Graduate School of Economics, Hitotsubashi University
Discussion Paper Series No. 2016-04

Testing for Speculative Bubbles in Large-Dimensional Financial Panel Data Sets

Tetsushi Horie Yohei Yamamoto

June 2016

Testing for Speculative Bubbles in Large-Dimensional Financial Panel Data Sets

Tetsushi Horie*
Hitotsubashi University

Yohei Yamamoto†
Hitotsubashi University

June 17, 2016

Abstract

Towards the financial crisis of 2007 to 2008, speculative bubbles prevailed in various financial assets. Whether these bubbles are an economy-wide phenomenon or market-specific events is an important question. This study develops a testing approach to investigate whether the bubbles lie in the common or in the idiosyncratic components of large-dimensional financial panel data sets. To this end, we extend the right-tailed unit root tests to common factor models, benchmarking the panel analysis of nonstationarity in idiosyncratic and common component (PANIC) proposed by Bai and Ng (2004). We find that when the PANIC test is applied to the explosive alternative hypothesis as opposed to the stationary alternative hypothesis, the test for the idiosyncratic component may suffer from the nonmonotonic power problem. In this paper, we newly propose a cross-sectional (CS) approach to disentangle the common and the idiosyncratic components in a relatively short explosive window. This method first estimates the factor loadings in the training sample and then uses them in cross-sectional regressions to extract the common factors in the explosive window. A Monte Carlo simulation shows that the CS approach is robust to the nonmonotonic power problem. We apply this method to 24 exchange rates against the U.S. dollar to identify the currency values that were explosive during the financial crisis period.

JEL Classification Number: C12, C38, F31

Keywords: speculative bubbles, explosive behaviors, factor model, moderate deviations, local asymptotic power, nonmonotonic power

*Hitotsubashi University, Graduate School of Economics, 2-1 Naka, Kunitachi, Tokyo, Japan 186-8601 (ed161002@g.hit-u.ac.jp).

†Corresponding author: Hitotsubashi University, Graduate School of Economics, 2-1 Naka, Kunitachi, Tokyo, Japan 186-8601 (yohei.yamamoto@econ.hit-u.ac.jp).

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) in which they pay attention to the link between speculative bubbles and explosive behaviors of the asset price data.[1] Their strategy is to fit a univariate autoregressive (AR) model and test whether the root is greater than unity. While this paper is motivated by these studies, it explicitly accounts for an empirical fact that during the financial crisis of 2007 to 2008, speculative bubbles prevailed in various financial assets such as real estate, stocks, bonds, commodities, as well as exchange rates. It is important to investigate whether these bubbles are an economy-wide phenomenon or market-specific events. In order to answer such a question, we formally analyze how panel data of asset prices comove in an explosive environment.[2]

In this paper, we tackle the abovementioned problem by using the large-dimensional common factor model with principal component estimation. The common factor model is now a driving force to effectively investigate comovements of large panel data sets. Bai (2003) and Bai and Ng (2006) have discovered that when the series have no time trends, the principal component method provides a consistent estimate for the common and the idiosyncratic components. When the series have stochastic trends of integrating order one, the standard practice is to induce stationarity by transforming the original data by first-differencing prior to identifying and estimating the common and the idiosyncratic components. See Bai (2004) and the seminal work of Stock and Watson (2002, 2005) for empirical examples. One the other hand, if one is interested in identifying whether these stochastic trends lie in the common or in the idiosyncratic components, Bai and Ng (2004) suggest applying the augmented Dickey–Fuller (hereafter, "ADF") tests (Dickey and Fuller, 1979) for either the common or the idiosyncratic components estimated by the first-differenced data. This method is referred to as the panel analysis of nonstationarity in idiosyncratic and common component (hereafter, "PANIC"). A strong advantage of this method is the common and the idiosyncratic components being separately identified under the null hypothesis of random walk by using the first-differenced data. More precisely, the standard ADF tests can be used

[1] Phillips et al. (2011) and Phillips and Yu (2011) developed a method for a single bubble. Phillips et al. (2015ab) modified it to account for multiple bubbles. For other testing methods, see Gürkaynak (2008) for a survey.

[2] When the cross-sectional dimension is not large, Phillips and Magdalinos (2008) developed a limit theory for least squares estimation of the cointegration system among explosive time series.

for the null hypothesis of random walk against the alternative hypothesis of stationarity (hereafter, the"stationary test"), because under the alternative hypothesis of stationarity, the first-differenced series is over-differenced but it remains stationary so that the tests have power. Further, Bai and Ng's (2004) simulation study shows that the common test has a good size and power regardless of the idiosyncratic components being stationary or random walk. The same can be said of the idiosyncratic tests. Therefore, the PANIC approach successfully disentangles the common and the idiosyncratic components.

The central question of this paper is whether this convenient property of the PANIC approach is available even when one tests against the alternative hypothesis of an explosive process, that is, the right-tailed version of the ADF test (hereafter,"the explosive test"). To this end, we first confirm that the common and the idiosyncratic explosive tests constructed by the PANIC method have local asymptotic power as the standard tests do. This is a local analysis in the sense that the first-order autoregressive coefficient is assumed to shrink to one at rate T^{-1}, where T is the time dimension of the panel data set. This particular rate enables the test statistics to have a limiting distribution under the alternative hypothesis and provides a meaningful approximation to the finite sample power. Most importantly, the local power result applies to either the stationary or the explosive tests.

A potential problem of the local asymptotic framework is that it only considers small deviations from the unit root. Recently, it is understood that the asymptotic results under the local asymptotic framework may not adequately approximate the finite sample behaviors of the test statistics if the true parameter value is distant from the null hypothesis (see, e.g., Deng and Perron, 2008). With this caveat in mind, we take a nonlocal approach that considers the autoregressive root that shrinks to one at a slower rate than T^{-1}. In particular, we use the moderate deviations framework developed by Phillips and Magdalinos (2007). Importantly, under this framework, we find that explosive idiosyncratic components may be identified as a common factor if some idiosyncratic components are explosive in a non-local order. This leads to the fact that the common and the idiosyncratic tests have size distortions and deterioration of power.

A Monte Carlo simulation illustrates our theoretical results. We first confirm that as long as the stationary test is concerned, the PANIC approach provides very good size and power in every case in study. This is consistent with Bai and Ng (2004). However, despite our local asymptotic results, the explosive tests behave very differently from the stationary tests when the process is moderately or strongly explosive in finite samples as follows. First, the common test shows significant size distortions when some idiosyncratic components are explosive. This is because the common factor is now identified as the moderately explosive individual response variables. Second, when the common component is explosive, the idiosyncratic test also suffers from size distortions for the same reason. Finally and most importantly, the idiosyncratic test shows an upward power function as long as it is weakly explosive, as supported by our local asymptotic result. However, the power function starts to decline at some point and then may reach zero as the explosiveness is strengthened, as our moderate deviations asymptotic result discovers. This phenomenon is the well-known nonmonotonic power problem that is widely documented in the context of structural change tests (see Perron, 1991 and Vogelsang, 1999). What is new in this paper is that the source of nonmonotonic power is the identification failure between the common factors and explosive idiosyncratic errors. Further, as far as the authors know, this is the first study that documents the nonmonotonic power problem in the unit root test.

Finally, we attempt to provide a new method to test for speculative bubbles in the common and the idiosyncratic components. This method takes advantage of the fact that in many empirical situations, bubbles appear only in a certain subperiod and the series are not explosive in the rest of the sample period. Therefore, we can set a training sample during which no or only weak explosive behaviors exist. We then use cross-sectional regressions to estimate the common components in the explosive window as coefficients attached to the factor loadings, while the factor loadings are estimated in the training sample. We call this the cross-sectional (hereafter, "CS") method. It is shown that both the common and idiosyncratic tests achieve the correct size asymptotically and are consistent under the moderate deviations framework. A Monte Carlo simulation shows that the CS common test considerably reduces the size distortions. More importantly, the CS idiosyncratic test is robust to the nonmonotonic power problem.

The usefulness of the proposed approach is illustrated in an empirical example of the exchange-rate system during the financial crisis of 2007 to 2008. We use 24 bilateral exchange rates against the U.S. dollar (USD) from August 1, 2007 to January 31, 2009. It is clearly observed that many of them exhibit extreme movements in this period. The question of interest is whether these explosive behaviors are due to the common components that are ascribed to the value of the USD or the idiosyncratic components that pertain to the values of the paired individual currencies. We find that the common component is explosive in both the PANIC and the CS tests. Hence, the USD exhibits an explosive behavior in this period. As for the idiosyncratic components, the PANIC tests are insignificant except for the Indonesian rupiah. This may suggest that the widespread explosive behaviors in the exchange-rate system during the financial crisis are mainly attributed to the USD and most other currencies are stable. However, when the CS approach is used, eight currencies are judged explosive in their idiosyncratic components at the 10% significance level. These include the so-called safe haven currencies: the Japanese yen and the Swiss franc. Therefore, the CS approach provides additional perspectives for widely observed explosive behaviors in the exchange-rate system during the financial crisis period.

The structure of the paper is as follows. Section 2 introduces the model, assumptions, and the existing PANIC tests. Section 3 uses a local and a nonlocal asymptotic framework to investigate the theoretical properties of the right-tailed PANIC tests. Section 4 studies the finite sample size and power of the PANIC tests via Monte Carlo simulations. In section 5, we propose a new CS method to disentangle the common and the idiosyncratic components. In addition, the theoretical and finite sample properties of the CS method are investigated in this section. Section 6 gives an empirical example to illustrate the usefulness of the proposed method and section 7 concludes the paper. Throughout the paper, the following notations are used. The Euclidean norm of vector x is denoted by $\|x\|$. For matrices, the vector-induced norm is used. The symbols $O(\cdot)$ and $o(\cdot)$ denote the standard asymptotic orders of sequences. The symbol $\xrightarrow{p}$ represents convergence in probability under the probability measure P and the symbol $\Rightarrow$ denotes convergence in distribution. $O_p(\cdot)$ and $o_p(\cdot)$ are the orders of convergence in probability under P.

2 Model and test statistics

We consider the common factor model

$$X_{it} = \lambda_i' F_t + U_{it}, \quad \text{for } i = 1, ..., N \text{ and } t = 1, ..., T, \tag{1}$$

where X_{it} is a scalar of the observed random variable, F_t and λ_i are $r \times 1$ vectors of common factors and factor loadings, respectively, and U_{it} is a scalar of idiosyncratic errors. In this paper, we focus on the essence of problem by considering the case of $r = 1$ without losing any substance. The common factor follows the first-order autoregressive (AR1) process so that

$$F_t = \alpha F_{t-1} + e_t, \tag{2}$$

where α is an autoregressive coefficient and e_t is a white noise disturbance. Furthermore, the idiosyncratic errors follow AR1 processes

$$U_{it} = \rho_i U_{i,t-1} + z_{it}, \tag{3}$$

where ρ_i is the autoregressive coefficient of the ith cross-section and z_{it} is a white noise disturbance. Let $F_0 = e_0$ and $U_{i0} = z_{i0}$.

We consider the following assumptions on this model. Let $M < \infty$ be a generic constant.

Assumption 1. For every $t = 0, 1, ..., T$, $e_t \sim i.i.d.(0, \sigma^2)$ and $E\left|e_t\right|^4 \leq M$.

Assumption 2.

(a) λ_i is a nonrandom quantity satisfying $|\lambda_i| \leq M$ or a random quantity satisfying $E\left|\lambda_i\right|^4 \leq M$.

(b) $N^{-1}\sum_{i=1}^{N}\lambda_i^2 \xrightarrow{p} \sigma_\lambda^2$, where σ_λ is a positive constant.

Assumption 3. For every $t, s = 0, 1, ..., T$, the following holds.

(a) $z_{it} \sim i.i.d.(0, \sigma_i^2)$ and $E\left|z_{it}\right|^8 \leq M$.

(b) Let $\gamma_{st} = N^{-1}\sum_{i=1}^{N}E(z_{is}z_{it})$. Then, $|\gamma_{ss}| \leq M$ and $T^{-1}\sum_{s=1}^{T}\sum_{t=1}^{T}|\gamma_{st}| \leq M$ for all s and t.

(c) Let $\phi_{ij} = E(z_{it}z_{jt})$. Then, $\sum_{i=1}^{N}\left|\phi_{ij}\right| \leq M$ for all j and $N^{-1}\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\phi_{ij}\right| \leq M$.

(d) Let $\zeta_{st} = E\left|N^{-1/2}\sum_{i=1}^{N}[z_{is}z_{it} - E(z_{is}z_{it})]\right|^4$. Then, $\zeta_{st} \leq M$.

Assumption 4. z_{is}, e_t, and λ_j are mutually independent for every (i, j, s, t).

The model and assumptions follow those of Bai and Ng (2004), who considered the unit root test against the alternative hypothesis of stationarity for the common and the idiosyncratic components. In this paper, we are interested in the test against the alternative hypothesis of explosive process. For the common component,

$$H_0 : \alpha = 1 \qquad \text{versus} \qquad H_1 : \alpha > 1. \tag{4}$$

We are also interested in the hypotheses for the ith idiosyncratic component,

$$H_0 : \rho_i = 1 \qquad \text{versus} \qquad H_1 : \rho_i > 1. \tag{5}$$

At first glance, this testing problem has already been explored. Under a restriction of $\alpha = 1$, the model is essentially the same as Bai and Ng's (2004) PANIC. They propose a method of separately identifying the common factors and the idiosyncratic errors under the null hypothesis of the common factors following random walks. It is based on the first-differenced data so that

$$x_{it} = \lambda_i f_t + u_{it}, \tag{6}$$

where $x_{it} = X_{it} - X_{i,t-1}$, $f_t = F_t - F_{t-1}$, and $u_{it} = U_{it} - U_{i,t-1}$. In the following, we assume that there are $T+1$ observations $t = 0, 1, ..., T$ for X_{it} (so that F_t and U_{it}) for notational simplicity. The common factors and factor loadings can be estimated using x_{it} by the principal component method so that

$$(\hat{f}_t, \hat{\lambda}_i) = \arg \min_{\{\lambda_i\}_{i=1}^N, \{f_t\}_{t=1}^T} \sum_{i=1}^N \sum_{t=1}^T (x_{it} - \lambda_i f_t)^2, \tag{7}$$

with normalization $T^{-1} \sum_{t=1}^T \hat{f}_t^2 = 1$. This minimization problem gives a common factor estimate $\hat{f} = [\hat{f}_1, ..., \hat{f}_T]'$ as the $\sqrt{T}$-times eigenvectors of xx' corresponding to the largest eigenvalue, where x is a $T \times N$ matrix with the (t,i)th element being x_{it}. The factor loadings are estimated by $\hat{\lambda}_i = \frac{1}{T} \sum_{t=1}^T \hat{f}_t x_{it}$. Furthermore, the level common factor is estimated by $\hat{F}_t = \sum_{s=1}^t \hat{f}_s$ and the level idiosyncratic errors are estimated by $\hat{U}_{it} = \sum_{s=1}^t \hat{u}_{is}$, where $\hat{u}_{is} = x_{is} - \hat{\lambda}_i \hat{f}_s$.

The unit root test for the common component (hereafter, the "common test") can be implemented by using a t-test for $H_0 : \delta = 0$ in the regression

$$\hat{f}_t = \delta \hat{F}_{t-1} + error, \tag{8}$$

so that

$$t_{\hat{F}} = \frac{\hat{\delta}}{se(\hat{\delta})}, \tag{9}$$

where $\hat{\delta}$ is an OLS estimator for δ and $se(\hat{\delta})$ represents its standard errors. Note that the regression can potentially include an intercept and a time trend with appropriate references to critical values of Dickey and Fuller (1979). When the errors are suspected to be serially correlated, we can also include the lags of $\hat{f}_t$ in the regression. However, the model with no lags is relevent in asset price data in which no serial correlations are present in their first differences.[3] If necessary, we can extend the fremework to the model with p lags under appropriate conditions for p as in Bai and Ng (2004).

The unit root test for the ith idiosyncratic component (referred to as the "idiosyncratic test") is implemented by using a t-test for $H_0 : \delta_i = 0$ in the regression

$$\hat{u}_{it} = \delta_i \hat{U}_{i,t-1} + error, \tag{10}$$

so that

$$t_{\hat{U}}(i) = \frac{\hat{\delta}_i}{se(\hat{\delta}_i)}, \tag{11}$$

[3]Phillips and Yu (2011) also consider only the model with $p = 0$.

where the same note as $t_{\hat{F}}$ applies.

As Bai and Ng (2004) point out this approach is convenient because the common and the idiosyncratic components are separately identified by using the first-differenced data, so that the test statistics (9) and (11) have the standard Dickey and Fuller's (1979) distribution under the null hypothesis. If the alternative hypothesis of stationarity is true, the series become over-differenced, but they remain stationary so that the test has nontrivial power. Further, their simulation study shows that the common test has a good size and power regardless of the idiosyncratic components being stationary or random walk. The same can be said of the idiosyncratic test. Therefore, the PANIC approach successfully disentangles the common and idiosyncratic components.

Remark 1 (Bai and Ng, 2004) *Let Assumptions 1–4 hold. (i) Under the null hypothesis of* $\alpha = 1$ *together with* $|\rho_i| \leq 1$ *for all* i*,*

$$t_{\hat{F}} \Rightarrow \frac{\int_0^1 W(r)dW(r)}{\left[\int_0^1 W(r)^2 dr\right]^{1/2}},$$

as $N, T \to \infty$*, where* $W(r)$ *is the standard Wiener process defined on* $r \in [0,1]$*. (ii) Under the null hypothesis of* $\rho_i = 1$ *together with* $\alpha = 1$ *and* $\left|\rho_j\right| \leq 1$ *for all* $j \neq i$*,*

$$t_{\hat{U}}(i) \Rightarrow \frac{\int_0^1 W_i(r)dW_i(r)}{\left[\int_0^1 W_i(r)^2 dr\right]^{1/2}},$$

as $N, T \to \infty$*, where* $W_i(r)$ *is the standard Wiener process defined on* $r \in [0,1]$*.*

This result is trivially applicable to our right-tailed testing problems (4) and (5) if the researcher knows $\alpha = 1$ and $\rho_i = 1$ for all i. However, the right-tailed test must confirm that the null distribution is available in (i) $\alpha = 1$ so that the null is true but $|\rho_i| > 1$ for some i. This is because the first-difference series x_{it} is no longer stationary and the consistent factor estimate may be unavailable. We must also ensure that $\rho_i = 1$ in case (ii), so that the null is true but $|\alpha| > 1$ and/or $\left|\rho_j\right| > 1$ for some $j \neq i$. More importantly, the above results do not provide any information on the power of the common and the idiosyncratic tests.

3 Theoretical results

Bai and Ng (2004) investigate the finite sample properties of the unit root tests (9) and (11) but against the stationary alternative hypothesis via Monte Carlo simulations. They show that the common test has a good size and nontrivial power regardless of the idiosyncratic components being stationary or random walk. The same can be said of the idiosyncratic test. However, the theoretical power properties of these tests under the explosive alternative hypothesis have not been discussed. To discover this issue, this section conducts theoretical investigations on the power properties of the explosive tests by using two asymptotic frameworks. The first approach is a local alternative framework and its result is expected to capture the finite sample properties of the test when the explosiveness is weak.[4] However, it is well-known that local asymptotic frameworks often fail to provide good approximations on the finite sample behavior of the test when the true parameter value is distant from the null. Therefore, we also use a nonlocal asymptotic framework, in particular, the moderate deviations framework developed by Phillips and Magdalinos (2007), to investigate situations of stronger explosiveness.

