

Economics Department
Working Papers in Economics

Boston College

Year 2001

A PANIC Attack on Unit Roots and
Cointegration

Jushan Bai
Boston College,

Serena Ng
Johns Hopkins University,

A *PANIC* ATTACK ON UNIT ROOTS AND COINTEGRATION

Jushan Bai*

Serena Ng[†]

December 2001

Abstract

This paper develops a new methodology that makes use of the factor structure of large dimensional panels to understand the nature of non-stationarity in the data. We refer to it as PANIC— a ‘Panel Analysis of Non-stationarity in Idiosyncratic and Common components’. PANIC consists of univariate and panel tests with a number of novel features. It can detect whether the nonstationarity is pervasive, or variable-specific, or both. It tests the components of the data instead of the observed series. Inference is therefore more accurate when the components have different orders of integration. PANIC also permits the construction of valid panel tests even when cross-section correlation invalidates pooling of statistics constructed using the observed data. The key to PANIC is consistent estimation of the components even when the regressions are individually spurious. We provide a rigorous theory for estimation and inference. In Monte Carlo simulations, the tests have very good size and power. PANIC is applied to a panel of inflation series.

Keywords: Panel data, common factors, common trends, principal components

*Dept. of Economics, Boston College, Chestnut Hill, MA 02467 Email Jushan.Bai@bc.edu

[†]Dept. of Economics, Johns Hopkins University, Baltimore, MD 21218 Email: Serena.Ng@jhu.edu

This paper was presented at the NSF 2001 Summer Symposium on Econometrics and Statistics in Berkeley, California, the CEPR/Banca d'Italia Conference in Rome, and at NYU. We thank the seminar participants and the discussants (Andrew Harvey and Marco Lippi) for many helpful comments. We also thank Todd Clark for providing us with the inflation data. The first author acknowledges financial support from the NSF (grant SBR 9709508).

1 Introduction

Knowledge of whether a series is stationary or non-stationary is important for a wide range of economic analysis. As such, unit root testing is extensively conducted in empirical work. But in spite of the development of many elegant theories, the power of univariate unit root tests is severely constrained in practice by the short span of macroeconomic time series. Panel unit root tests have since been developed with the goal of increasing power through pooling information across units. But pooling is valid only if the units are independent, an assumption that is perhaps unreasonable given that many economic models imply, and the data support, the comovement of economic variables.

In this paper, we propose a new approach to understanding non-stationarity in the data, both on a series by series basis, and from the viewpoint of a panel. Rather than treating the cross-section correlation as a nuisance, we exploit these comovements to develop new univariate statistics and valid pooled tests for the null hypothesis of non-stationarity. Our tests are applied to two components of the data, one with the characteristic that it is strongly correlated with many series, and one with the characteristic that it is largely unit specific. More precisely, we consider a factor analytic model:

$$X_{it} = D_{it} + \lambda_i' F_t + e_{it}$$

where D_{it} is a polynomial trend function of order p , F_t is a $r \times 1$ vector of common factors, and λ_i is a vector of factor loadings. The series X_{it} is the sum of a deterministic component D_{it} , a common component $\lambda_i' F_t$, and an error e_{it} that is largely idiosyncratic. A factor model with N variables will have N idiosyncratic components but a small number of common factors.¹

A series with a factor structure is non-stationary if one or more of the common factors are non-stationary, or the idiosyncratic error is non-stationary, or both. Except by assumption, there is nothing that restricts F_t to be all I(1) or all I(0). There is also nothing that rules out the possibility that F_t and e_{it} are integrated of different orders. These are not merely cases of theoretical interest, but also of empirical relevance. As an example, let X_{it} be real output of country i . It may consist of a global trend component F_{1t} , a global cyclical component F_{2t} , and an idiosyncratic component (e_{it}) that may or may not be stationary. As another example, the inflation rate of durable goods may consist of a component that is common to all prices, and a component that is specific to durable goods.

