even more flexible and realistic lag structures in the identified idiosyncratic components may shed light on rich dynamics among asset prices, such as spillover effects across individual assets. Finally, the method proposed in this study contributes to not only the house price dynamics but also a variety of large dimensional financial panel data sets that exhibit comovements with widespread booms and busts in highly integrated economies.

References

- Ang, A., and M. Piazzesi. 2003. "A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables." *Journal of Monetary Economics* 50 (4): 745–87.
- Bai, J. 2003. "Inferential Theory for Factor Models of Large Dimensions." *Econometrica* 71 (1): 135–71.
- Bai, J., and S. Ng. 2002. "Determining the Number of Factors in Approximate Factor Models." *Econometrica* 70 (1): 191–221.
- Bai, J., and S. Ng. 2004. "A Panic Attack on Unit Roots and Cointegration." *Econometrica* 72 (4): 1127–77.
- Chen, Y., P. C. B. Phillips, and S. Shi. 2022. "Common Bubble Detection in Large Dimensional Financial Systems." *Journal of Financial Econometrics*: 1–75, <https://doi.org/10.1093/jfinec/nbab027>.
- Choi, C.-Y. 2019. *Understanding Geographic Comovement of House Prices among U.S. Cities: The Role of Financial Integration*. New York, USA: Mimeo.
- Cohen, J. P., C. C. Coughlin, and D. Soques. 2021. "House Price Growth Independencies and Comovement." In *Federal Reserve Bank of St. Louise Working Paper 2019–028*.
- Del Negro, M., and C. Otrok. 2007. "99 Luftballons: Monetary Policy and the House Price Boom across U.S. States." *Journal of Monetary Economics* 54: 1962–85.
- Fama, E. F. 1970. "Efficient Capital Market: A Review of Theory and Empirical Work." *The Journal of Finance* 25 (2): 383–417.
- Fama, E. F., and K. R. French. 1993. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics* 33: 3–56.
- Favara, G., and J. Imbs. 2015. "Credit Supply and the Price of Housing." *The American Economic Review* 105 (3): 958–92.
- Federal Reserve Board. 2005. *Testimony of Chairman Alan Greenspan: The Economic Outlook Before the Joint Economic Committee*. Washington, D.C: U.S. Congress.
- Gürkaynak, R. S. 2008. "Econometric Tests of Asset Price Bubbles: Taking Stock." *Journal of Economic Surveys* 22 (1): 166–86.
- Homm, U., and J. Breitung. 2012. "Testing for Speculative Bubbles in Stock Markets: A Comparison of Alternative Methods." *Journal of Financial Econometrics* 10 (1): 198–231.
- Hu, Y., and L. Oxley. 2018. "Bubbles in US Regional House Prices: Evidence from House Price-Income Ratios at the State Level." *Applied Economics* 50 (29): 3196–229.
- Kallberg, J. G., C. H. Liu, and P. Pasquariello. 2014. "On the Price Comovement of U.S. Residential Real Estate Markets." *Real Estate Economics* 42: 71–108.
- Krugman, P. 2022. *Working Out: Are We in Another Housing Bubble?* New York: The New York Times.
- Kuchler, T., M. Piazzesi, and J. Stroebel. 2022. "Housing Market Expectations." In *CEPR Press Discussion Paper No. 17158*.
- Landier, A., D. Sraer, and D. Thesmar. 2017. "Banking Integration and House Price Co-movement." *Journal of Financial Economics* 125: 1–25.

- Landvoigt, T., M. Piazzesi, and M. Schneider. 2015. "The Housing Market(s) of San Diego." *The American Economic Review* 105 (4): 1371–407.
- Litterman, R. B., and J. Scheinkman. 1991. "Common Factors Affecting Bond Returns." *Journal of Fixed Income* 1 (1): 54–61.
- Moench, E., and S. Ng. 2011. "A Hierarchical Factor Analysis of U.S. Housing Market Dynamics." *The Econometrics Journal* 14 (1): C1–C24.
- Ng, S., and P. Perron. 2001. "Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power." *Econometrica* 69 (6): 1519–54.
- Nieuwerburgh, S. V., and P.-O. Weill. 2010. "Why Has House Price Dispersion Gone up?" *The Review of Economic Studies* 77: 1567–606.
- Phillips, P. C. B., and T. Magdalinos. 2007. "Limit Theory for Moderate Deviations from a Unit Root." *Journal of Econometrics* 136 (1): 115–30.
- Phillips, P. C. B., and J. Yu. 2011. "Dating the Timeline of Financial Bubbles during the Subprime Crisis." *Quantitative Economics* 2 (3): 455–91.
- Phillips, P. C. B., Y. Wu, and J. Yu. 2011. "Explosive Behavior in 1990s Nasdaq: When Did Exuberance Escalate Asset Values?" *International Economic Review* 51 (2): 201–26.
- Phillips, P. C. B., S. Shi, and J. Yu. 2015a. "Testing for Multiple Bubbles: Historical Episodes of Exuberance and Collapse in the S&P 500." *International Economic Review* 56 (4): 1043–78.
- Phillips, P. C. B., S. Shi, and J. Yu. 2015b. "Testing for Multiple Bubbles: Limit Theory of Real Time Detectors." *International Economic Review* 56 (4): 1079–134.
- Smith, A. A. 1993. "Estimating Nonlinear Time-Series Models Using Simulated Vector Autoregressions." *Journal of Applied Econometrics* 8: S63–84.
- Stock, J. H., and M. W. Watson. 2002. "Forecasting Using Principal Components from a Large Number of Predictors." *Journal of the American Statistical Association* 97 (460): 1167–79.
- Yamamoto, Y., and T. Horie. 2022. "A Cross-Sectional Method for Right-Tailed PANIC Tests under a Moderately Local to Unity Framework." *Econometric Theory* (forthcoming), <https://doi.org/10.1017/s0266466622000044>.