3.1 Results under the local deviations framework

We first show that the asymptotic local power of the PANIC tests is obtained for the common and the idiosyncratic components. The following assumption is considered in this subsection.

Assumption 5. The autoregressive coefficients satisfy $\alpha = 1 + c/T$ and $\rho_i = 1 + c_i/T$, where c and c_i are fixed constants.

As in the literature, Assumption 5 considers the local alternative hypothesis that shrinks to the null hypothesis at rate T^{-1}. The noncentrality parameters c and c_i could either be positive or negative, where $c > 0$ and $c_i > 0$ consider the explosive tests and $c < 0$ and $c_i < 0$ pertain to the stationary tests. Therefore, the local asymptotic result is symmetric in the sense that it is valid either against the explosive alternative hypothesis or the stationary alternative hypothesis. More importantly, the specific rate T^{-1} helps us derive the limiting distributions of the test statistics[5] and the autoregressive coefficient estimators as follows.

Theorem 1 *Suppose that Assumptions 1–5 hold. Let $W(r)$ and $W_i(r)$ be independent standard Wiener processes and $W_c(r)$ and $W_{ci}(r)$ be independent Ornstein and Uhlenbeck processes defined on $r \in [0, 1]$.*

[4] It is also pointed out that the result is symmetric in the sense that it applies to either stationary or explosive test.

[5] See Phillips (1988), Perron (1989), and Nabeya and Tanaka (1990) for detail.

(i-a) If $c = 0$ and $c_i \geq 0$ for any i, then

$$t_{\hat{F}} \Rightarrow \frac{\int_0^1 W(r)dW(r)}{\left[\int_0^1 W(r)^2 dr\right]^{1/2}},$$

as $N, T \to \infty$.

(i-b) If $c > 0$ and $c_i \geq 0$ for any i, then

$$T\hat{\delta} \Rightarrow c + \frac{\int_0^1 W_c(r)dW(r)}{\int_0^1 W_c(r)^2 dr},$$

as $N, T \to \infty$.

(ii-a) If $c_i = 0$, $c_j \geq 0$ for any $j \neq i$, and $c \geq 0$, then

$$t_{\hat{U}}(i) \Rightarrow \frac{\int_0^1 W_i(r)dW_i(r)}{\left[\int_0^1 W_i(r)^2 dr\right]^{1/2}},$$

as $N, T \to \infty$.

(ii-b) If $c_i > 0$, $c_j \geq 0$ for any $j \neq i$, and $c \geq 0$, then

$$T\hat{\delta}_i \Rightarrow c_i + \frac{\int_0^1 W_{c,i}(r)dW_i(r)}{\int_0^1 W_{c,i}(r)^2 dr},$$

as $N, T \to \infty$.

We provide a proof of this theorem in the appendix. This theorem confirms the following facts. First, parts (i-a) and (ii-a) show that the size of the common (idiosyncratic) test is robust to local deviations in the idiosyncratic (common and other idiosyncratic) components. As for the power, parts (i-b) and (ii-b) ensure that the common (idiosyncratic) test has the standard local power even though the idiosyncratic (common and other idiosyncratic) components deviate from the random walk as long as the deviations are local. Therefore, this theorem theoretically confirms Bai and Ng's (2004) Monte Carlo findings in both stationary and explosive tests and implies that the PANIC can disentangle the common and the idiosyncratic explosive components as well. However, the potential problem of this asymptotic framework is that it only considers small deviations from the unit root.

3.2 Results under the moderate deviations framework

It is well-known that the asymptotic results under local asymptotic frameworks may not adequately approximate the finite sample behaviors of the test statistics. For example, in the context of structural change testing in linear regression models, a certain type of test statistics may have good power when the magnitude of change is assumed to quickly shrink to zero as the sample size increases, but they lose power when the break is assumed to be fixed. In finite samples, this class of tests typically draws a concave-shaped power function, called the nonmonotonic power problem.[6] One reason of this phenomenon is that under the alternative hypothesis, a change in the conditional mean and a change in the persistence parameter are not separately identified. Yamamoto and Tanaka (2015) extend this idea to structural change tests in factor loadings, pointing out that the factor loading structural change and the extra common factors may not separately be identified under the alternative hypothesis of common breaks. In such a case, the standard tests of Breitung and Eickmeier (2011) suffer from the nonmonotonic power problem.

The purpose of this subsection is to theoretically explain why the PANIC tests potentially have size distortions and the nonmonotonic power problem despite our local asymptotic results. Here, we claim that an identification problem between the common factors and the explosive idiosyncratic errors occurs under the alternative hypothesis of explosive idiosyncratic components. To this end, we take a nonlocal approach that assumes the explosive root shrinking to one at a slower rate than T^{-1}. In particular, we use the moderate deviations framework developed by Phillips and Magdalinos (2007).

Assumption 6. The autoregressive coefficients satisfy $\alpha = 1 + \frac{c}{k_T}$ and $\rho_i = 1 + \frac{c_i}{k_T}$, where $c \geq 0$, $c_i \geq 0$ and k_T is a deterministic sequence such that $k_T \to \infty$ and $k_T = o(T)$.

The quantities c and c_i ($i = 1, ..., N$) are again noncentrality parameters but now take a nonnegative value to focus on the explosive case. Of interest is the fact that the scaling factor k_T is an arbitrary deterministic function of T that satisfies $k_T \to \infty$ strictly slower than T. This way, we can consider stronger explosiveness than that in the local assumption. A typical formulation is $k_T = T^{\kappa}$, where $0 < \kappa < 1$.

[6] As far as the authors know, Perron (1991) is the first paper to point out this problem in structural change tests. See Vogelsang (1999), Perron and Yamamoto (2016), and the references therein.

Under this nonlocal setting, the principal component estimate may not appropriately identify the common components. We illustrate this fact in the following theorem by considering two cases. The first case assumes that $c > 0$ but $c_i = 0$ for all i, so that only the common factor is explosive. The second case is $c_i > 0$ for some or all i but $c = 0$, so that only the idiosyncratic errors are explosive.

Theorem 2 *Let Assumptions 1–4 and 6 hold. If k_T grows slow enough so that $\alpha^T T^{-1/2}$ and $\rho_i^T T^{-1/2}$ go to infinity, as $T \to \infty$, then the following equation holds for the factor estimate*

$$V\hat{f}_t = Af_t + Bf_t + N^{-1}\textstyle\sum_{i=1}^N a_i u_{it} + N^{-1}\sum_{i=1}^N b_i u_{it}, \tag{12}$$

where V is the largest eigenvalue of $N^{-1}T^{-1}xx'$, and the quantities A, B, a_i, and b_i satisfy the following properties.

(i) If $c > 0$ and $c_i = 0$ for all i, then

$$\begin{aligned} A &= O_p(\alpha^T T^{-1/2}),\ B = O_p(1), \\ a_i &= O_p(\alpha^T T^{-1/2}),\ \text{and } b_i = O_p(1). \end{aligned}$$

(ii) If $c = 0$ and $c_i > 0$, then

$$\begin{aligned} A &= O_p(1),\ B = O_p(\rho_i^T T^{-1/2}), \\ a_i &= O_p(1),\ \text{and } b_i = O_p(\rho_i^T T^{-1/2}). \end{aligned}$$

In a nutshell, the principal component estimate may have an asymptotic bias if the true factor or idiosyncratic errors are moderately explosive. This is in contrast to the local case where the factors are always consistently estimated. In the following, let the bound for V given in (A.12) be tight. Part (i) shows that when the true factor is explosive, the asymptotic bias consists of the first and the third terms of (12), because the weights $V^{-1}A$ and $V^{-1}a_i$ do not vanish, whereas $V^{-1}B$ and $V^{-1}b_i$ do. The implication of these results to the unit root tests is clear. The factor estimate is a sum of the true explosive factors contaminated by nonexplosive idiosyncratic errors. However, since the true factor dominates this bias component, the factor estimate remains explosive and the power of the common test remains. As for the idiosyncratic test, the idiosyncratic component estimate comprises the residuals from the regression using the biased factor estimate as a regressor. Therefore, the residuals are also biased and the idiosyncratic test may suffer from size distortions.

On the other hand, part (ii) considers the case where the idiosyncratic errors are explosive. Here, the factor estimate is contaminated in a different way. The nonvanishing asymptotic bias now consists of the second and the fourth terms in (12), where the weights $V^{-1}B$ and $V^{-1}b_i$ are bounded. Now, the latter becomes a dominant component in the factor estimates, because the true factor is not explosive while the idiosyncratic errors are so. Therefore, we identify the weighted average of explosive idiosyncratic errors as a common factor. This has an interesting insight for the common factor estimate in an explosive environment because even if the factor is not explosive, as long as the data includes explosive idiosyncratic components, they may be identified as a common factor.

We can also derive clear implications of part (ii) to the testing. First, when some idiosyncratic errors are explosive, the size of the common test would be distorted because the factor estimate is now dominated by the explosive idiosyncratic errors. More interestingly, since the factor estimate consists of explosive idiosyncratic errors, the residuals from the regression of an explosive response variable on the series including itself become nonexplosive even though the true process is explosive. This causes a power loss of the idiosyncratic tests as the idiosyncratic errors become more explosive.

Remark 2 *We illustrate the power loss of the idiosyncratic test by taking a special case of (ii) where only the ith cross-sectional unit has explosive idiosyncratic errors. We have*

$$\hat{f}_t \approx b_i u_{it}, \tag{13}$$

where "$\approx$" denotes asymptotic equality. In such a case,

$$\begin{aligned}
\hat{\lambda}_i &= \left(\textstyle\sum_{t=1}^T \hat{f}_t^2\right)^{-1}\left(\textstyle\sum_{t=1}^T \hat{f}_t x_{it}\right), \\
&\approx \left(b_i^2 \textstyle\sum_{t=1}^T u_{it}^2\right)^{-1}\left(b_i \textstyle\sum_{t=1}^T u_{it} x_{it}\right), \\
&= \left(b_i^2 T^{-1} \textstyle\sum_{t=1}^T u_{it}^2\right)^{-1}\left(b_i \lambda_i T^{-1} \textstyle\sum_{t=1}^T u_{it} f_t + b_i T^{-1} \textstyle\sum_{t=1}^T u_{it}^2\right), \\
&\approx b_i^{-1},
\end{aligned} \tag{14}$$

because the numerator of the third line is dominated by the second term. Then,

$$\begin{aligned}
\hat{u}_{it} &= u_{it} + \lambda_i f_t - \hat{\lambda}_i \hat{f}_t, \\
&\approx u_{it} + \lambda_i f_t - u_{it}, \\
&= \lambda_i f_t,
\end{aligned} \tag{15}$$

by using (13) and (14), so that the idiosyncratic residuals inherit the time-series properties of the true common factor. Hence, the idiosyncratic test loses power.

Remark 3 *We illustrate the identification problem between the factors and errors in the next section via Monte Carlo simulation.*

4 Finite sample size and power of the PANIC tests

This section investigates the finite sample properties of the PANIC tests via Monte Carlo simulations. Although our main focus is the empirical size and power of the explosive test, those of the stationary test are also presented for reference. Further, while the latter experiment overlaps Bai and Ng's (2004) simulation results, it is instructive to illustrate how differently the explosive and stationary PANIC tests behave. The data is generated by the models (1), (2), and (3) with λ_i, u_{it}, z_{it}, F_0, and U_{0i} independently drawn from the standard normal quasi random variables in each replication. In order to evaluate the size and power, we vary the values of α and ρ_i from 1.0 to 1.01 for the explosive test and from 1.0 to 0.0 for the stationary test. Based on the same data, the results using the regression models that include (A) no intercept and no time trend, (B) an intercept but no time trend, and (C) an intercept and a linear time trend are reported. The common test uses the estimated factor and the idiosyncratic test presents the size and power of the first response variable, that is, $i = 1$, but this is without loss of generality because the Monte Carlo design is symmetric for any i. Each panel of Figures 1 to 5 contains two lines. The solid line is the result using the estimated common and idiosyncratic components and they are labeled as "estimated.". The dotted line is the result that hypothetically uses the true common and idiosyncratic components and is labeled as "observed" for reference. We use $N = 100$ and $T = 100$, unless otherwise stated. The number of replications is 5,000.

We first consider the size of the tests by using the 5% nominal level. Figure 1 and Figure 2 report the size of the common test as a function of ρ_i and the size of the idiosyncratic test as a function of α, respectively. In both figures, the left-hand-side panels show the size of the explosive test and the right-hand-side panels show the size of the stationary test. The size of the stationary test is very good in every case, in the same line as Bai and Ng (2004). As they claim, the PANIC approach successfully disentangles the common and idiosyncratic components and the common and idiosyncratic tests can be implemented without worrying about the behavior of the other component. However, if we look at the size of the explosive test, the results are very different. The left-hand-side panels of Figure 1 show that the size of the common test is close to the nominal level when the idiosyncratic component is not very explosive (ρ_i is approximately smaller than 1.002); however, the size quickly reaches one as ρ_i increases. Further, as shown in the left-hand-side panels of Figure 2, the size of the idiosyncratic test is also distorted towards zero as α increases. These size distortions are expected from Theorem 2 under the moderate deviations framework. However, it is clearly suggested that the convenient property of Bai and Ng (2004) no longer applies to the explosive test.

Next, we consider the power of the tests at the 5% nominal level. Figure 3 reports the empirical power functions of the common test. Again, the left-hand-side and the right-hand-side panels correspond to the explosive and the stationary tests, respectively. Here, we present the power functions of the common test under $\rho_i = 1$ for all i, that is, no contamination from any idiosyncratic components. This is because the setting at $\rho_i > 1$ does not show any unique features of power of the common tests, except for the size distortions that are already reported in Figure 1.[7] The power functions of the stationary test are again standard as the right-hand-side panels of Figure 3 illustrate. This feature is consistent with Bai and Ng's (2004) finding. It is also observed that the power functions of the stationary and explosive alternative hypotheses are very close to the ones using the true common factor (dotted lines).

Of interest are the power functions of the idiosyncratic test in Figure 4. Again, we see the standard power functions for the stationary test in the right-hand-side panels. Although there are discrepancies between the powers of the stationary test using the estimated idiosyncratic component and the true idiosyncratic component, these are minor. On the other hand, the power functions for the explosive alternative hypothesis are very different from the ones for the stationary alternative hypothesis. They show a clear concave shape, which suggests the nonmonotonic power problem documented in the aforementioned literature. As we discussed in section 3, this can be explained as follows. When the explosive coefficient ρ_i is only slightly larger than one, the power increases as ρ_i diverges from one. This is consistent with our Theorem 1 and the Monte Carlo experiment illustrates that the test indeed has a local power. However, when the explosive coefficient becomes moderately larger than one, the explosive response variable is identified as a common factor and the residuals become nonexplosive as we discussed in Theorem 2 (ii). Hence, the power function starts to decrease eventually into zero as ρ_i increases. This property is empirically important, because the moderate or strong individual explosive behaviors may not be detected even though they are the ones that should not be overlooked compared to weak (or local) explosive behaviors.

So far, when we set $\rho_i > 1$, it applies to all i. An interesting question is whether these problems are mitigated when not all idiosyncratic components are explosive or only one is. To answer this question, Figure 5 investigates the following two cases: (a) only one

[7]Therefore, the power functions of the explosive test in the case of $\rho_i > 1$ start at a point above 0.05 but draw an upward curve. The starting point of the power function with $\rho_i > 1$ can be referred to in Figure 1.

idiosyncratic component is explosive ($\rho_1 \geq 1.0$, but $\rho_i = 1.0$ for all $i \neq 1$) and (b) ten idiosyncratic components are explosive ($\rho_i \geq 1.0$ for $i = 1, ..., 10$, but $\rho_i = 1.0$ for all $i \geq 11$). Case (a) is reported in the left-hand-side panels and case (b) is reported in the right-hand-side panels. For both cases, the top panels show the size of the common test for explosive alternative. We observe that the common test has considerable size distortions even when one idiosyncratic component is explosive. This is because only one explosive response variable is identified as a common factor. The size of the idiosyncratic test is presented in the middle panels. The idiosyncratic test shows size distortions even when the common factor is not explosive but some other idiosyncratic components are so. Again, this is because only one explosive individual response variable is identified as a common factor.[8] Finally, the bottom panels are the power functions of the idiosyncratic test for the first idiosyncratic component $t_{\hat{U}}(1)$. Similar to Figure 4, they clearly show nonmonotonic power here. Therefore, the size distortion and the power problem remain even if only one idiosyncratic component is explosive.

Finally, we validate the identification problem suggested in Theorem 2 (ii) by investigating the correlation between the estimated and true common factors, as well as the correlation between the estimated factor and the idiosyncratic errors. If the stated identification problem occurs, the former decreases but the latter increases as the idiosyncratic errors become more explosive. To this end, in Figure 6-1, we generate the same data as in the left-hand-side panels of Figure 4 and compute the average of (absolute value of) correlation coefficients between the estimated and true common factors $Corr(\hat{f}_t, f_t)$ as well as the correlation coefficient between the estimated factor and the idiosyncratic errors $Corr(\hat{f}_t, u_{1t})$ in the left-hand-side panel. In addition, we consider the case where only one idiosyncratic component has explosive errors as in the bottom left-hand-side panel of Figure 5. The results are presented in the right-hand-side panel of Figure 6-1. Both panels of Figure 6-1 clearly show that as the idiosyncratic errors become more explosive, the estimated factor is less correlated with the true factor but more correlated with the idiosyncratic errors. This is consistent with Theorem 2 (ii). Next, as equation (15) in Remark 2 suggests, we compute the average (absolute) correlation coefficients between the estimated and true idiosyncratic components $Corr(\hat{u}_{1t}, u_{1t})$ and the correlation between the estimated idiosyncratic components and the true factor $Corr(\hat{u}_{1t}, f_t)$. This is presented in Figure 6-2. Again, the left-hand-side panel considers the model with all the idiosyncratic components being explosive, while the right-hand-side panel corresponds

[8] Note that we can consider an idiosyncratic test for any observations $i \geq 11$ but particularly use the idiosyncratic test for the last observation $t_{\hat{U}}(N)$. This is again without loss of generality.

to the model with only one idiosyncratic unit having explosive errors. They show that as the idiosyncratic errors become more explosive, the estimated idiosyncratic components become less correlated with the true idiosyncratic errors but more correlated with the true common factor. Therefore, the estimated idiosyncratic components inherit the time-series properties of the true common factor, resulting in a power reversal.

5 Cross-sectional approach for testing speculative bubbles

5.1 Algorithm

This section attempts to provide a new method to test speculative bubbles in the common and the idiosyncratic components. The method is based on the following two key ingredients. First, it takes advantage of the fact that bubbles appear only in a certain subperiod and the series are not explosive in the rest of the sample period. If this is the case, we can time-wise localize the explosive behaviors by considering the models

$$F_t = \begin{cases} F_{t-1} + e_t, & \text{for } t = 1, ..., T \\ \alpha F_{t-1} + e_t & \text{for } t = T+1, ..., T+h \end{cases}, \tag{16}$$

and

$$U_{it} = \begin{cases} U_{i,t-1} + z_{it}, & \text{for } t = 1, ..., T \\ \rho_i U_{i,t-1} + z_{it} & \text{for } t = T+1, ..., T+h \end{cases}, \tag{17}$$

with h being the length of the window, so that the data is assumed to have a certain period $t \in [1, T]$ in which no explosive behaviors exist in either common or idiosyncratic components. We call this the training sample.[9] On the other hand, the period of interest $t \in [T+1, T+h]$ is called the explosive window. If identifying the timing of initiation of explosive window is an issue, an existing time-stamping method can be implemented for x_{it} series-by-series. See Phillips et al. (2011).