It is well known that the sum of two time series can have dynamic properties very different from the individual series themselves. If one component is I(1) and one is I(0), it could be difficult

¹This is a static factor model, and is to be distinguished from the dynamic factor model being analyzed in Forni, Hallin, Lippi and Reichlin (2000).

to establish that a unit root exists from observations on X_{it} alone, especially if the stationary component is large. Unit root tests on X_{it} can be expected to be oversized while stationarity tests will have no power. The issue is documented in Schwert (1989), and formally analyzed in Pantula (1991), Ng and Perron (2001), among others, in the context of a negative moving-average component in the first-differenced data.

Instead of testing for the presence of unit roots in X_{it} , the approach proposed in this paper is to test the common factors and the idiosyncratic components separately. We refer such a Panel Analysis of Non-stationarity in the Idiosyncratic and Common components as *PANIC*. PANIC allows us to determine if nonstationarity comes from a pervasive or an idiosyncratic source. To our knowledge, there does not exist a test in the literature for this purpose. PANIC can also potentially resolve three econometric problems. The first is the size issue relating to summing series with different orders of integration just mentioned. The second is a consequence of the fact that the idiosyncratic components in a factor model can only be weakly correlated across i by design. In contrast, X_{it} will be strongly correlated across units if the data obey a factor structure. Thus, pooled tests based upon e_{it} are more likely to satisfy the cross-section independence assumption required for pooling. The third relates to power, and follows from the fact that pooled tests exploit cross-section information and are more powerful than univariate unit root tests.

Since the factors and the idiosyncratic components are both unobserved, and our objective is to test if they have unit roots, the key to our analysis is consistent estimation of these components irrespective of their stationarity properties. To this end, we propose a robust common-idiosyncratic (I-C) decomposition of the data using large dimensional panels. That is, datasets in which the number of observations in the time (T) and the cross-section (N) dimensions are both large. Loosely speaking, the large N permits consistent estimation of the common variation whether or not they are stationary, while a large T enables application of the relevant central limit theorems so that limiting distributions of the tests can be obtained. Robustness is achieved by a ‘differencing and re-cummulating’ estimation procedure so that I(1) and I(0) errors can be accommodated. Our results add to the growing literature on large dimensional factor analysis by showing how consistent estimates of the factors can be obtained using the method of principal components even without imposing stationarity on the errors.

Our framework differs from conventional multivariate time series models in which N is small. In small N analysis of cointegration, common trends and cycles, the estimation methodology being employed typically depends on whether the variables considered are all I(1) or all I(0).² Pretesting

²See, for example, King, Plosser, Stock and Watson (1991), Engle and Kozicki (1993), and Gonzalo and Granger (1995).

for unit roots is thus necessary. Because N is small, what is extracted is the trend or the cycle common to just a small number of variables. Not only is the information in many potentially relevant series left unexploited, consistent estimation of common factors is in fact not possible when the number of variables is small. In our analysis with N and T large, the common variation can be extracted without appealing to stationarity assumptions and/or cointegration restrictions. This makes it possible to decouple the extraction of common trends and cycles from the issue of testing stationarity.

The rest of the paper is organized as follows. In Section 2, we describe the PANIC procedures and present asymptotic results for the Dickey-Fuller t test of the unit root hypothesis. As an intermediate result, we establish uniform consistency of the factor estimates even when the individual regressions are spurious. As this result is important in its own right, we devote Section 3 to the large sample properties of the factor estimates. Section 4 uses simulations to illustrate the properties of the factor estimates and the tests in finite samples. PANIC is then applied to a panel of inflation data. Proofs are given in the Appendix.

2 PANIC

The data X_{it} are assumed to be generated by

$$X_{it} = c_i + \beta_i t + \lambda_i' F_t + e_{it}, \quad t = 1, \dots, T, \quad (1)$$

$$F_{mt} = \alpha_m F_{m,t-1} + u_{mt} \quad m = 1, \dots, r \quad (2)$$

$$e_{it} = \rho_i e_{i,t-1} + \epsilon_{it}, \quad i = 1, \dots, N. \quad (3)$$

Factor m is stationary if $\alpha_m < 1$. The idiosyncratic error e_{it} is stationary if $\rho_i < 1$. The objective is to understand the stationarity property of F_{mt} and e_{it} when these are all unobserved, and for which we estimate by the method of principal components.