The second key element is using cross-sectional regressions to estimate the common factors in the explosive window instead of using the principal component estimation of the first-differenced series. This is because the first-differenced series of the explosive process remains explosive and violates Assumption 3. Hence, the common factors are not consistently estimated. To address this problem, we estimate the factor loadings in the training sample

[9] We can easily show that the weak (local) explosive process with the autoregressive coefficient $1 + \frac{c}{T}$ and $1 + \frac{c_i}{T}$ can exist in the training sample.

under a nonexplosive environment. Then, we use them as regressors in the cross-sectional regressions in the explosive window, to estimate the common components as the coefficients attached to the factor loadings. In this way, we can avoid the identification problem between the common and the explosive idiosyncratic components that are investigated in section 3.2. In doing this, we consider an asymptotic framework of $N, T \to \infty$, but h is fixed.[10] We call this approach the cross-sectional (CS) method and the steps are described as follows.

Algorithm:

Step 1. Use the first-differenced data x_{it} for $t = 1, ..., T$ to estimate the factor loadings λ_i by using the principal components method (7). Denote the factor loadings estimated in the training sample by $\hat{\lambda}_i^*$.

Step 2. At $t = T + 1$, estimate the level of common factors by the cross-sectional regression of $\{X_{it}\}_{i=1}^N$ on $\left\{\hat{\lambda}_i^*\right\}_{i=1}^N$ so that

$$\widetilde{F}_t = \left(\sum_{i=1}^N \hat{\lambda}_i^* \hat{\lambda}_i^{*\prime}\right)^{-1} \left(\sum_{i=1}^N \hat{\lambda}_i^* X_{it}\right),$$

and the idiosyncratic components by

$$\widetilde{U}_{it} = X_{it} - \hat{\lambda}_i^{*\prime} \widetilde{F}_t.$$

Then, repeat Step 2 for $t = T + 2, ..., T + h$.

Step 3. Construct the common test $t_{\widetilde{F}}^*$ by using $\widetilde{F}_t$ and $\widetilde{f}_t = \widetilde{F}_t - \widetilde{F}_{t-1}$ in the regression (8) and the idiosyncratic test $t_{\widetilde{U}}^*(i)$ by using $\widetilde{U}_{it}$ and $\widetilde{u}_{it} = \widetilde{U}_{it} - \widetilde{U}_{i,t-1}$ in the regression (10) for $t = T + 1, ..., T + h$.

We now discuss an asymptotic justification of the CS method in the following theorem. Note that the time dimension of the testing period is now h instead of T, and hence, we now denote $\alpha = 1 + \frac{c}{k_h}$ and $\rho_i = 1 + \frac{c_i}{k_h}$ in Assumption 7.

Theorem 3 *(i) Let Assumptions 1–4 hold. Under the null hypothesis of $\alpha = 1$,*

$$t_{\widetilde{F}}^* \Rightarrow \frac{\int_0^1 W(r) dW(r)}{\left[\int_0^1 W(r)^2 dr\right]^{1/2}},$$

[10] This is a standard assumption in the panel data model with a short time dimension, where the common factors are regarded as parameters. See, for example, Robertson and Sarafidis (2015).

and under the null hypothesis of $\rho_i = 1$,

$$t^*_{\widetilde{U}}(i) \Rightarrow \frac{\int_0^1 W_i(r) dW_i(r)}{\left[\int_0^1 W_i(r)^2 dr\right]^{1/2}},$$

as $h \to \infty$ *after* $N, T \to \infty$, *where* $W(r)$ *and* $W_i(r)$ *are independent standard Wiener processes defined on* $r \in [0, 1]$.

(ii) Let the OLS estimates for δ *and* δ_i *in Step 3 be* $\hat{\delta}^*$ *and* $\hat{\delta}^*_i$ *and Assumptions 1–5 hold. Under the alternative hypothesis of* $c \neq 0$,

$$h\hat{\delta}^* \Rightarrow c + \frac{\int_0^1 W_c(r) dW(r)}{\int_0^1 W_c(r)^2 dr},$$

and under the alternative hypothesis of $c_i \neq 0$,

$$h\hat{\delta}^*_i \Rightarrow c_i + \frac{\int_0^1 W_{c,i}(r) dW_i(r)}{\int_0^1 W_{c,i}(r)^2 dr},$$

as $h \to \infty$ *after* $N, T \to \infty$, *where* $W_c(r)$ *and* $W_{c,i}(r)$ *are independent Ornstein and Uhlenbeck processes defined on* $r \in [0, 1]$.

(iii) Let Assumptions 1–4 and 6 hold. Under the alternative hypothesis of $c > 0$ *and* $(1 + \frac{c}{k_h})^h \to \infty$,

$$\alpha^{-h} t^*_{\widetilde{F}} \Rightarrow \sqrt{\frac{c}{2\sigma^2}} |\Theta| > 0,$$

where $\Theta \equiv N(0, \sigma^2/2c)$ *and under the alternative hypothesis of* $c_i > 0$ *and* $(1 + \frac{c_i}{k_h})^h \to \infty$,

$$\rho_i^{-h} t^*_{\widetilde{U}}(i) \Rightarrow \sqrt{\frac{c_i}{2\sigma_i^2}} |\Theta_i| > 0,$$

where $\Theta_i \equiv N(0, \sigma_i^2/2c_i)$ *as* $h \to \infty$ *after* $N, T \to \infty$.

This theorem shows that the common test asymptotically achieves the correct size and is consistent under the moderate deviations framework. The idiosyncratic test also attains the correct size asymptotically and it is consistent under the moderate deviations framework. Notice that these results are obtained only under a sequential limit: first N and $T \to \infty$, then $h \to \infty$. The proof is as follows. In the first step ($N, T \to \infty$), we obtain the consistent estimate for the common factors. In the second step ($h \to \infty$), we can simply replace the estimated factors with its true counterparts to construct the tests. As the literature shows, this is a stronger assumption than the limit obtained under N, T, and $h \to \infty$ simultaneously. Hence, we investigate whether this asymptotic approximation reasonably approximates finite sample behaviors of the tests in the following subsection.

5.2 Finite sample properties of the CS method

This subsection investigates the finite sample property of the CS method via Monte Carlo simulation. The data is generated by (1), (16), and (17) with $r = 1$. All of λ_i, u_{it}, z_{it}, F_0, and U_{0i} are independently drawn from the standard normal quasi random variables in each replication. The sample size is N, $T = 100$ and we consider three cases for the length of explosive window $h = 10$, 20, and 30. The size and power of the following tests at the 5% nominal level are computed through 5,000 replications: the CS tests and the PANIC tests using the explosive window. Further, we present the tests by using the true common and idiosyncratic components for the explosive window, labeled as "observed." The "observed" is a counterfactual experiment and would merely give us a benchmark to evaluate performance of the CS and the PANIC tests.

Figure 7 presents the size of the common and idiosyncratic tests in the left-hand-side and right-hand-side panels, respectively. The top, middle, and bottom panels correspond to $h = 10$, 20, and 30 cases. Consistent with our findings in section 4, the PANIC common test has serious size distortions when the idiosyncratic components are explosive and the PANIC idiosyncratic test becomes undersized when the common component is explosive. Although the CS common test also shows size distortions, they are considerably smaller than the PANIC tests. As for the CS idiosyncratic test, we now see over rejections especially when h is large. Figure 8 reports the power of both tests in the same format as Figure 7. The power functions of the CS and PANIC common tests are very similar and almost equivalent to the counterfactual test. Most importantly, the right-hand-side panels of Figure 8 show the power of idiosyncratic tests and suggest that the CS idiosyncratic test is robust to the nonmonotonic power problem. In summary, the CS test suffers from size distortions that are not present in the PANIC test when we have a relatively long explosive period. However, the CS test performs well in general with a short explosive window and outweighs the PANIC approach with respect to power in all cases.

6 Empirical example

Towards the financial crisis of 2007 to 2008, speculative bubbles prevailed in various financial assets such as real estate, stock, bond, and commodity prices. This has been documented in empirical literature; for example, Phillips and Yu (2011) consider examples of a home price index, the crude oil price, and the spread between Baa and Aaa bond rates in the United States. However, relatively less attention has been paid to explosive behaviors in exchange rates. One difficulty arises in testing bubbles in this case because when a bilateral exchange rate, say the U.S. dollar (USD) measured by the Euro, exhibits a bubble, we hardly tell which currency or both of them are susceptible to speculative investments. Here, we try to shed light on this question by using the common factor approach considered in this paper. Applying the common factor model to the exchange-rate system recently becomes increasingly popular from various viewpoints. For example, Lustig et al. (2011) use the common factor model for exchange-rate returns to identify the risk factors. Greenaway-McGrevy et al. (2015) tries to identify the value of a single currency, that is, USD and Euro factors, based on the principal component estimate. Engel et al. (2015) use the common factors extracted from a cross-sectional data of bilateral exchange returns for forecasting.

We construct a currency portfolio roughly by using an intersection of the aforementioned three papers.[11] Figure 9 plots 24 bilateral exchange rates against the USD in logarithm from January 1, 2004 to January 31 2009, where their respective titles are the ISO4217 currency codes that are fully described in Table 1. Following the literature, we set the financial crisis period from August 1, 2007 to January 31, 2009 (the sample size $h = 393$). It is clearly observed that many bilateral series exhibit extreme behaviors in this period. Most currencies depreciated against the USD (increased in figure), while the JPY is the one that rapidly appreciated against the USD (decreased in figure). We then performed a battery of tests considered in this paper. The results using the regression with an intercept but no time trend are reported in Table 1.[12] We first test whether these explosive behaviors of individual bilateral rates are statistically significant series-by-series. The column "individual" shows the right-tailed DF test applied to the bilateral rates in the financial crisis period. We see that 13 bilateral exchange rates are significantly explosive at the 10% level. This is consistent with our visual inspection of Figure 9.

The question of interest is whether these explosive behaviors are due to the common or the idiosyncratic component. According to the literature, the common components are ascribed to the value of the USD and the idiosyncratic components pertain to the values of the paired individual currencies. To this end, we employ the PANIC tests for the explosive period from August 1, 2007 to January 31, 2009. Further, we perform the CS tests for the same explosive period by using January 1, 2004 to July 31, 2007 as the training sample. We

[11] By following Lustig et al. (2011), we included some currencies that are partly pegged to the USD, for example, the Hong Kong dollar and the Singapore dollar.

[12] The results using the augmented regression with an intercept and a linear time trend are qualitatively similar, and hence, they are not separately reported.

present the results by using two common factors following the literature[13]; however, other variations of the number of factors provide qualitatively similar results.

The first common factor is significantly explosive in both the PANIC and the CS tests at the 5% level. It is concluded that the USD exhibits bubbles in this period, although there is a caveat that the common tests may have size distortions. The second factor is not explosive both in the PANIC and the CS tests. Therefore, the two approaches reach the same conclusion for the common component. Let us move on to the idiosyncratic tests. When we use the PANIC approach, all the idiosyncratic tests are insignificant, except for the Indonesian rupiah. If we rely on this result, it is concluded that the widespread explosive behaviors in exchange rates are mainly attributed to the USD and most other currencies are stable. However, if we use the CS method, the results are strikingly different. We see that eight currencies are significant at the 10% level, three currencies out of them are significant at the 5% level. This finding suggests that the explosive behaviors in the exchange-rate system are because of not only the USD but also other currencies. For instance, idiosyncratic tests of the Japanese yen and the Swiss franc are significant at the 5% level. This may come from the speculative demand for the safe-haven properties of these currencies as pointed out by Habib and Stracca (2012) and Fatum and Yamamoto (2015). Therefore, we discovered that explosive behaviors prevailed in the exchange-rate system during the financial crisis period is not solely attributed to the value of the USD but other currencies as well. This example shows that information extracted by the CS tests and the PANIC tests can be very different and the CS approach provides additional perspectives.

7 Conclusions

Towards the financial crisis of 2007 to 2008, speculative bubbles prevailed in various financial assets such as real estate, stocks, bonds, commodities, as well as exchange rates. Whether these bubbles are an economy-wide phenomenon or market-specific events is an important question. To address this question, in this paper, we have developed a testing approach to investigate whether the speculative bubbles lie in the common or in the idiosyncratic components in large-dimensional financial panel data sets. We first show that when the existing PANIC tests are applied to the explosive alternative hypothesis as opposed to the stationary alternative hypothesis, both the common and the idiosyncratic tests exhibit serious size distortions. More importantly, the idiosyncratic tests suffer from the nonmonotonic power

[13] Lustig et al. (2011) and Greenaway-McGrevy et al. (2015) use two factor model while Engel et al. (2015) consider three factors.

problem. By using the moderate deviations framework of Phillips and Magdalinos (2007), we find that the source of nonmonotonic power is an identification failure between the common factors and the explosive idiosyncratic components. This paper attempts to provide a cross-sectional method to disentangle the common and the idiosyncratic components in a relatively short explosive window. The method is justified by a sequential asymptotic framework and it is robust to the nonmonotonic power problem. In an empirical example using the 24 bilateral exchange rates, we discover that explosive behaviors prevailed in the exchange-rate system during the financial crisis period is not solely attributed to the value of USD but other currencies as well.

Acknowledgement: Yamamoto acknowledges financial support for this work from the MEXT Grants-in-Aid for Scientific Research No. 13K03593.

Appendix: Technical derivations

Throughout the appendix, we use the notation $\theta = \min\left\{\sqrt{N}, \sqrt{T}\right\}$ and denote k_T by k for simplicity.

Lemma A1. Under Assumptions 1, 3(a), 4, and 5, the following hold.
(a) $T^{-1/2}F_{[Tr]} \Rightarrow \sigma W_c(r)$,
(b) $T^{-3/2}\sum_{t=1}^{T} F_t \Rightarrow \sigma \int W_c(r)dr$,
(c) $T^{-1}\sum_{t=1}^{T} F_{t-1}e_t \Rightarrow \sigma^2 \int W_c(r)dW(r)$,
(d) $T^{-2}\sum_{t=1}^{T} F_t^2 \Rightarrow \sigma^2 \int W_c^2(r)dr$,
(e) $T^{-1/2}U_{i,[Tr]} \Rightarrow \sigma_i W_{c,i}(r)$,
(f) $T^{-3/2}\sum_{t=1}^{T} U_{it} \Rightarrow \sigma_i \int W_{c,i}(r)dr$,
(g) $T^{-1}\sum_{t=1}^{T} U_{i,t-1}z_{it} \Rightarrow \sigma_i^2 \int W_{c,i}(r)dW(r)$,
(h) $T^{-2}\sum_{t=1}^{T} U_{it}^2 \Rightarrow \sigma_i^2 \int W_{c,i}^2(r)dr$,
where $W_c(r)$ and $W_{c,i}(r)$ are independent Ornstein and Uhlenbeck processes defined on $r \in [0,1]$.

Proof of Lemma A1. See Phillips (1987).

Lemma A2. Under Assumptions 1, 3, 4, and 5, the following hold.
(a) $T^{-1}\sum_{t=1}^{T} f_t^2 \xrightarrow{p} \Sigma_f$, a positive constant,
(b) $E(u_{it}) = 0$ and $E\left|u_{it}\right|^8 = O(1)$,
(c) $|\gamma_{ss}^*| = O(1)$ for all s and $T^{-1}\sum_{s=1}^{T}\sum_{t=1}^{T}|\gamma_{st}^*| = O(1)$, where $\gamma_{st}^* = N^{-1}\sum_{i=1}^{N} E(u_{is}u_{it})$,
(d) $\sum_{i=1}^{N}\left|\phi_{ij}^*\right| = O(1)$ for all j and $N^{-1}\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\phi_{ij}^*\right| = O(1)$, where $\phi_{ij}^* = E(u_{it}u_{jt})$,
(e) $\zeta_{st}^* = O(1)$, where $\zeta_{st}^* = E\left|N^{-1/2}\sum_{i=1}^{N}[u_{is}u_{it} - E(u_{is}u_{it})]\right|^4$.

Proof of Lemma A2. (a) We start with

$$f_t = \frac{c}{T}F_{t-1} + e_t.$$

Squaring both sides, summing over t, and multiplying both sides by T^{-1} would yield

$$\begin{aligned} \frac{1}{T}\sum_{t=1}^{T} f_t^2 &= \frac{c^2}{T^3}\sum_{t=1}^{T} F_{t-1}^2 + 2\frac{c}{T^2}\sum_{t=1}^{T} F_{t-1}e_t + \frac{1}{T}\sum_{t=1}^{T} e_t^2, \\ &= I + II + T^{-1}\sum_{t=1}^{T} e_t^2 \xrightarrow{p} \sigma^2, \end{aligned}$$

because $I = O_p(T^{-1})$ by using Lemma A1 (d) and $II = o_p(T^{-1})$ by using Lemma A1 (c). The convergence of the third term is implied by the weak law of large numbers by Assumption 1. Hence, the result follows.

(b) It is straightforward that

$$E(u_{it}) = E(U_{it}) - E(U_{i,t-1}) = 0,$$

by Assumptions 3 (a). Next,

$$\begin{aligned}
E\left|u_{it}\right|^{8} &= E\left|\frac{c}{T}U_{i,t-1} + z_{it}\right|^{8}, \\
&\leq 2^{8} \times \max\left\{\frac{c_i^8}{T^8}E\left|U_{i,t-1}\right|^{8}, E\left|z_{it}\right|^{8}\right\},
\end{aligned}$$

but

$$E\left|U_{i,t-1}\right|^{8} \leq T^{8}\rho_i^{8T}E\left|z_{it}\right|^{8},$$

so that

$$\frac{c_i^8}{T^8}E\left|U_{i,t-1}\right|^{8} \leq c_i^8\rho_i^{8T}E\left|z_{it}\right|^{8},$$

where $\rho_i^{8T} = (1+\frac{c_i}{T})^{8T} \to \exp(8c_i)$ and $E\left|z_{it}\right|^{8} \leq M$ by Assumption 3 (a). Hence, the result follows.

(c) Without loss of generality, let $s \geq t$. Consider

$$\begin{aligned}
E(u_{is}u_{it}) &= E\left[\left(\frac{c_i}{T}U_{i,s-1} + z_{is}\right)\left(\frac{c_i}{T}U_{i,t-1} + z_{it}\right)\right], \\
&= \frac{c_i^2}{T^2}E(U_{i,s-1}U_{i,t-1}) + \frac{c_i}{T}E\left(U_{i,s-1}z_{it}\right) + \frac{c_i}{T}E\left(U_{i,t-1}z_{is}\right) + E(z_{is}z_{it}), \\
&= I + II + III + IV.
\end{aligned}$$

However,

$$I \leq \frac{c_i^2}{T}E(T^{-1}U_{i,s-1}^2) = O(T^{-1}),$$

by using Lemma A1 (e). For II,

$$\begin{aligned}
II &= \frac{c_i}{T}E\left(U_{i,s-1}z_{it}\right) = \frac{c_i}{T}E[(U_{i,s-1} - U_{it})z_{it} + u_{it}z_{it} + U_{i,t-1}z_{it}], \\
&= \frac{c_i}{T}E[(U_{i,s-1} - U_{it})z_{it}] + \frac{c_i}{T}E(u_{it}z_{it}) + \frac{c_i}{T}E(U_{i,t-1}z_{it}), \\
&= IIa + IIb + IIc.
\end{aligned}$$

However, since $U_{i,s-1} = z_{i,s-1} + \rho_i z_{i,s-2} + \cdots + \rho_i^{s-t-1}U_{it}$,

$$\begin{aligned}
IIa &= \frac{c_i}{T}E[\left\{z_{i,s-1} + \rho_i z_{i,s-2} + \rho_i^2 z_{i,s-3} + \cdots + (\rho_i^{s-t-1} - 1)U_{it}\right\}z_{it}], \\
&= \frac{c_i}{T}E[\left\{z_{i,s-1} + \rho_i z_{i,s-2} + \rho_i^2 z_{i,s-3} + \cdots + (\rho_i^{s-t-1} - 1)z_{it} + \rho_i(\rho_i^{s-t-1} - 1)U_{i,t-1}\right\}z_{it}], \\
&= \frac{c_i}{T}(\rho_i^{s-t-1} - 1)E(z_{it}^2) = O(T^{-1}),
\end{aligned}$$

by using Assumption 3 (a),

$$\begin{aligned} IIb &= \frac{c_i}{T}E[(z_{it}+\frac{c_i}{T}z_{i,t-1}+\frac{c_i}{T}\rho_i z_{i,t-2}+\cdots+\frac{c_i}{T}\rho_i^t U_{i0})z_{it}], \\ &= \frac{c_i}{T}E(z_{it}^2)=O(T^{-1}), \end{aligned}$$

by using Assumption 3 (a), and

$$IIc=\frac{c_i}{T^{1/2}}\underbrace{E(T^{-1/2}U_{i,t-1})}_{=O(1)\text{ by Lemma A1 (e)}}E(z_{it})=0,$$

so that $II=O(T^{-1})$. For III,

$$III=\frac{c_i}{T}E\left(U_{i,t-1}z_{is}\right)=\frac{c_i}{T}E\left(U_{i,t-1}\right)E(z_{is})=0,$$

since $U_{i,t-1}$ and z_{is} are independent as long as $s\geq t$. Therefore,

$$\begin{aligned} E(u_{is}u_{it}) &= O(T^{-1})+O(T^{-1})+O(T^{-1})+0+E(z_{is}z_{it}), \\ &= \begin{cases} \sigma_i^2+O(T^{-1}) & \text{if } s=t \\ O(T^{-1}) & \text{if } s\neq t \end{cases}. \end{aligned}$$

We now consider

$$\begin{aligned} \gamma_{st}^* &= E\left[N^{-1}\textstyle\sum_{i=1}^N u_{is}u_{it}\right], \\ &= \begin{cases} N^{-1}\sum_{i=1}^N \sigma_i^2+O(T^{-1}) & \text{if } s=t \\ O(T^{-1}) & \text{if } s\neq t \end{cases}. \end{aligned}$$

We also have

$$\textstyle\sum_{s=1}^T|\gamma_{st}^*|=N^{-1}\sum_{i=1}^N\sigma_i^2+O(1),$$

so that

$$T^{-1}\textstyle\sum_{s=1}^T\sum_{t=1}^T|\gamma_{st}^*|=N^{-1}\sum_{i=1}^N\sigma_i^2+O(1)=O(1).$$

(d) Consider

$$\begin{aligned} \phi_{ij}^* &= E(u_{it}u_{jt})=E\left[\left(\frac{c_i}{T}U_{i,t-1}+z_{it}\right)\left(\frac{c_j}{T}U_{j,t-1}+z_{jt}\right)\right], \\ &= \frac{c_ic_j}{T^2}E(U_{i,t-1}U_{j,t-1})+\frac{c_i}{T}E(U_{i,t-1}z_{jt})+\frac{c_j}{T}E(U_{j,t-1}z_{it})+E(z_{it}z_{jt}), \\ &= I+II+III+IV. \end{aligned}$$

For I,

$$I=\frac{c_ic_j}{T^2}E(U_{i,t-1}U_{j,t-1})=\frac{c_ic_j}{T^2}\phi_{ij}\left[\textstyle\sum_{l=0}^{t-1}(1+\frac{c_i}{T})^l(1+\frac{c_j}{T})^l\right],$$

and by assuming without loss of generality $c_i \geq c_j$ we obtain

$$\left[\sum_{l=0}^{t-1}(1+\frac{c_i}{T})^{2l}\right] \leq T(1+\frac{c_i}{T})^{2T} = O(T),$$

so that $I = \phi_{ij} \times O(T^{-1})$. For II,

$$II = \frac{c_i}{T^{1/2}} E(T^{-1/2}U_{i,t-1})E(z_{jt}) = 0,$$

by Assumption 3 (a), and similarly, $III = 0$. $IV = \phi_{ij}$ by definition. Therefore,

$$\phi_{ij}^* = \phi_{ij}[1 + O(T^{-1})],$$

so that

$$\sum_{i=1}^N \left|\phi_{ij}^*\right| = [1 + O(T^{-1})] \sum_{i=1}^N \left|\phi_{ij}\right| = O(1),$$

by Assumption 3 (c) and

$$N^{-1} \sum_{i=1}^N \sum_{j=1}^N \left|\phi_{ij}^*\right| = [1 + O(T^{-1})]N^{-1} \sum_{i=1}^N \sum_{j=1}^N \left|\phi_{ij}\right| = O(1),$$

by Assumption 3 (c) as well. Hence, the result follows.

(e) Since

$$u_{is}u_{it} = \frac{c^2}{T^2}U_{i,s-1}U_{i,t-1} + \frac{c}{T}U_{i,s-1}z_{it} + \frac{c}{T}U_{i,t-1}z_{is} + z_{is}z_{it},$$

$$\begin{aligned}
\zeta_{st}^* &= E\left|N^{-1/2}\sum_{i=1}^N \left[u_{is}u_{it} - E(u_{is}u_{it})\right]\right|^4, \\
&= E\left|\frac{c^2}{T^2N^{1/2}}\sum_{i=1}^N \left[U_{i,s-1}U_{i,t-1} - E(U_{i,s-1}U_{i,t-1})\right]\right. \\
&\quad + \frac{c}{TN^{1/2}}\sum_{i=1}^N \left[U_{i,s-1}z_{it} - E(U_{i,s-1}z_{it})\right] \\
&\quad + \frac{c}{TN^{1/2}}\sum_{i=1}^N \left[U_{i,t-1}z_{is} - E(U_{i,t-1}z_{is})\right] \\
&\quad \left. + \frac{1}{N^{1/2}}\sum_{i=1}^N \left[z_{is}z_{it} - E(z_{is}z_{it})\right]\right|^4, \\
&= E\left|\Phi_1 + \Phi_2 + \Phi_3 + \Phi_4\right|^4, \\
&\leq 4^4 \times \max\left\{E\left|\Phi_1\right|^4, E\left|\Phi_2\right|^4, E\left|\Phi_3\right|^4, \zeta_{st}\right\}.
\end{aligned}$$

Consider $E\left|\Phi_1\right|^4$. Since $U_{i,s-1}=\sum_{l=0}^{s-1}\rho_i^{s-1-l}z_{il}$ and $U_{i,t-1}=\sum_{m=0}^{t-1}\rho_i^{t-1-m}z_{im}$,

$$\begin{aligned}
E\left|\Phi_1\right|^4 &= \frac{c_i^8}{T^8}E\left|N^{-1/2}\sum_{i=1}^N\left[U_{i,s-1}U_{i,t-1}-E(U_{i,s-1}U_{i,t-1})\right]\right|^4, \\
&= \frac{c_i^8}{T^8}E\left|N^{-1/2}\sum_{i=1}^N\left[\sum_{l=0}^{s-1}\rho_i^{s-1-l}\sum_{m=0}^{t-1}\rho_i^{t-1-m}z_{il}z_{im}\right.\right. \\
&\quad \left.\left.-\sum_{l=1}^{s-1}\rho_i^{s-1-l}\sum_{m=0}^{t-1}\rho_i^{t-1-m}E(z_{il}z_{im})\right]\right|^4, \\
&= \frac{c_i^8}{T^8}E\left|N^{-1/2}\sum_{i=1}^N\sum_{l=0}^{s-1}\rho_i^{s-1-l}\sum_{m=0}^{t-1}\rho_i^{t-1-m}(z_{il}z_{im}-E(z_{il}z_{im}))\right|^4, \\
&\leq \frac{c_i^8}{T^8}E\left|\sum_{l=0}^{s-1}\rho_i^{s-1-l}\sum_{m=0}^{t-1}\rho_i^{t-1-m}\left|N^{-1/2}\sum_{i=1}^N(z_{il}z_{im}-E(z_{il}z_{im}))\right|\right|^4, \\
&\leq \frac{c_i^8}{T^8}T^8\rho_i^{8T}E\left|N^{-1/2}\sum_{i=1}^N(z_{il}z_{im}-E(z_{il}z_{im}))\right|^4, \\
&= c_i^8\rho_i^{8T}M=O(1),
\end{aligned}$$

by Assumption 3(d). Next,

$$\begin{aligned}
E\left|\Phi_2\right|^4 &= \frac{c_i^4}{T^4}E\left|N^{-1/2}\sum_{i=1}^N\left[U_{i,s-1}z_{it}-E(U_{i,s-1}z_{it})\right]\right|^4, \\
&= \frac{c_i^4}{T^4}E\left|N^{-1/2}\sum_{i=1}^N\left[\sum_{l=0}^{s-1}\rho_i^{s-1-l}z_{il}z_{it}-\sum_{l=0}^{s-1}\rho_i^{s-1-l}E(z_{il}z_{it})\right]\right|^4, \\
&= \frac{c_i^4}{T^4}E\left|\sum_{l=0}^{s-1}\rho_i^{s-1-l}N^{-1/2}\sum_{i=1}^N\left[z_{il}z_{it}-E(z_{il}z_{it})\right]\right|^4, \\
&\leq \frac{c_i^4}{T^4}T^4\rho_i^{4T}E\left|N^{-1/2}\sum_{i=1}^N\left[z_{il}z_{it}-E(z_{il}z_{it})\right]\right|^4, \\
&= c_i^4\rho_i^{4T}M=O(1),
\end{aligned}$$

and $E\left|\Phi_3\right|^4=O(1)$ is similarly shown. Therefore,

$$\zeta_{st}^*\leq 4^4\times\max\left\{O(1),\zeta_{st}\right\}=O(1),$$

by using Assumption 3 (d). Hence, the result follows.

Lemma A3. Under Assumptions 1–5, the following hold.

(a) $T^{-1/2}\sum_{t=1}^T(\hat{f}_t-Hf_t)=O_p(\theta^{-1})$,
(b) $T^{-1}\sum_{t=1}^T(\hat{f}_t-Hf_t)^2=O_p(\theta^{-2})$,
(c) $T^{-1}\sum_{t=1}^T(\hat{f}_t-Hf_t)u_{it}=O_p(\theta^{-2})$,
(d) $T^{-1}\sum_{t=1}^T(\hat{f}_t-Hf_t)f_t=O_p(\theta^{-2})$,
(e) $T^{-1}\sum_{t=1}^T(\hat{f}_t-Hf_t)\hat{f}_t=O_p(\theta^{-2})$,
(f) $\hat{\lambda}_i-H^{-1}\lambda_i=O_p\left(\frac{1}{\min\left\{N,T^{1/2}\right\}}\right)$.

Proof of Lemma A3. Part (a) is a direct consequence from Theorem 1 of Bai (2003). For part (b), the proof is straightforward by following Theorem 1 of Bai and Ng (2002) and replacing their assumptions with our Lemma A2. For parts (c), (d), and (e), the proof is obtained by following Lemmas B1, B2, and B3 of Bai (2003), respectively, by replacing their assumptions with our Lemma A2. For part (f), we have

$$
\begin{aligned}
\hat{\lambda}_i - \lambda_i H^{-1} &= T^{-1} H \textstyle\sum_{t=1}^T f_t u_{it} \\
&\quad + T^{-1} \textstyle\sum_{t=1}^T (\hat{f}_t - H f_t) \hat{f}_t \lambda_i + T^{-1} \sum_{t=1}^T (\hat{f}_t - H f_t) u_{it}, \\
&= T^{-1} H \textstyle\sum_{t=1}^T f_t u_{it} + O_p(\theta^{-2}),
\end{aligned}
$$

by using Lemma A3 (e) and (c). Now

$$
\begin{aligned}
T^{-1} \textstyle\sum_{t=1}^T f_t u_{it} &= T^{-1} \textstyle\sum_{t=1}^T \left(\frac{c}{T} F_{t-1} + e_t \right) \left(\frac{c_i}{T} U_{i,t-1} + z_{it} \right) \\
&= c c_i T^{-3} \textstyle\sum_{t=1}^T F_{t-1} U_{i,t-1} + c T^{-2} \sum_{t=1}^T F_{t-1} z_{it} \\
&\quad + c_i T^{-2} \textstyle\sum_{t=1}^T U_{i,t-1} e_t + T^{-1} \sum_{t=1}^T e_t z_{it}, \\
&= I + II + III + IV.
\end{aligned}
$$

However, if we use Cauchy–Schwarz inequality, Lemma A1 (d) and (h), Assumptions 1 and 3 (a), we obtain

$$
\begin{aligned}
I &\leq c c_i T^{-1} \left(T^{-2} \textstyle\sum_{t=1}^T F_{t-1}^2 \right)^{1/2} \left(T^{-2} \sum_{t=1}^T U_{i,t-1}^2 \right)^{1/2} = O_p(T^{-1}), \\
II &\leq c T^{-1/2} \left(T^{-2} \textstyle\sum_{t=1}^T F_{t-1}^2 \right)^{1/2} \left(T^{-1} \sum_{t=1}^T z_{i,t-1}^2 \right)^{1/2} = O_p(T^{-1/2}), \\
III &\leq c_i T^{-1/2} \left(T^{-2} \textstyle\sum_{t=1}^T U_{i,t-1}^2 \right)^{1/2} \left(T^{-1} \sum_{t=1}^T e_{t-1}^2 \right)^{1/2} = O_p(T^{-1/2}).
\end{aligned}
$$

For IV, Assumptions 1, 3 (a), and 4 imply that $\{e_t z_{it}\}_{t=2}^T$ is a white noise sequence so that $IV = O_p(T^{-1/2})$. Therefore,

$$
\hat{\lambda}_i - H^{-1} \lambda_i = O_p(T^{-1/2}) + O_p(\theta^{-2}) = O_p \left(\frac{1}{\min \{N, T^{1/2}\}} \right). \tag{A.1}
$$

Lemma A4. Under Assumptions 1–5, the following hold.

(a) $T^{-1} \sum_{t=1}^T \hat{f}_t^2 = T^{-1} H^2 \sum_{t=1}^T f_t^2 + O_p(\theta^{-2})$,
(b) $T^{-2} \sum_{t=1}^T \hat{F}_{t-1}^2 = T^{-2} H^2 \sum_{t=1}^T F_{t-1}^2 + O_p(\theta^{-1})$,
(c) $T^{-1} \sum_{t=1}^T \hat{F}_{t-1} \hat{f}_t = T^{-1} H^2 \sum_{t=1}^T F_{t-1} f_t + O_p(\theta^{-1})$,
(d) $T^{-1} \sum_{t=1}^T \hat{u}_{it}^2 = T^{-1} \sum_{t=1}^T u_{it}^2 + O_p(\theta^{-2})$,
(e) $T^{-2} \sum_{t=1}^T \hat{U}_{i,t-1}^2 = T^{-2} \sum_{t=1}^T U_{i,t-1}^2 + O_p(\theta^{-1})$,
(f) $T^{-1} \sum_{t=1}^T \hat{U}_{i,t-1} \hat{u}_{it} = T^{-1} \sum_{t=1}^T U_{i,t-1} u_{it} + O_p(\theta^{-1})$.

Proof of Lemma A4. Note that $\hat{F}_0 = 0$ and $\hat{U}_{i0} = 0$ for all i by definition. (a) We start with the identity

$$\begin{aligned}
T^{-1}\sum_{t=1}^{T}\hat{f}_t^2 &= T^{-1}\sum_{t=1}^{T}\left[Hf_t+(\hat{f}_t-Hf_t)\right]^2, \\
&= T^{-1}H^2\sum_{t=1}^{T}f_t^2+T^{-1}\sum_{t=1}^{T}(\hat{f}_t-Hf_t)^2, \\
&\quad +2T^{-1}H\sum_{t=1}^{T}f_t(\hat{f}_t-Hf_t) \\
&= T^{-1}H^2\sum_{t=1}^{T}f_t^2+I+II,
\end{aligned}$$

However, $I=O_p(\theta^{-2})$ by using Lemma A3 (b) and $II=O_p(\theta^{-2})$ by using Lemma A3 (d). Hence, the result follows.

(b) This part closely follows Bai and Ng's (2004) Lemma B2. Since

$$\hat{F}_{t-1}=HF_{t-1}+\sum_{s=1}^{t-1}(\hat{f}_s-Hf_s), \tag{A.2}$$

squaring both sides, summing over t, and multiplying by T^{-2} would yield

$$\begin{aligned}
T^{-2}\sum_{t=1}^{T}\hat{F}_{t-1}^2 &= T^{-2}H^2\sum_{t=1}^{T}F_{t-1}^2+T^{-1}\sum_{t=1}^{T}\left[T^{-1/2}\sum_{s=1}^{t-1}(\hat{f}_s-Hf_s)\right]^2 \\
&\quad +2T^{-1}H\sum_{t=1}^{T}F_{t-1}\left[T^{-1/2}\sum_{s=1}^{t-1}(\hat{f}_s-Hf_s)\right], \\
&= T^{-2}H^2\sum_{t=1}^{T}F_{t-1}^2+I+II.
\end{aligned}$$

However, $I=O_p(\theta^{-1})$ by using Lemma A3 (a). For term II, we use Cauchy–Schwarz inequality to get

$$\begin{aligned}
II &\leq 2\left(T^{-2}\sum_{t=1}^{T}F_{t-1}^2\right)^{1/2}\left[T^{-1}\sum_{t=1}^{T}\left(T^{-1/2}\sum_{s=1}^{t-1}(\hat{f}_s-Hf_s)\right)^2\right]^{1/2}, \\
&= O_p(1)\times O_p(\theta^{-1}),
\end{aligned}$$

by using Lemma A1 (d) for the first term and Lemma A3 (a) for the second term. Hence, the result follows.