When e_{it} is $I(0)$, the principal components estimators for F_t and λ_i have been shown to be consistent when all the factors are $I(0)$ and when some or all of them are $I(1)$. But consistent estimation of the factors when e_{it} is $I(1)$ has not been considered in the literature. Indeed, when e_{it} has a unit root, a regression of X_{it} on F_t is spurious even if F_t was observed, and the estimates of λ_i and thus of e_{it} will not be consistent. The validity of PANIC thus hinges on the ability to obtain estimates of F_t and e_{it} that preserve their orders of integration, both when e_{it} is $I(1)$ and when it is $I(0)$. We now outline a set of procedures that accomplish this goal. Essentially, the trick is to apply the method of principal components to the first differenced data. We show in this section that inference about unit roots is not affected by the fact that F_t and e_{it} are not observed.

We defer the discussion on the theoretical underpinnings of PANIC and the properties of factor estimates to Section 3 so as to keep unit root testing the main focus of this section.

We consider two specifications of the deterministic trend function, leading to what will be referred as the intercept only model and the linear trend model. We assume the number of common factors (r) is known.³ To simplify the proof, we let ϵ_{it} and u_t be serially uncorrelated. This allows us to consider the t statistic on the first order autoregressive parameter developed in Dickey and Fuller (1979). More general errors can be permitted, provided they satisfy the assumptions stated in Section 3. Remarks to this effect will be made below.

2.1 The Intercept Only Case

The factor model in the intercept only case is

$$X_{it} = c_i + \lambda'_i F_t + e_{it}. \quad (4)$$

We assume $E(\Delta F_t) = 0$. This is without loss of generality because if $F_t = a + \xi_t$ such that $E(\Delta \xi_t) = 0$, then $X_{it} = c_i + \lambda'_i a + \lambda'_i \xi_t + e_{it}$. The first differenced model $\Delta X_{it} = \lambda'_i \Delta \xi_t + \Delta e_{it}$ is thus observationally equivalent to $\Delta X_{it} = \lambda'_i \Delta F_t + \Delta e_{it}$. Denote

$$x_{it} = \Delta X_{it}, \quad f_t = \Delta F_t, \quad \text{and} \quad z_{it} = \Delta e_{it}. \quad (5)$$

Then the model in first-differenced form is:

$$x_{it} = \lambda'_i f_t + z_{it}. \quad (6)$$

The test statistics are constructed as follows:

1. Difference the data and estimate f_t and λ_i from (6) by the method of principal components. To be precise, let x be the $(T-1) \times N$ data matrix such that the i^{th} column is $(x_{i2}, x_{i3}, \dots, x_{iT})'$, $i = 1, 2, \dots, N$. Let $f = (f_2, f_3, \dots, f_T)'$ and $\Lambda = (\lambda_1, \dots, \lambda_N)'$. The principal component estimator of f , denoted \hat{f} , is $\sqrt{T-1}$ times the r eigenvectors corresponding to the first r largest eigenvalues of the $(T-1) \times (T-1)$ matrix xx' . The estimated loading matrix is $\hat{\Lambda} = x'\hat{f}/(T-1)$. Define $\hat{z}_{it} = x_{it} - \hat{\lambda}'_i \hat{f}_t$.
2. Given \hat{f}_t , define for each $m = 1, \dots, r$,

$$\hat{F}_{mt} = \sum_{s=2}^t \hat{f}_{ms}.$$

³Consistent estimation of r is possible using the method of Bai and Ng (2002) with data in differences. It can be shown that this will not affect the limiting distribution of the test statistics when the number of factors is estimated.

Test the null hypothesis that \widehat{F}_{mt} has a unit root for each $m = 1, \dots, r$ with an intercept included in the autoregression. Denote this test by $DF_F^c(m)$.