(c) Since $F_t^2=(F_{t-1}+f_t)^2=F_{t-1}^2+f_t^2+2F_{t-1}f_t$ by construction, we obtain

$$F_{t-1}f_t=\frac{1}{2}(F_t^2-F_{t-1}^2-f_t^2).$$

Summing over t and multiplying by T^{-1} would yield

$$T^{-1}\sum_{t=1}^{T}F_{t-1}f_t=\frac{1}{2}(T^{-1}F_T^2-T^{-1}F_0^2-T^{-1}\sum_{t=1}^{T}f_t^2). \tag{A.3}$$

We also have by construction

$$\hat{F}_{t-1}\hat{f}_t=\frac{1}{2}(\hat{F}_t^2-\hat{F}_{t-1}^2-\hat{f}_t^2),$$

so that

$$T^{-1}\sum_{t=1}^{T}\hat{F}_{t-1}\hat{f}_t=\frac{1}{2}(T^{-1}\hat{F}_T^2-T^{-1}\hat{F}_0^2-T^{-1}\sum_{t=1}^{T}\hat{f}_t^2). \tag{A.4}$$

Subtracting (A.3) multiplied by H^2 from (A.4) would yield

$$\begin{aligned}
T^{-1}\textstyle\sum_{t=1}^T \hat{F}_{t-1}\hat{f}_t &= T^{-1}H^2\textstyle\sum_{t=1}^T F_{t-1}f_t + \frac{1}{2T}(\hat{F}_T^2 - H^2F_T^2) - \frac{1}{2T}(\hat{F}_0^2 - H^2F_0^2) \\
&\quad - \left(T^{-1}\textstyle\sum_{t=1}^T \hat{f}_t^2 - T^{-1}H^2\sum_{t=1}^T f_t^2\right), \\
&= T^{-1}H^2\textstyle\sum_{t=1}^T F_{t-1}f_t + I + II + III.
\end{aligned}$$

For I, updating (A.2) to the period T, squaring both sides, and multiplying by T^{-1} would yield

$$\begin{aligned}
T^{-1}\hat{F}_T^2 &= T^{-1}H^2F_T^2 + \underbrace{\left[T^{-1/2}\textstyle\sum_{s=1}^T(\hat{f}_s - Hf_s)\right]^2}_{=O_p(\theta^{-2}) \text{ by Lemma A3(a)}} \\
&\quad +2 \underbrace{T^{-1/2}F_T}_{=O_p(1) \text{ by Lemma A1(a)}} \underbrace{\left[T^{-1/2}\textstyle\sum_{s=1}^T(\hat{f}_s - Hf_s)\right]}_{=O_p(\theta^{-1}) \text{ by Lemma A3(a)}},
\end{aligned}$$

so that $I = O_p(\theta^{-1})$. For II,

$$\hat{F}_0^2 - H^2F_0^2 = -H^2(\alpha^2F_0^2 + e_1^2 + 2\alpha F_0 e_1),$$

is bounded as $T \to \infty$ so that $II = O_p(T^{-1})$. Term III is $O_p(\theta^{-2})$ by using Lemma A4 (a). Hence, the result follows.

(d) Since $\hat{u}_{it} = x_{it} - \hat{\lambda}_i\hat{f}_t$ and $x_{it} = u_{it} + \lambda_i H^{-1}Hf_t$,

$$\begin{aligned}
\hat{u}_{it} &= u_{it} + \lambda_i H^{-1}Hf_t - \hat{\lambda}_i\hat{f}_t, \\
&= u_{it} - \lambda_i H^{-1}(\hat{f}_t - Hf_t) - (\hat{\lambda}_i - \lambda_i H^{-1})\hat{f}_t. \qquad \text{(A.5)}
\end{aligned}$$

Squaring both sides, summing over t, and multiplying by T^{-1} would yield

$$\begin{aligned}
T^{-1}\textstyle\sum_{t=1}^T \hat{u}_{it}^2 &= T^{-1}\textstyle\sum_{t=1}^T u_{it}^2 + \lambda_i^2H^{-2}T^{-1}\sum_{t=1}^T(\hat{f}_t - Hf_t)^2 + (\hat{\lambda}_i - \lambda_iH^{-1})^2T^{-1}\sum_{t=1}^T \hat{f}_t^2, \\
&\quad -2\lambda_iH^{-1}T^{-1}\textstyle\sum_{t=1}^T(\hat{f}_t - Hf_t)u_{it} - 2(\hat{\lambda}_i - \lambda_iH^{-1})T^{-1}\sum_{t=1}^T \hat{f}_tu_{it}, \\
&\quad +2\lambda_i(\hat{\lambda}_i - \lambda_iH^{-1})T^{-1}\textstyle\sum_{t=1}^T(\hat{f}_t - Hf_t)\hat{f}_t, \\
&= T^{-1}\textstyle\sum_{t=1}^T u_{it}^2 + I + II + III + IV + V.
\end{aligned}$$

However, $I = O_p(\theta^{-2})$ by Lemma A3 (b), $II = O_p(\frac{1}{\min\{T,N^2\}})$ by Lemma A3 (f) and $T^{-1}\sum_{t=2}^T \hat{f}_t = 1$, and $III = O_p(\theta^{-2})$ by Lemma A3 (c). We also have

$$\begin{aligned}
IV &= -2(\hat{\lambda}_i - \lambda_iH^{-1})T^{-1}\textstyle\sum_{t=1}^T \hat{f}_tu_{it}, \\
&= -2(\hat{\lambda}_i - \lambda_iH^{-1})T^{-1}\textstyle\sum_{t=1}^T(\hat{f}_t - Hf_t)u_{it} - 2(\hat{\lambda}_i - \lambda_iH^{-1})T^{-1}H\sum_{t=1}^T f_tu_{it}, \\
&= O_p\left(\frac{1}{\min\{T,N^2\}}\right) \times O_p(\theta^{-2}) + O_p\left(\frac{1}{\min\{T,N^2\}}\right) \times O_p(T^{-1/2}),
\end{aligned}$$

by using Lemma A3 (f) and Lemma A3 (c) for the first term and by using Lemma A3 (f) for the second term, $V = O_p\left(\frac{1}{\min\{T,N^2\}}\right) \times O_p(\theta^{-2})$ by using Lemma A3 (f) and Lemma A3 (e). Hence, the result follows.

(e) We have

$$\begin{aligned}
\hat{U}_{it} &= \textstyle\sum_{s=1}^{t} \hat{u}_{is}, \\
&= \textstyle\sum_{s=1}^{t} u_{is} - \lambda_i H^{-1} \sum_{s=1}^{t}(\hat{f}_s - Hf_s) - (\hat{\lambda}_i - \lambda_i H^{-1}) \sum_{s=1}^{t} \hat{f}_s, \\
&= U_{it} - U_{i0} - \lambda_i H^{-1} \textstyle\sum_{s=1}^{t}(\hat{f}_s - Hf_s) - (\hat{\lambda}_i - \lambda_i H^{-1}) \sum_{s=1}^{t} \hat{f}_s,
\end{aligned}$$

by using (A.5). Multiplying both sides by $T^{-1/2}$ would yield

$$\begin{aligned}
T^{-1/2}\hat{U}_{it} &= T^{-1/2}U_{it} - T^{-1/2}U_{i0} - \lambda_i H^{-1}\left[T^{-1/2} \textstyle\sum_{s=1}^{t}(\hat{f}_s - Hf_s)\right] - (\hat{\lambda}_i - H^{-1}\lambda_i)T^{-1/2} \sum_{s=1}^{t} \hat{f}_s, \\
&= T^{-1/2}U_{it} + I + II + III.
\end{aligned}$$

but $I = O_p(T^{-1/2})$ by Assumption 3(a), $II = O_p(\theta^{-1})$ by using Lemma A3 (a), $III = O_p\left(\frac{1}{\min\left\{N,T^{1/2}\right\}}\right)$ by using Lemma A3 (f) and

$$\begin{aligned}
T^{-1/2} \textstyle\sum_{s=1}^{t} \hat{f}_s &= T^{-1/2}\hat{F}_t = T^{-1/2}F_t + T^{-1/2} \textstyle\sum_{s=1}^{t}(\hat{f}_s - Hf_s), \\
&= O_p(1) + O_p(\theta^{-1}),
\end{aligned}$$

by using Lemma A1 (b) and Lemma A3 (a). This results in $T^{-1/2}\hat{U}_{it} = T^{-1/2}U_{it} + O_p(\theta^{-1})$ so that squaring both sides would yield

$$\begin{aligned}
T^{-1}\hat{U}_{it}^2 &= T^{-1}U_{it}^2 + O_p(\theta^{-2}) + O_p(\theta^{-1}) \times T^{-1/2}U_{it}, \\
&= T^{-1}U_{it}^2 + O_p(\theta^{-1}),
\end{aligned} \tag{A.6}$$

by using Lemma A1 (e). Furthermore, summing over t would yield

$$T^{-1} \textstyle\sum_{t=1}^{T} \hat{U}_{it}^2 = T^{-1} \sum_{t=1}^{T} U_{it}^2 + O_p(\theta^{-1})T^{-1/2} \sum_{t=1}^{T} U_{it}.$$

Multiplying both sides by T^{-1} would yield

$$\begin{aligned}
T^{-2} \textstyle\sum_{t=1}^{T} \hat{U}_{it}^2 &= T^{-2} \textstyle\sum_{t=1}^{T} U_{it}^2 + O_p(\theta^{-1}) \underbrace{T^{-3/2} \textstyle\sum_{t=1}^{T} U_{it}}_{=O_p(1) \text{ by Lemma A1(g)}}, \\
&= T^{-2} \textstyle\sum_{t=1}^{T} U_{it}^2 + O_p(\theta^{-1}).
\end{aligned}$$

Hence, the result follows.

(f) We use a similar identity as (A.4) for $\hat{U}_{it}$

$$T^{-1} \textstyle\sum_{t=1}^{T} \hat{U}_{i,t-1}\hat{u}_{it} = \dfrac{\hat{U}_{iT}^2}{2T} - \dfrac{\hat{U}_{i0}^2}{2T} - \dfrac{1}{2T} \sum_{t=1}^{T} \hat{u}_{it}^2, \tag{A.7}$$

and a similar identity as (A.3) for U_{it}

$$T^{-1}\sum_{t=1}^{T}U_{i,t-1}u_{it}=\frac{U_{iT}^2}{2T}-\frac{U_{i0}^2}{2T}-\frac{1}{2T}\sum_{t=1}^{T}u_{it}^2. \tag{A.8}$$

Subtracting (A.8) from (A.7) would yield

$$\begin{aligned} & T^{-1}\sum_{t=1}^{T}\hat{U}_{i,t-1}\hat{u}_{it}-T^{-1}\sum_{t=1}^{T}U_{i,t-1}u_{it}, \\ = & \frac{1}{2T}(\hat{U}_{iT}^2-U_{iT}^2)-\frac{1}{2T}(\hat{U}_{i0}^2-U_{i0}^2)-\frac{1}{2T}\left(\sum_{t=1}^{T}\hat{u}_{it}^2-\sum_{t=1}^{T}u_{it}^2\right), \\ = & I+II+III. \end{aligned}$$

However, I and II are $O_p(\theta^{-1})$ by (A.6) and III is $O_p(\theta^{-2})$ by using Lemma A4 (d). Hence, the result follows.

Proof of Theorem 1. The common test is

$$t_{\hat{F}}=\frac{T\hat{\delta}}{\hat{\sigma}\left(T^{-2}\sum_{t=2}^{T}\hat{F}_{t-1}^2\right)^{-1/2}}. \tag{A.9}$$

Under Assumptions 1–5, we can use Lemma A4 (b) and (c) so that the numerator becomes

$$\begin{aligned} T\hat{\delta} &= \frac{T^{-1}\sum_{t=1}^{T}\hat{F}_{t-1}\hat{f}_t}{T^{-2}\sum_{t=1}^{T}\hat{F}_{t-1}^2}, \\ &= \frac{T^{-1}H^2\sum_{t=1}^{T}F_{t-1}f_t+O_p(\theta^{-1})}{T^{-2}H^2\sum_{t=1}^{T}F_{t-1}^2+O_p(\theta^{-1})}. \end{aligned} \tag{A.10}$$

The denominator has two components. One is

$$T^{-2}\sum_{t=1}^{T}\hat{F}_{t-1}^2=T^{-2}H^2\sum_{t=1}^{T}F_{t-1}^2+O_p(\theta^{-1}), \tag{A.11}$$

by Lemma A4 (b) and the other is

$$\begin{aligned} \hat{\sigma}^2 &= T^{-1}\sum_{t=1}^{T}\left(\hat{f}_t-\hat{\delta}\hat{F}_{t-1}\right)^2, \\ &= T^{-1}\sum_{t=1}^{T}\hat{f}_t^2-2T^{-1}\hat{\delta}\sum_{t=1}^{T}\hat{f}_t\hat{F}_{t-1}+(T\hat{\delta})^2T^{-3}\sum_{t=1}^{T}\hat{F}_{t-1}^2, \\ &= T^{-1}H^2\sum_{t=1}^{T}f_t^2-2(T\hat{\delta})T^{-2}H^2\sum_{t=1}^{T}f_tF_{t-1}+(T\hat{\delta})^2T^{-3}H^2\sum_{t=1}^{T}F_{t-1}^2+O_p(\theta^{-1}), \\ &= T^{-1}H^2\sum_{t=1}^{T}f_t^2-O_p(T^{-1})+O_p(T^{-1})+O_p(\theta^{-1}), \end{aligned}$$

by Lemma A4 (a), (b), and (c). Therefore, the variance estimate satisfies $\hat{\sigma}^2\xrightarrow{p}Q^{-2}\sigma^2$ for any fixed c.

(i-a) If $c=0$, then (A.10)$\Rightarrow\int_0^1W(r)dW(r)/\int_0^1W(r)^2dr$ and (A.11)$\Rightarrow\sigma^2Q^{-2}\int_0^1W(r)^2dr$ so that

$$t_{\hat{F}}\Rightarrow\frac{\int_0^1W(r)dW(r)}{\left[\int_0^1W(r)^2dr\right]^{1/2}},$$

as $N, T \to \infty$.

(i-b) If $c > 0$, then Lemma A1 (c) and (d) imply that (A.10)$\Rightarrow c + \int_0^1 W_c(r)dW(r) / \int_0^1 W_c(r)^2 dr$. The result follows.

For parts (ii-a) and (ii-b), we follow the same steps as above by replacing $\hat{f}_t$ and $\hat{F}_{t-1}$ with $\hat{u}_{it}$ and $\hat{U}_{i,t-1}$ and using Lemma A4 (d)–(f) to show the results. Hence, the proof is suppressed.

Lemma A5. Under Assumptions 1 and 6, the following hold.
(a) $k^{-1/2} \sum_{t=1}^{T} \alpha^{-t} e_t \Rightarrow N(0, \sigma^2/2c)$,
(b) $\sum_{t=1}^{T} F_t = O_p(kT^{1/2}) + O_p(\alpha^T k^{3/2})$,
(c) $\alpha^{-2T+1} k^{-1} \sum_{t=1}^{T} F_{t-1} e_t = o_p(1)$,
(d) $\alpha^{-2T} k^{-2} \sum_{t=1}^{T} F_t^2 \Rightarrow \frac{1}{2c} \Theta^2$, where $\Theta = N(0, \sigma^2/2c)$,
(e) $T^{-1} \sum_{t=1}^{T} f_t^2 = O_p(\alpha^{2T} T^{-1}) + O_p(1)$.

Under Assumptions 3 (a) and 6, the following hold for all i:

(f) $k^{-1/2} \sum_{t=1}^{T} \rho_i^{-t} z_{it} \Rightarrow N(0, \sigma_i^2/2c_i)$,
(g) $\sum_{t=1}^{T} U_{it} = O_p(kT^{1/2}) + O_p(\rho_i^T k^{3/2})$,

(h) ${\rho_i}^{-2T+1} k^{-1} \sum_{t=1}^{T} U_{i,t-1} z_{it} = o_p(1)$,
(i) ${\rho_i}^{-2T} k^{-2} \sum_{t=1}^{T} U_{it}^2 \Rightarrow \frac{1}{2c_i} \Theta_i^2$, where $\Theta_i = N(0, \sigma_i^2/2c_i)$,
(j) $T^{-1} \sum_{t=1}^{T} u_{it}^2 = O_p(\rho_i^{2T} T^{-1}) + O_p(1)$.

Proof of Lemma A5. Here, we present the proof of only parts (a) to (e). Proof of parts (f) to (j) is shown in the same way but using U_{it} instead of F_t and replacing Assumption 1 with Assumption 3 (a). Thus, it is suppressed to conserve space.

(a) See Lemma 4.2 of Phillips and Magdalinos (2007).

(b) We start with the expression

$$\begin{aligned}
\sum_{t=1}^{T} F_t &= \sum_{t=0}^{T} \alpha^t e_0 + \sum_{t=0}^{T-1} \alpha^t e_1 + \sum_{t=0}^{T-2} \alpha^t e_2 + \cdots + e_T, \\
&= \frac{1}{1-\alpha} \left[(\alpha - \alpha^{T+1}) F_0 + (1-\alpha^T) e_1 + (1-\alpha^{T-1}) e_2 + \cdots + (1-\alpha) e_T \right], \\
&= \frac{k}{c} \left[\sum_{t=1}^{T} e_t - \sum_{t=1}^{T} \alpha^{T+1-t} e_t + (\alpha - \alpha^{T+1}) F_0 \right], \\
&= \frac{k}{c} \sum_{t=1}^{T} e_t - \frac{\alpha^{T+1} k}{c} \sum_{t=1}^{T} \alpha^{-t} e_t + \frac{k}{c} (\alpha - \alpha^{T+1}) F_0, \\
&= I + II + III.
\end{aligned}$$

However, $I = O_p(kT^{1/2})$ by Assumption 1, $II = O_p(\alpha^T k^{3/2})$ by using Lemma A5 (a), and $III = O_p(\alpha^T k)$ by Assumption 1. Hence, the result follows.

(c) We start with the expression for F_{t-1}

$$F_{t-1} = e_{t-1} + \alpha e_{t-2} + ... + \alpha^{t-1} e_1 + \alpha^t F_0 = \alpha^{t-1} \sum_{s=1}^{t-1} \alpha^{-s} e_s + \alpha^{t-1} F_0.$$

Multiplying both sides by $\alpha^{-2T+1} k^{-1} e_t$ and summing over t would yield

$$\begin{aligned}
\alpha^{-2T+1} k^{-1} \sum_{t=1}^T F_{t-1} e_t &= \alpha^{-2T+1} k^{-1} \sum_{t=1}^T \left(\sum_{s=1}^{t-1} \alpha^{t-s-1} e_s \right) e_t + \alpha^{-2T+1} k^{-1} F_0 \sum_{t=1}^T \alpha^{t-1} e_t, \\
&= I + II.
\end{aligned}$$

The expected value of this is zero because of Assumption 1. In order to show that this is $o_p(1)$, we show that the second moment of both terms diminishes as $T \to \infty$. For I, by using Assumption 1, we can simplify the second moment as follows.

$$\begin{aligned}
& E \left[\alpha^{-2T+1} k^{-1} \sum_{t=1}^T \left(\sum_{s=0}^{t-1} \alpha^{t-s-1} e_s \right) e_t \right]^2, \\
=\ & \alpha^{-4T+2} k^{-2} \sigma^4 \sum_{t=1}^T \sum_{s=0}^{t-1} \alpha^{2(t-s-1)}, \\
=\ & \alpha^{-4T} \frac{\alpha^4 \sigma^4}{k(\alpha^2 - 1)} \left(\frac{\alpha^{2T} - 1}{k(\alpha^2 - 1)} - k^{-1} T \right).
\end{aligned}$$

However, since $k(\alpha^2 - 1) \to 2c$, $\alpha^4 \to 1$, and $k^{-1}T = o(1)$, this is $O(\alpha^{-2T})$. This is $o(kT^{-1})$ by using Proposition A.1 (b) of Phillips and Magdalinos (2007). Further, by Assumption 6, it is $o(1)$. For II,

$$\begin{aligned}
& E \left[\alpha^{-2T+1} k^{-1} F_0 \sum_{t=1}^T \alpha^{t-1} e_t \right]^2, \\
=\ & \alpha^{-4T} k^{-1} \frac{F_0^2 \sigma^2}{k(\alpha^2 - 1)} (\alpha^{2T+2} - \alpha^2) = O(\alpha^{-2T} k^{-1}),
\end{aligned}$$

so that the second moment of both I and II diminishes. Therefore, the result follows.

(d) See the derivation of equation (9) of Phillips and Magdalinos (2007).

(e) We start with

$$f_t = \frac{c}{k} F_{t-1} + e_t.$$

Squaring both sides, summing over t, and multiplying by $(T-1)^{-1}$ would yield

$$\begin{aligned}
\frac{1}{T-1} \sum_{t=2}^T f_t^2 &= \frac{c^2}{(T-1)k^2} \sum_{t=2}^T F_{t-1}^2 + \frac{2c}{(T-1)k} \sum_{t=2}^T F_{t-1} e_t + \frac{1}{T-1} \sum_{t=2}^T e_t^2, \\
&= I + II + III.
\end{aligned}$$

However, $I = O_p(\alpha^{2T} T^{-1})$ by using Lemma A5 (d), $II = o_p(\alpha^T T^{-1})$ by using Lemma A5 (c), and $III = O_p(1)$ by Assumption 1. Hence, the result follows.

Proof of Theorem 2. We start with equation (A.1) of Bai and Ng (2004). Let $u_t = [u_{1t}\ u_{2t} \cdots u_{Nt}]$ be an $1 \times N$ vector of first differences of the idiosyncratic errors at time t.

$$\begin{aligned}
\hat{f}_t &= H f_t + V^{-1} N^{-1} T^{-1} \hat{f}' u \Lambda f_t + V^{-1} N^{-1} T^{-1} \hat{f}' f \Lambda' u_t' \\
&\quad + V^{-1} N^{-1} T^{-1} \hat{f}' u u_t',
\end{aligned}$$

or

$$V\hat{f}_t = Af_t + Bf_t + N^{-1}\sum_{i=1}^N a_i u_{it} + N^{-1}\sum_{i=1}^N b_i u_{it},$$

where $A = N^{-1}T^{-1}\hat{f}'f\Lambda'\Lambda$ by definition of H matrix. We also have $B = N^{-1}T^{-1}\hat{f}'u\Lambda$, $a_i = T^{-1}\hat{f}'f\lambda_i'$, and $b_i = T^{-1}\hat{f}'u_i$.

(i) If $c > 0$ and $c_i = 0$ for all i, and the stated condition is satisfied, then

$$\begin{aligned}
A &\leq \underbrace{\left|T^{-1}\textstyle\sum_{s=1}^T \hat{f}_s^2\right|^{1/2}}_{=1} \underbrace{\left|T^{-1}\textstyle\sum_{s=1}^T f_s^2\right|^{1/2}}_{=O_p(\alpha^T T^{-1/2}) \text{ by Lemma A5(e)}} \underbrace{\left|N^{-1}\textstyle\sum_{i=1}^N \lambda_i^2\right|}_{=O_p(1) \text{ by Assumption 2}} = O_p(\alpha^T T^{-1/2}),\\
B &\leq \underbrace{\left|T^{-1}\textstyle\sum_{s=1}^T \hat{f}_s^2\right|^{1/2}}_{=1} \underbrace{\left|N^{-1}\textstyle\sum_{i=1}^N T^{-1}\sum_{s=1}^T u_{is}^2\right|^{1/2}}_{=O_p(1)} \underbrace{\left|N^{-1}\textstyle\sum_{i=1}^N \lambda_i^2\right|}_{=O_p(1)} = O_p(1),\\
a_i &\leq \underbrace{\left|T^{-1}\textstyle\sum_{s=1}^T \hat{f}_s^2\right|^{1/2}}_{=1} \underbrace{\left|T^{-1}\textstyle\sum_{s=1}^T f_s^2\right|^{1/2}}_{=O_p(\alpha^T T^{-1/2})} |\lambda_i| = O_p(\alpha^T T^{-1/2}),\\
b_i &\leq \underbrace{\left|T^{-1}\textstyle\sum_{s=1}^T \hat{f}_s^2\right|^{1/2}}_{=1} \underbrace{\left|T^{-1}\textstyle\sum_{s=1}^T u_{is}^2\right|^{1/2}}_{=O_p(1)} = O_p(1).
\end{aligned}$$

(ii) If $c = 0$ and $c_i > 0$ for all i, then

$$\begin{aligned}
A &\leq \underbrace{\left|T^{-1}\textstyle\sum_{s=1}^T \hat{f}_s^2\right|^{1/2}}_{=1} \underbrace{\left|T^{-1}\textstyle\sum_{s=1}^T f_s^2\right|^{1/2}}_{=O_p(1)} \underbrace{\left|N^{-1}\textstyle\sum_{i=1}^N \lambda_i^2\right|}_{=O_p(1)} = O_p(1),\\
B &\leq \underbrace{\left|T^{-1}\textstyle\sum_{s=1}^T \hat{f}_s^2\right|^{1/2}}_{=1} \underbrace{\left|N^{-1}\textstyle\sum_{i=1}^N T^{-1}\sum_{s=1}^T u_{is}^2\right|^{1/2}}_{=O_p(\rho_i^T T^{-1/2}) \text{ by Lemma A5 (j)}} \underbrace{\left|N^{-1}\textstyle\sum_{i=1}^N \lambda_i^2\right|}_{=O_p(1)} = O_p(\rho_i^T T^{-1/2}),\\
a_i &\leq \underbrace{\left|T^{-1}\textstyle\sum_{s=1}^T \hat{f}_s^2\right|^{1/2}}_{=1} \underbrace{\left|T^{-1}\textstyle\sum_{s=1}^T f_s^2\right|^{1/2}}_{=O_p(1)} |\lambda_i| = O_p(1),\\
b_i &\leq \underbrace{\left|T^{-1}\textstyle\sum_{s=1}^T \hat{f}_s^2\right|^{1/2}}_{=1} \underbrace{\left|T^{-1}\textstyle\sum_{s=1}^T u_{is}^2\right|^{1/2}}_{=O_p(\rho_i^T T^{-1/2})} = O_p(\rho_i^T T^{-1/2}).
\end{aligned}$$

The result follows. Note that the largest eigenvalue V of $N^{-1}T^{-1}xx'$ satisfies $V^{1/2} = \left\|N^{-1/2}T^{-1/2}x\right\|$, where $\|\cdot\|$ denotes the Euclidean norm, so that

$$\begin{aligned}
V &= N^{-1}T^{-1}\|x\|^2,\\
&= N^{-1}T^{-1}\textstyle\sum_{i=1}^N \sum_{t=1}^T x_{it}^2,\\
&= \begin{cases} O_p(\alpha^T T^{-1/2}), & \text{for case (i)} \\ O_p(\rho_i^T T^{-1/2}), & \text{for case (ii)} \end{cases}.
\end{aligned} \tag{A.12}$$

Proof of Theorem 3. Without loss of generality, we confine the case of $r=1$. We first show the common test results in parts (i), (ii), and (iii). We have

$$\widetilde{F}_t = \frac{\sum_{i=1}^{N}\hat{\lambda}_i^{*}X_{it}}{\sum_{i=1}^{N}\hat{\lambda}_i^{*2}} = \frac{N^{-1}\sum_{i=1}^{N}\hat{\lambda}_i^{*}\lambda_i}{N^{-1}\sum_{i=1}^{N}\hat{\lambda}_i^{*2}}F_t + \frac{N^{-1}\sum_{i=1}^{N}\hat{\lambda}_i^{*}U_{it}}{N^{-1}\sum_{i=1}^{N}\hat{\lambda}_i^{*2}} = I + II.$$

For I, by using Lemma 1 (c) of Bai and Ng (2004), we know that $p\lim_{N,T\to\infty}\hat{\lambda}_i^{*} = Q^{-1}\lambda_i$ for any i, where $Q = p\lim_{N,T\to\infty} H$. Hence, $I \to QF_t$ for any $t = T+1, ..., T+h$ as $N, T \to \infty$. For II, the numerator converges in probability to σ_λ^2 and the numerator converges in probability to zero because $E(U_{it}) = 0$ and U_{it} is independent of λ_i. Therefore, after N, $T\to\infty$, we obtain $\widetilde{F}_t \xrightarrow{p} QF_t$ for any $t = T+1, ..., T+h$. The common test is

$$t_{\widetilde{F}}^{*} = \frac{\hat{\delta}^{*}}{\hat{\sigma}Q^{-1}(\sum_{t=T+1}^{T+h}F_{t-1}^2)^{-1/2}}, \tag{A.13}$$

where

$$\begin{aligned} \hat{\delta}^{*} &= \frac{\sum_{t=T+1}^{T+h}F_{t-1}f_t}{\sum_{t=T+1}^{T+h}F_{t-1}^2} = \frac{c}{k_h} + \frac{\sum_{t=T+1}^{T+h}F_{t-1}e_t}{\sum_{t=T+1}^{T+h}F_{t-1}^2}, \\ \hat{\sigma}^2 &= h^{-1}Q^2\sum_{t=T+1}^{T+h}(f_t - \hat{\delta}^{*}F_{t-1})^2, \end{aligned} \tag{A.14}$$

or (A.13) can be written as

$$t_{\widetilde{F}}^{*} = \frac{c}{\hat{\sigma}k_h}Q\left(\sum_{t=T+1}^{T+h}F_{t-1}^2\right)^{1/2} + \frac{\sum_{t=T+1}^{T+h}F_{t-1}e_t}{\hat{\sigma}Q^{-1}\left(\sum_{t=T+1}^{T+h}F_{t-1}^2\right)^{1/2}}. \tag{A.15}$$

We also have

$$\begin{aligned} \hat{\sigma}^2 &= h^{-1}Q^2\sum_{t=T+1}^{T+h}(f_t - \hat{\delta}^{*}F_{t-1})^2, \\ &= h^{-1}Q^2\sum_{t=T+1}^{T+h}\left(f_t - \frac{c}{k_h}F_{t-1} - \frac{\sum_{t=T+1}^{T+h}F_{t-1}e_t}{\sum_{t=T+1}^{T+h}F_{t-1}^2}F_{t-1}\right)^2, \\ &= h^{-1}Q^2\sum_{t=T+1}^{T+h}\left(e_t - \alpha^{-h}k_h^{-1}\frac{\alpha^{-h}k_h^{-1}\sum_{t=T+1}^{T+h}F_{t-1}e_t}{\alpha^{-2h}k_h^{-2}\sum_{t=T+1}^{T+h}F_{t-1}^2}F_{t-1}\right)^2, \\ &= h^{-1}Q^2\sum_{t=T+1}^{T+h}e_t^2 - 2h^{-1}Q^2D_h\underbrace{\alpha^{-h}k_h^{-1}\sum_{t=T+1}^{T+h}F_{t-1}e_t}_{=o_p(1)\text{ by Lemma A5 (c)}} + h^{-1}Q^2D_h^2\underbrace{\alpha^{-2h}k_h^{-2}\sum_{t=T+1}^{T+h}F_{t-1}^2}_{=O_p(1)\text{ by Lemma A5 (d)}}. \end{aligned}$$

However, since $D_h = \frac{\alpha^{-h}k_h^{-1}\sum_{t=T+1}^{T+h}F_{t-1}e_t}{\alpha^{-2h}k_h^{-2}\sum_{t=T+1}^{T+h}F_{t-1}^2} = o_p(1)$ by Lemma A5 (c) and (d), we obtain

$$\hat{\sigma}^2 = h^{-1}Q^2\sum_{t=T+1}^{T+h}e_t^2 + o_p(1) \xrightarrow{p} Q^2\sigma^2. \tag{A.16}$$

(i) We consider the case of $\alpha = 1$ or $c = 0$. In this case, the first term of (A.15) disappears. The t test becomes

$$\begin{aligned} t^*_{\widetilde{F}} &= \frac{h^{-1}\sum_{t=T+1}^{T+h} F_{t-1}e_t}{\hat{\sigma}Q^{-1}\left(h^{-2}\sum_{t=T+1}^{T+h} F_{t-1}^2\right)^{1/2}}, \\ &\Rightarrow \frac{\int_0^1 W(r)dr}{\left[\int_0^1 W(r)^2 dr\right]^{1/2}}, \end{aligned}$$

as $h \to \infty$, by using (A.16).

(ii) We now consider the case of $c \neq 0$ and Assumption 5 holds with $k_h = h$. The coefficient estimate (A.14) becomes

$$\begin{aligned} h\hat{\delta}^* &= c + \frac{h^{-1}\sum_{t=T+1}^{T+h} F_{t-1}e_t}{h^{-2}\sum_{t=T+1}^{T+h} F_{t-1}^2}, \\ &\Rightarrow c + \frac{\int_0^1 W_c(r)dr}{\int_0^1 W_c(r)^2 dr}, \end{aligned}$$

as $h \to \infty$.

(iii) We finally consider the case of $c > 0$ but $(1 + \frac{c}{k_h})^h = \alpha^h \to \infty$ as $h \to \infty$. If we consider the t test statistic (A.15) scaled by α^{-h},

$$\begin{aligned} \alpha^{-h} t^*_{\widetilde{F}} &= \frac{c}{\hat{\sigma}} Q \left(\alpha^{-2h} k_h^{-2} \sum_{t=T+1}^{T+h} F_{t-1}^2\right)^{1/2} + \alpha^{-h} \underbrace{\frac{\alpha^{-h} k_h^{-1}\sum_{t=T+1}^{T+h} F_{t-1}e_t}{\hat{\sigma}Q^{-1}\left(\alpha^{-2h}k_h^{-2}\sum_{t=T+1}^{T+h} F_{t-1}^2\right)^{1/2}}}_{=o_p(1) \text{ by Lemma A5 (c) and (d)}}, \\ &\Rightarrow \sqrt{\frac{c}{2\sigma^2}}\,|\Theta| > 0, \end{aligned}$$

by using Lemma A5 (d). Therefore, the result follows. This proves the result for the common test.

We can follow the same steps to derive the results for the idiosyncratic test of parts (1), (ii), and (iii). This is because

$$\widetilde{U}_{it} = U_{it} - \lambda_i H^{-1} \underbrace{(\widetilde{F}_t - HF_t)}_{=o_p(1)} - \underbrace{(\hat{\lambda}_i^* - \lambda_i H^{-1})}_{=O_p(\frac{1}{\min\{\sqrt{T}, N\}})} \widetilde{F}_t \xrightarrow{p} U_{it},$$

as $N, T \to \infty$ by using Lemma 1 (b) and (c) of Bai and Ng (2004). The rest of the proof follows that of the common test (Q becomes identity in this case).

References

Bai, J., 2003, "Inferential theory for factor models of large dimensions," *Econometrica* 71(1), 135-171.

Bai, J., 2004, "Estimating cross-section common stochastic trends in nonstationary panel data," *Journal of Econometrics* 122, 137-183.

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 cointergration," *Econometrica* 72(4), 1127-1177.

Bai, J. and S. Ng, 2006, "Confidence intervals for diffusion index forecasts and inference for factor-augmented regressions," *Econometrica* 74, 1133-1150.

Breitung, J. and S. Eickmeier, 2011, "Testing for structural breaks in dynamic factor models," *Journal of Econometrics* 163, 71-84.

Deng, A. and P. Perron, 2008, "A non-local perspective on the power properties of the CUSUM and CUSUM of squares tests for structural change," *Journal of Econometrics* 142, 212-240.

Dickey, D.A. and W.A. Fuller, 1979, "Distribution of the estimators for autoregressive time series with a unit root," *Journal of the American Statistical Association* 74, 427-431.

Engel. C., N. C. Mark, and K. D. West, 2015, "Factor model forecasts of exchange rates," *Econometric Reviews* 34(1-2), 32-55.

Fatum, R. and Y. Yamamoto, 2015, "Intra-safe haven currency behavior during the global financial crisis," forthcoming in *Journal of International Money and Finance.*

Greenaway-McGrevy, R., N.C. Mark, D. Sul, and J. L. Wu, 2015, "Identifying exchange rate common factors," University of Texas.

Gürkaynak, R.S., 2008, "Econometric tests of asset price bubbles: Taking stock," *Journal of Economic Surveys*, 22(1), 166-186.

Habib, M.M., and L. Stracca, 2012, "Getting beyond carry trade: what makes a safe haven currency?" *Journal of International Economics* 87, 50-64.

Lustig, H., N. Roussanov, and A. Verdelhan, 2011, "Common risk factors in currency market," *The Review of Financial Studies* 24(11), 3731-3777.

Nabeya, S. and K. Tanaka, 1990, "A general approach to the limiting distribution for estimators in time series regression with nonstable autoregressive errors," *Econometrica* 58, 145-164.

Perron, P., 1989, "The calculation of the limiting distribution of the autoregressive parameter in a near-integrated model," *Econometric Theory* 5, 241-255.

Perron, P., 1991, "A test for changes in a polynomial trend function for a dynamic time series," Econometric Research Program Research Memorandum No. 363, Princeton University.

Perron, P. and Y. Yamamoto, 2016, "On the usefulness or lack thereof of optimality criteria in structural change tests," *Econometric Reviews* 35(5) 782-844.

Phillips, P.C.B., 1987, "Towards a unified asymptotic theory for autoregression," *Biometrica* 74(3), 535-547.

Phillips, P.C.B., 1988, "Regression theory for near-integrated time series," *Econometrica* 56, 1021-1044.

Phillips, P.C.B. and T. Magdalinos, 2007, "Limit theory for moderate deviations from a unit root," *Journal of Econometrics* 136, 115-130.

Phillips, P.C.B. and T. Magdalinos, 2008, "Limit theory for explosively cointegrated systems," *Econometric Theory* 24, 865-887.

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 Economics Review* 56(4), 1043-1078.

Phillips, P.C.B., S. Shi, and J. Yu, 2015b, "Testing for multiple bubbles: limit theory of real time detectors," *International Economics Review* 56(4), 1079-1134.

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), 201-226.

Phillips, P.C.B., and J. Yu, 2011, "Dating the timeline of financial bubbles during the subprime crisis," *Quantitative Economics* 2, 455-491.

Robertson, D. and V. Sarafidis, 2015, "IV estimation of panels with factor residuals," *Journal of Econometrics* 185, 526-541.

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, 1167-1179.

Stock, J. H., and M. W. Watson, 2005, "Implications of dynamic factor models for VAR analysis," Working Paper 11467, National Bureau of Economic Research.

Vogelsang, T. J., 1999, "Sources of nonmonotonic power when testing for a shift in mean of a dynamic time series," *Journal of Econometrics* 88, 283-299.

Yamamoto, Y. and S. Tanaka, 2015, "Testing for factor loading structural change under common breaks," *Journal of Econometrics* 189(1), 187-206.