3. For each $i = 1, \dots, N$, denote by $DF_e^c(i)$ the test for the null hypothesis that there is a unit root in \widehat{e}_{it} , where $\widehat{e}_{it} = \sum_{s=2}^t \widehat{z}_{it}$ ($t = 2, \dots, T$). No deterministic term is included in the autoregressions.

Theorem 1 (*the intercept only case*): Suppose the data are generated by (4), (2) and (3), where ϵ_{it} , u_{mt} are mean zero processes with finite variances and are iid over t . Suppose Assumptions A-E stated in Section 3 hold. Let W_{um} ($m = 1, \dots, r$) and W_{ei} ($i = 1, \dots, N$) be standard Brownian motions, and denote their demeaned counterparts by W_{um}^c and W_{ei}^c , respectively.

1. Consider testing the null hypothesis that $\alpha_m = 1$ for $m = 1, \dots, r$. As $N, T \rightarrow \infty$,

$$DF_F^c(m) \Rightarrow \frac{\int_0^1 W_{um}^c(s) dW_{um}(s)}{\left(\int_0^r W_{um}^c(s)^2 ds \right)^{1/2}}.$$

2. Consider testing the null hypothesis that $\rho_i = 1$ for $i = 1, \dots, N$. As $N, T \rightarrow \infty$,

$$DF_e^c(i) \Rightarrow \frac{\int_0^1 W_{ei}(s) dW_{ei}(s)}{\left(\int_0^1 W_{ei}(s)^2 ds \right)^{1/2}}.$$

When the first differenced data ΔX_{it} contain no deterministic terms, $\frac{\widehat{F}_t}{\sqrt{T}} = H \frac{F_t}{\sqrt{T}} + o_p(1)$, where H is a full rank matrix. As shown in the Appendix, the $o_p(1)$ term is uniform in t . Testing for a unit root in demeaned \widehat{F}_t is asymptotically the same as testing for a unit root in demeaned F_t . The $DF_F^c(m)$ statistic thus has the same limiting distribution as derived in Fuller (1976) for the constant only case. The 5% asymptotic critical value is -2.86.

As also shown in the appendix, $\frac{1}{\sqrt{T}} \widehat{e}_{it} = \frac{1}{\sqrt{T}} e_{it} + o_p(1)$, where the $o_p(1)$ term is uniform in t . This is the underlying reason (though not the only one) that e_{it} can be replaced by \widehat{e}_{it} . Since no demeaning is performed, the asymptotic distribution $DF_e^c(i)$ is the DF unit root test with no constant. The critical value at the 5% significance level is -1.95.

2.2 The Linear Trend Case: $p = 1$

Consider now the factor model in the case of a linear trend:

$$X_{it} = c_i + \beta_i t + \lambda_i' F_t + e_{it} \tag{7}$$

and thus $\Delta X_{it} = \beta_i + \lambda'_i \Delta F_t + \Delta e_{it}$. Let $\overline{\Delta F} = \frac{1}{T-1} \sum_{t=2}^T \Delta F_t$, $\overline{\Delta e_i} = \frac{1}{T-1} \sum_{t=2}^T \Delta e_{it}$, and $\overline{\Delta X_i} = \frac{1}{T-1} \sum_{t=2}^T \Delta X_{it}$. Then

$$\Delta X_{it} - \overline{\Delta X_i} = \lambda'_i (\Delta F_t - \overline{\Delta F}) + (\Delta e_{it} - \overline{\Delta e_i})$$

which can be rewritten as

$$x_{it} = \lambda'_i f_t + z_{it}. \quad (8)$$

where

$$x_{it} = \Delta X_{it} - \overline{\Delta X_i}, \quad f_t = \Delta F_t - \overline{\Delta F}, \quad z_{it} = \Delta e_{it} - \overline{\Delta e_i}. \quad (9)$$

Note that the differenced and demeaned data x_{it} are invariant to c_i and β_i . As a consequence, there is no loss of generality to assume $E(\Delta F_t) = 0$. For example, if $F_t = a + bt + \xi_t$ such that $E(\Delta \xi_t) = 0$, then we can rewrite model (7) with F_t replaced by ξ_t and $c_i + \beta_i t$ replaced by $c_i + \lambda'_i a + (\beta_i + \lambda'_i b)t$.