Figure 1: Size of the PANIC common test as a function of ρ_i

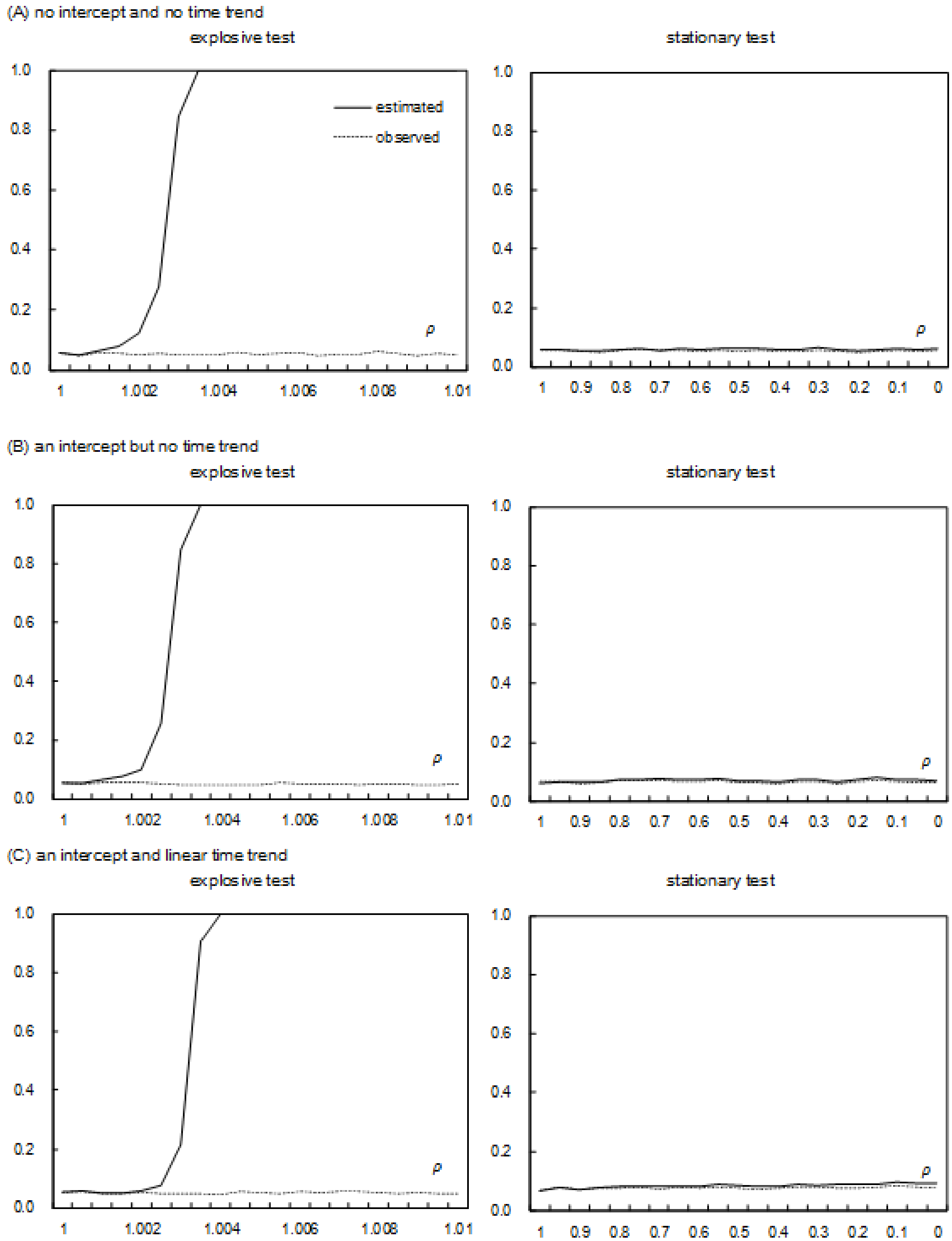


Figure 2: Size of the PANIC idiosyncratic test as a function of α

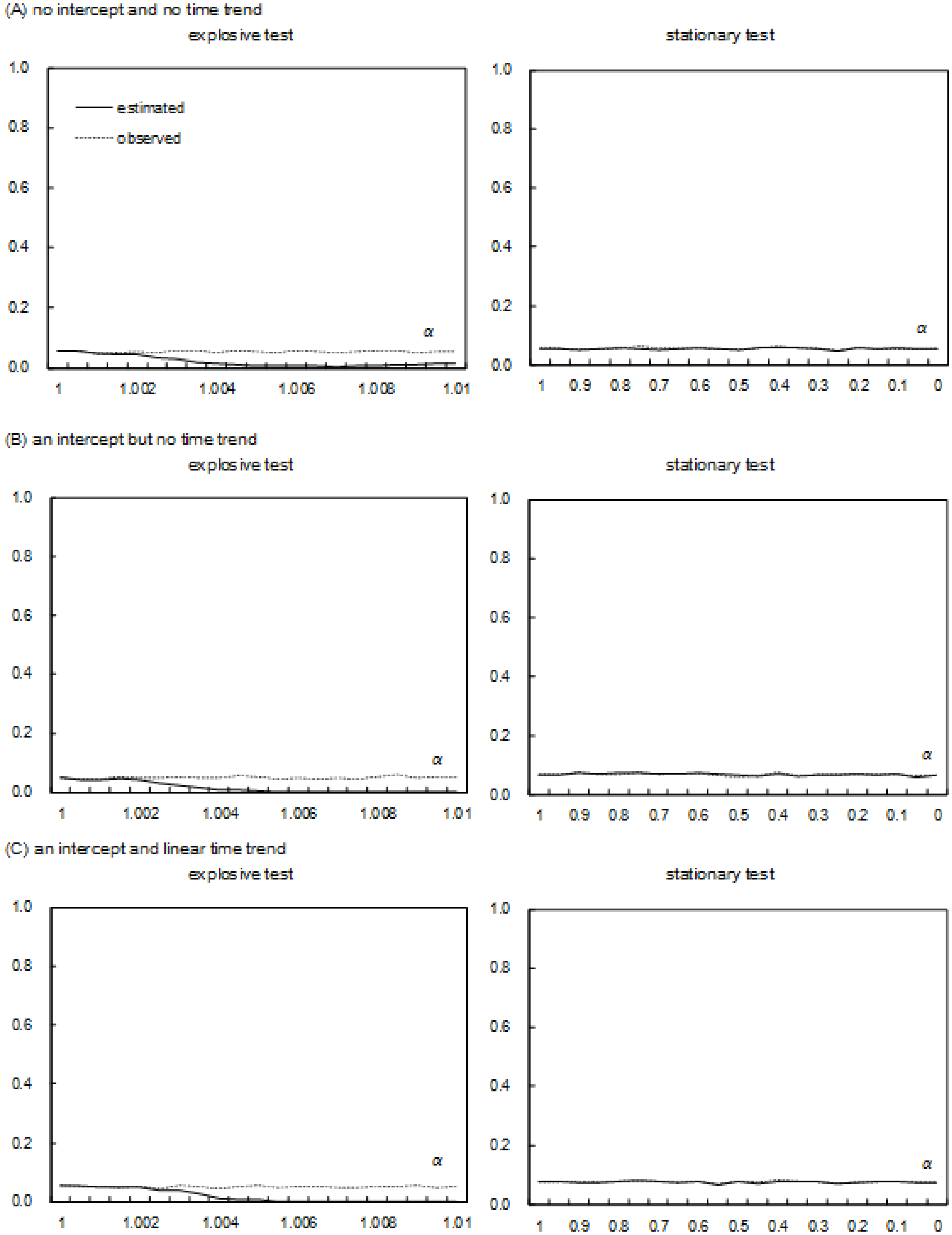


Figure 3: Power of the PANIC common test as a function of α

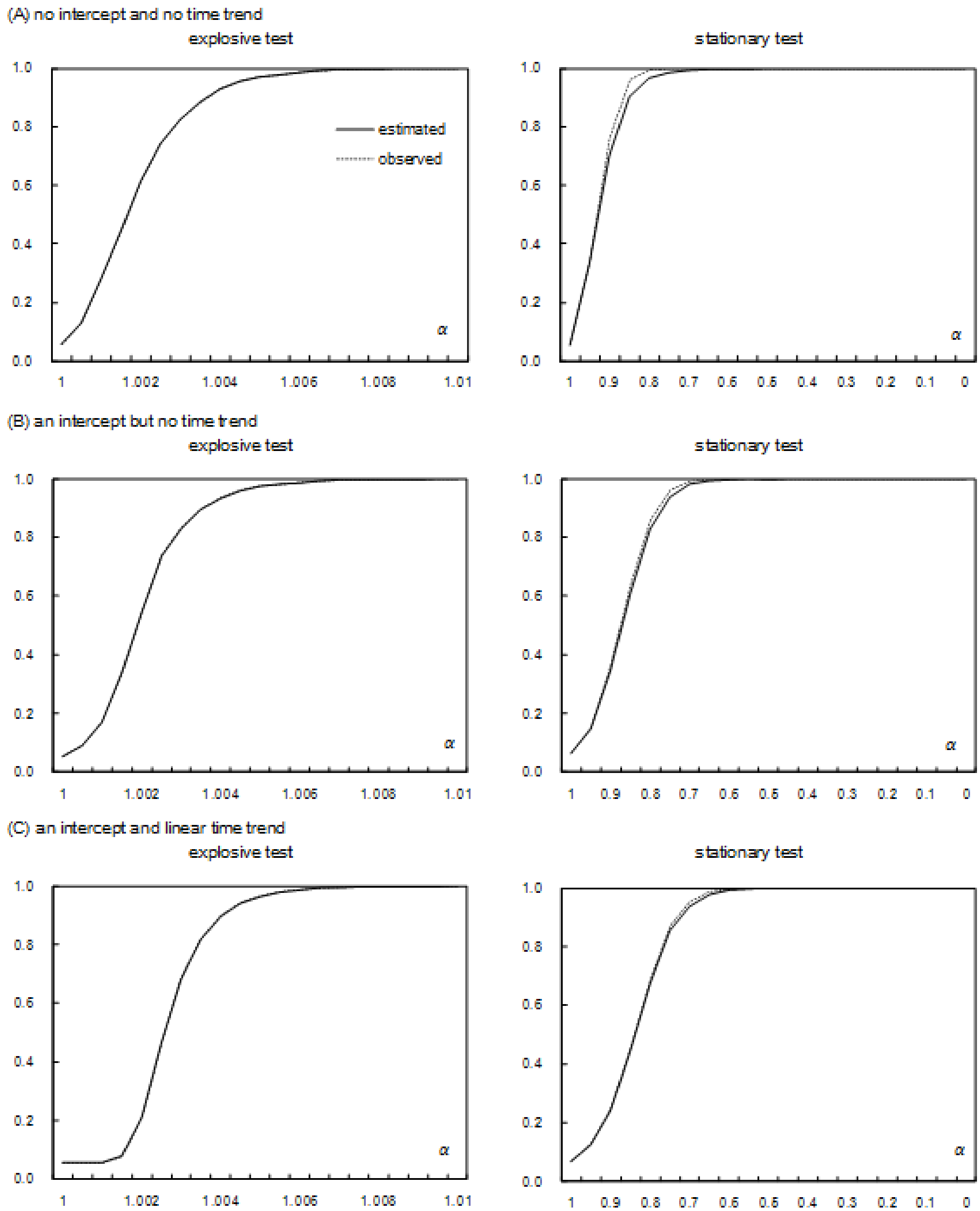


Figure 4: Power of the PANIC idiosyncratic test as a function of ρ_i

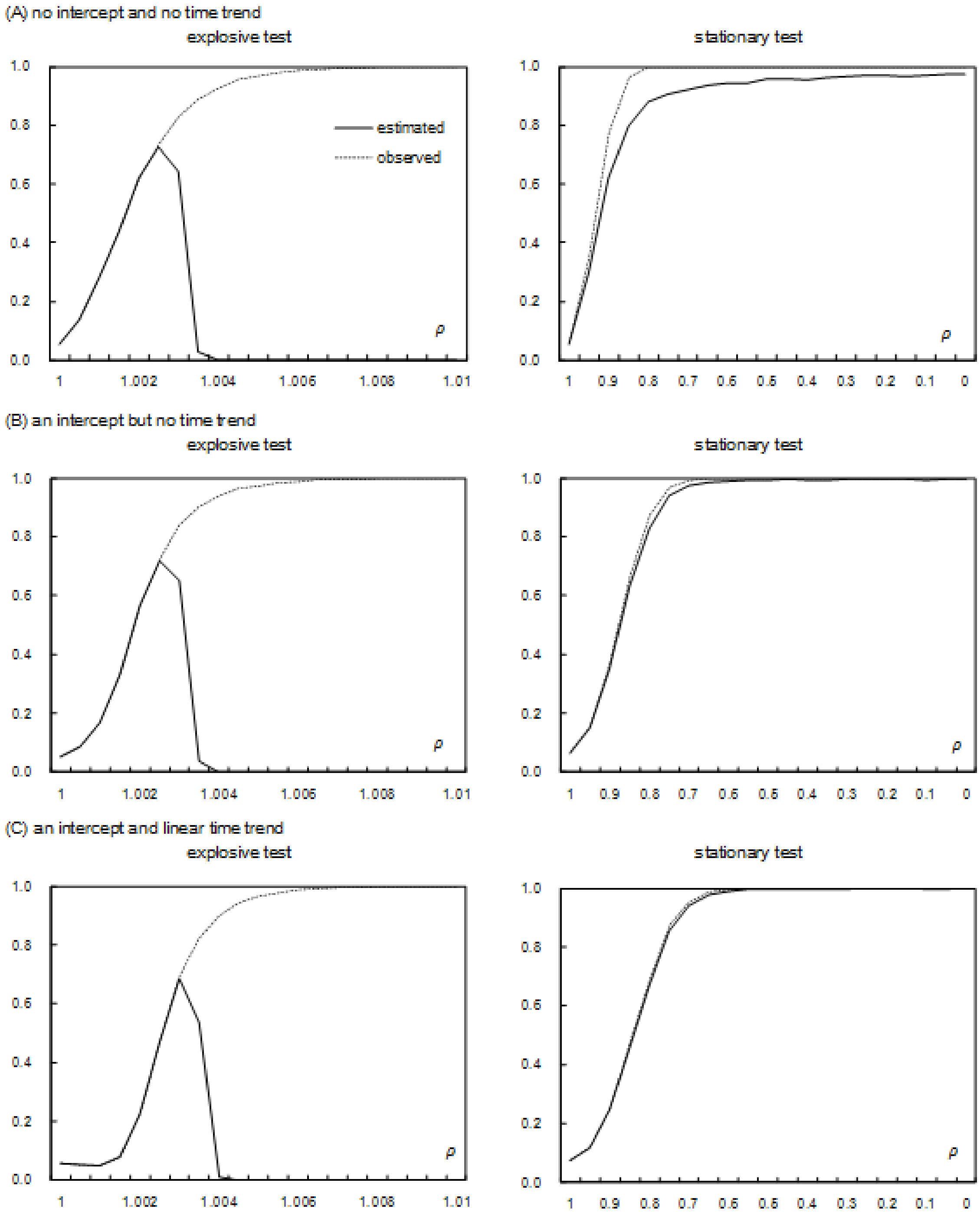
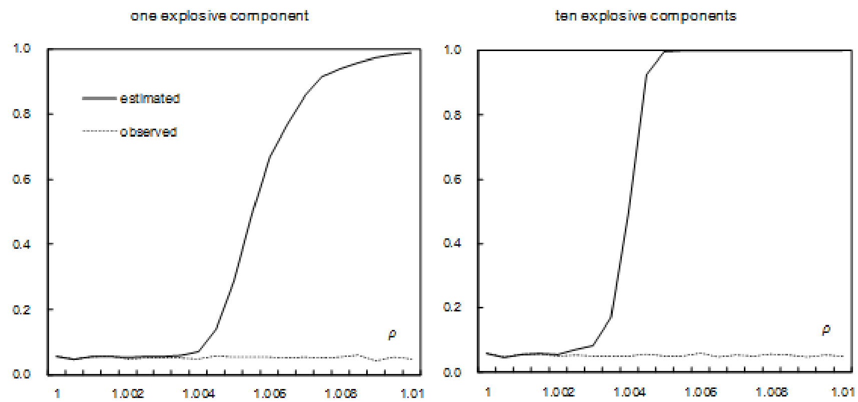
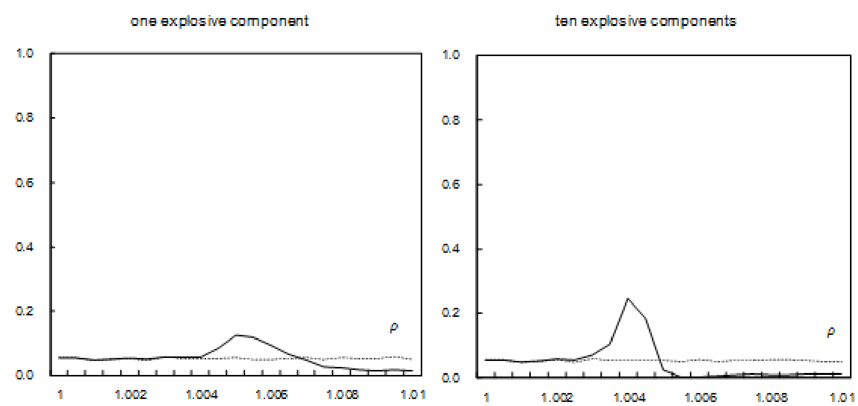


Figure 5. Size and power of the PANIC explosive tests when not all idiosyncratic components are explosive

Size of the common test as a function of ρ_1



Size of the idiosyncratic test as a function of ρ_1



Power of the idiosyncratic test as a function of ρ_1

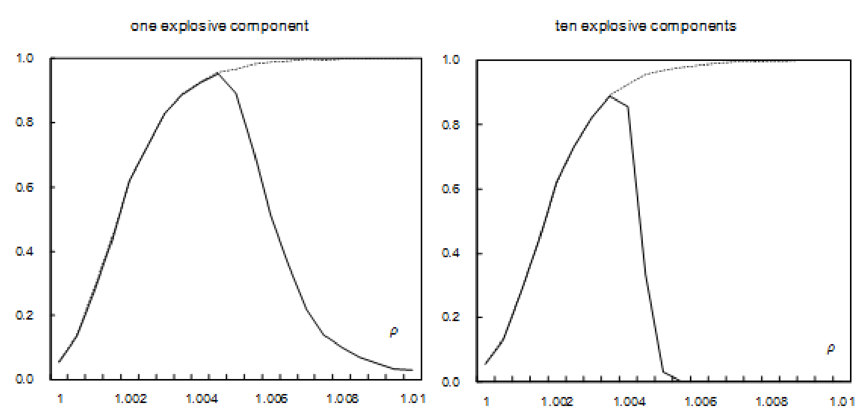


Figure 6-1. Correlation coefficients of the estimated common component
(with the true common factor and with the true idiosyncratic errors)

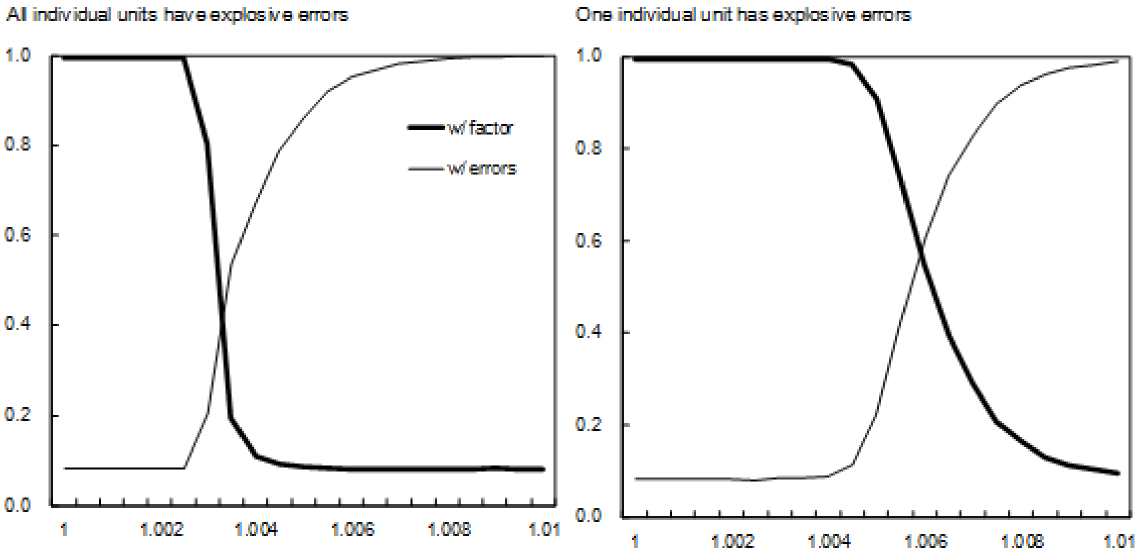


Figure 6-2. Correlation coefficients of the estimated individual component
(with the true common factor and with the true idiosyncratic errors)

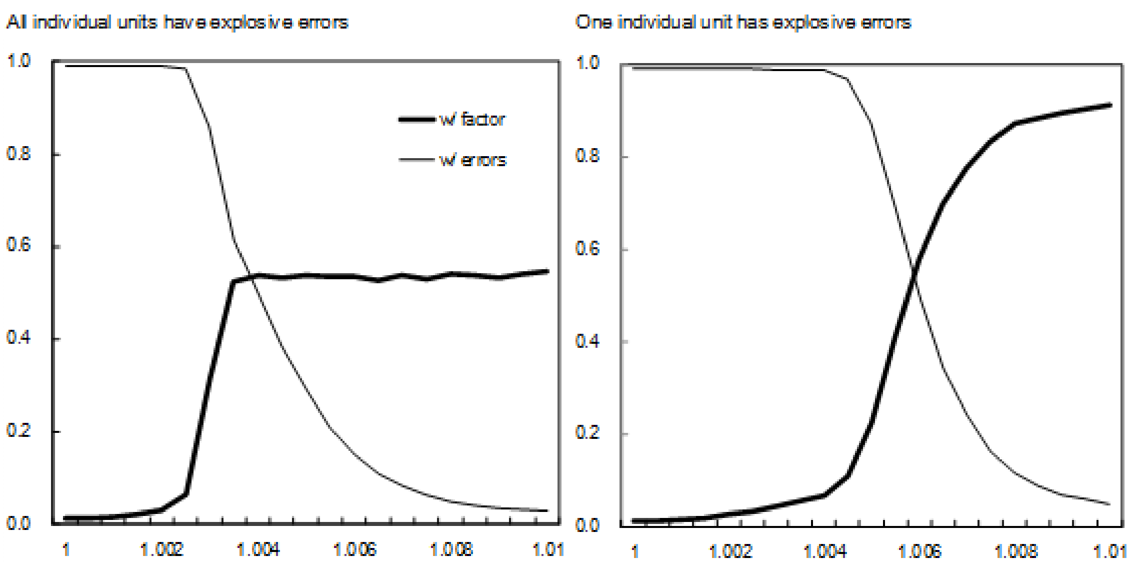


Figure 7. Size of the CS tests

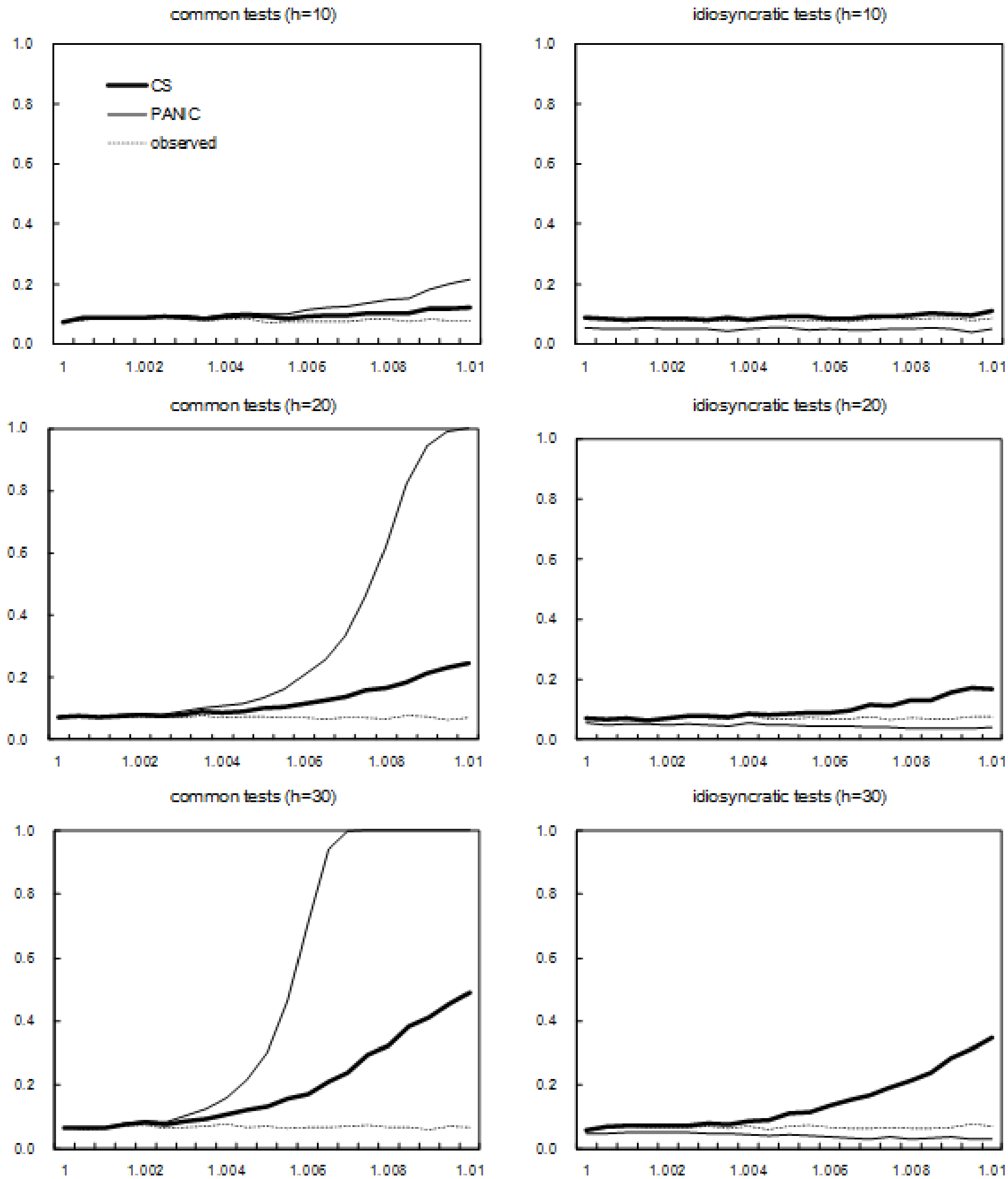


Figure 8. Power of the CS tests

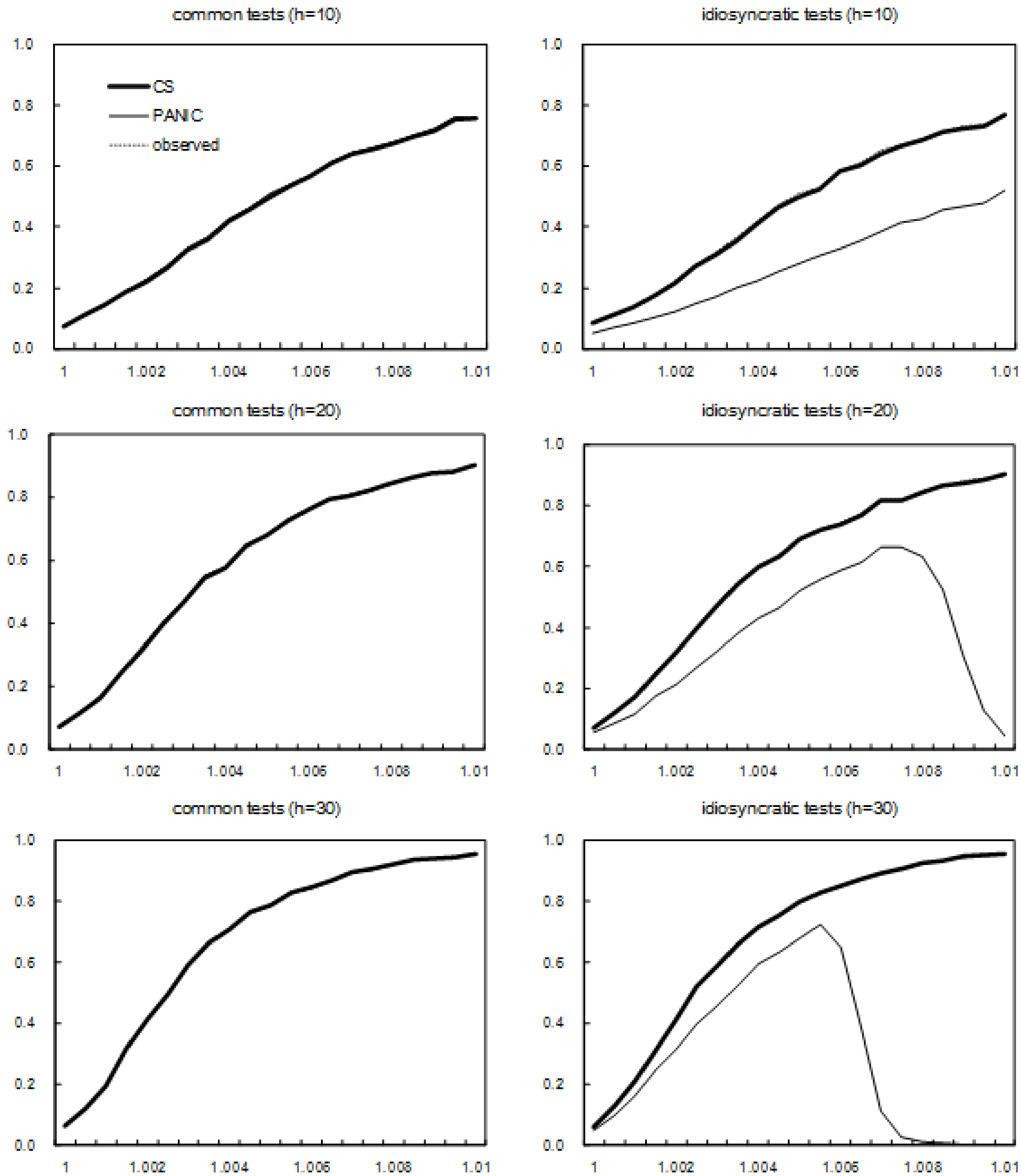


Figure 9. Exchange rates against the USD (in logarithm)

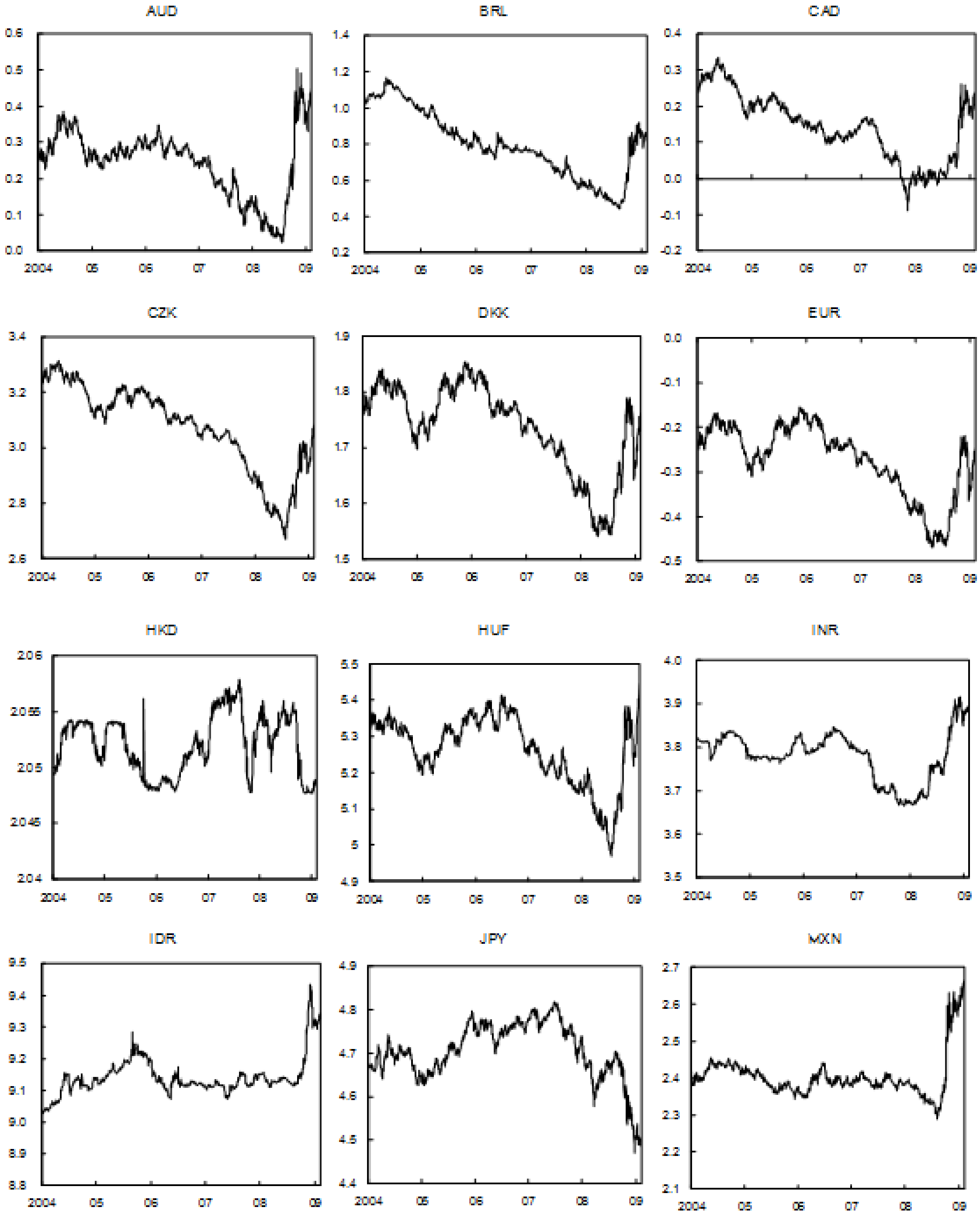


Figure 9. Exchange rates against the USD (in logarithm; continued)

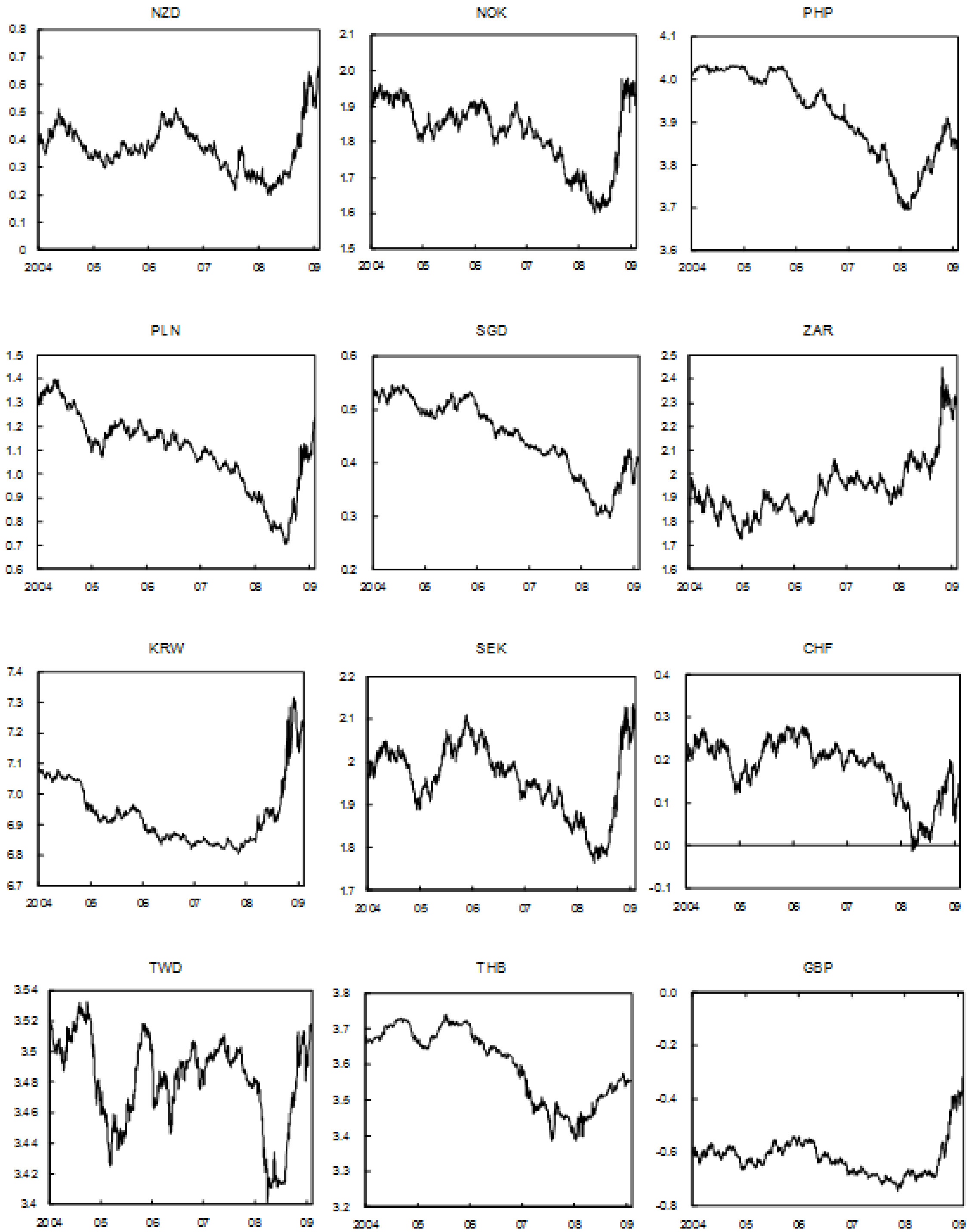


Table 1. Common and idiosyncratic explosive tests in the exchange-rate system

		individual	PANIC	CS
Common tests				
factor 1		-	0.38**	0.28**
factor 2		-	-1.68	-0.76
Idiosyncratic tests				
AUD	Australian Dollar	-0.16*	-1.80	-1.16
BRL	Brazilian Real	-0.48	-1.57	-2.66
CAD	Canadian Dollar	-0.23*	-1.39	-1.06
CZK	Czech Republic Koruna	-0.67	-2.27	-2.12
DKK	Danish Krone	-0.55	-1.06	-0.02**
EUR	Euro	-0.51	-1.19	-0.33*
HKD	Hong Kong Dollar	-2.48	-2.55	-1.83
HUF	Hungrian Forint	0.04**	-1.95	-1.78
INR	Indian Rupee	-0.03**	-0.71	-0.30*
IDR	Indonesian Rupiah	0.11**	0.06**	-0.27*
JPY	Japanese Yen	-0.81	-2.16	0.20**
MXN	Mexican Peso	0.20**	-2.24	-0.46
NZD	New Zealand Dollar	0.39**	-2.37	-1.47
NOK	Norwegian Krone	-0.22*	-1.06	-0.77
PHP	Philippines Peso	-1.00	-1.38	-1.19
PLN	Polish Zloty	0.42**	-1.14	-1.50
SGD	Singaporean Dollar	-1.32	-1.97	-2.25
ZAR	South African Rand	-0.48	-1.30	-0.35*
KRW	South Korean Won	-0.18*	-1.03	-0.38*
SEK	Swedish Krona	0.08**	-1.13	-0.91
CHF	Swiss Franc	-1.80***	-1.60	-0.02**
TWD	Taiwan Dollar	-0.31*	-1.08	-1.35
THB	Thai Baht	-1.63	-1.92	-2.18
GBP	UK Pound	0.50**	-0.54	-0.68

Notes: 1. The sample period is from August 1, 2007 to January 31, 2009. The CS sets the training period from January 1, 2004 to July 31,2007.
2. ***, **, and * denote significance at the 1%, the 5%, and the 10% level, respectively.