The tests are constructed as follows:

1. Difference and demean the data. Compute the principal components estimator \hat{f} as $\sqrt{T-1}$ times the first r eigenvectors of the matrix xx' , and estimate Λ as $\hat{\Lambda} = x'\hat{f}/(T-1)$. Define $\hat{z}_{it} = x_{it} - \hat{\lambda}'_i \hat{f}_t$.
2. Given \hat{f}_t , define for each $m = 1, \dots, r$

$$\hat{F}_{mt} = \sum_{s=2}^t \hat{f}_{ms}.$$

Test the null hypothesis that \hat{F}_{mt} has a unit root around a deterministic trend using the Dickey-Fuller test with an intercept and a linear trend included in the autoregression. Denote this test by $DF_F^T(m)$.

- 3 Given \hat{z}_{it} , define

$$\hat{e}_{it} = \sum_{s=2}^t \hat{z}_{is}.$$

For each $i = 1, \dots, N$, denote by $DF_e^T(i)$ the test for the null hypothesis that there is a unit root in \hat{e}_{it} with no intercept or linear trend included in the autoregression.

Theorem 2 (the linear trend case): *Suppose the data are generated by (7), (2), and (3) and the assumptions of Theorem 1 hold. Let W_{um}^τ and W_{ei}^τ be detrended standard Brownian motions.*

1. Consider testing the null hypothesis that $\alpha_m = 1$ for $m = 1, \dots, r$. Then as $N, T \rightarrow \infty$,

$$DF_F^\tau(m) \Rightarrow \frac{\int_0^1 W_{um}^\tau(s) dW_{um}(s)}{\left(\int_0^1 W_{um}^\tau(s)^2 ds \right)^{1/2}},$$

where W_{um}^τ is a demeaned and detrended Brownian motion.

2. Consider testing for each $i = 1, \dots, N$ the null hypothesis that $\rho_i = 1$. Then as $N, T \rightarrow \infty$,

$$DF_e^\tau(i) \Rightarrow -\frac{1}{2} \left(\int_0^1 V_{ei}(s)^2 ds \right)^{-1/2},$$

where $V_{ei}(s) = W_{ei}(s) - sW_{ei}(1)$ is a Brownian bridge.

The limiting distribution of $DF_F^\tau(m)$ coincides with the Dickey Fuller distribution for the case with a constant and a linear trend. As shown in the Appendix, the consequence of having to demean ΔX_{it} is that $\frac{\hat{\epsilon}_{it}}{\sqrt{T}}$ converges to a Brownian bridge instead of a Brownian motion. The limiting distribution of the DF statistic is now proportional to the reciprocal of a Brownian bridge. This is the same distribution as the test considered in Schmidt and Lee (1991), a modified version of the LM test for the presence of a unit root around a linear trend developed in Schmidt and Phillips (1992).

Remarks:

1. Two desirable features of $DF_e^c(i)$ and $DF_e^\tau(i)$ are worthy of highlighting. First, the limiting distribution of the tests do not depend on the behavior of the common stochastic trends F_t . That is to say, the same distributions are obtained whether F_t is I(1) or I(0). Second, the limiting distributions do not depend on Brownian motions driven by u_t , where we recall that these are the innovations to the common factors. This property is useful for constructing pooled tests, which will be considered below.
2. If ϵ_{it} and u_{mt} are weakly dependent, the ADF test of Said and Dickey (1984) should be used instead. Theorems 1 and 2 remain valid provided the order of the augmented autoregression, M , is chosen such that $M^3/T \rightarrow 0$ as M and $T \rightarrow \infty$.
3. Asymptotic results are provided here for the DF test. However, PANIC can also be used in conjunction with other unit root tests. Using the Lemmas provided in the Appendix, the limiting distributions of any unit roots test can, in principle, be obtained by adapting the proofs.

4. A test on \widehat{F}_t can be constructed without demeaning or detrending. For $p = 0$, the limiting distribution is the same except that W_{um}^c is replaced by W_{um} . For $p = 1$, the limiting distribution is related to a Brownian bridge. It has the same form as $DF_e^\tau(i)$ reported in Theorem 2, part 2. In unreported simulations, the test performs better with demeaning and detrending.
5. A test on \widehat{e}_{it} can likewise be constructed by demeaning or detrending \widehat{e}_{it} . If this is done, the limiting distribution has the same form as $DF_e(i)$ in Theorem 1, part 2, but with $W_{ei}(r)$ replaced by the demeaned and the detrended Brownian motion, respectively. Overall, the test performs better without demeaning and detrending of the idiosyncratic errors.

2.3 Pooled Tests

Thus far, we have introduced PANIC as univariate tests on the common factors and the idiosyncratic components. A common criticism of univariate unit root tests is low power, especially when T is small. This has generated substantial interest to improve power. A popular method is to pool information across units, leading to panel unit root tests. Recent surveys of panel unit root tests can be found in Maddala and Wu (1999) and Baltagi and Kao (2001). The early test developed in Quah (1994) imposed substantial homogeneity in the cross section dimension. Subsequent tests such as that of Levin and Lin (1993) and Im, Pesaran and Shin (1997) allow for heterogeneous intercepts and slopes, while maintaining the assumption of independence across units. This assumption is restrictive, and if violated, can lead to over-rejections of the null hypothesis. Banerjee, Marcellino and Osbat (2001) argued against use of panel unit root test because of this potential problem. O'Connell (1998) provides a GLS solution to this problem, but the approach is theoretically valid only when N is fixed. When N also tends to infinity, as is the case under consideration, consistent estimation of the GLS transformation matrix is not a well defined concept since the sample cross-section covariance matrix will have rank T when $N > T$ even when the population covariance matrix is rank N .

If cross-section correlation can be represented by common factors, then Theorems 1 and 2 show that univariate tests based on \widehat{e}_{it} do not depend on Brownian motions driven by the common innovations u_t asymptotically. Thus, if e_{it} is independent across i , tests based upon \widehat{e}_{it} are asymptotically independent across i . This leads to the following.

Theorem 3 *Suppose e_{it} is independent across i and consider testing $H_0 : \rho_i = 1 \ \forall i$ against $H_1 : \rho_i < 1$ for some i . Then unit root tests based upon \widehat{e}_{it} can be pooled.*

Theorem 3 can be used to obtain pooled tests, such as that developed by Im et al. (1997), provided

they are based upon \widehat{e}_{it} . A test we consider is the following. Let $p^c(i)$ and $p^\tau(i)$ be the p-values associated with $DF_e^c(i)$ and $DF_e^\tau(i)$, respectively. Consider the pooled tests $P^c = -2 \sum_{i=1}^N \log p^c(i)$ and $P^\tau = -2 \sum_{i=1}^N \log p^\tau(i)$. Then

$$\frac{P^c - 2N}{\sqrt{4N}} \Rightarrow N(0, 1) \quad \text{and} \quad \frac{P^\tau - 2N}{\sqrt{4N}} \Rightarrow N(0, 1).$$

Under the assumption that e_{it} is independent across i , tests based on \widehat{e}_{it} are independent across i asymptotically. The p -values are thus independent $U[0,1]$ random variables. This implies that minus two times the logarithm of the p -value is a χ^2 random variable with two degrees of freedom. The test $-2 \sum_{i=1}^N \ln p(i)$ was first proposed in Maddala and Wu (1999) for fixed N . Choi (2001) extended the analysis to allow $N \rightarrow \infty$ by standardization. An implication of Theorem 4 is that pooling over p values of tests based on the estimated residuals are also asymptotically normal upon standardization. Pooling on the basis of p values is widely used in meta analysis. It has advantage of allowing for as much heterogeneity across units as possible. For example, it can be used even when the panel is non-balanced.

Pooled tests based upon \widehat{e}_{it} can be seen as a panel test of no cointegration, as the null hypothesis that $\rho_i = 1$ for every i holds only if no stationary combination of X_{it} can be formed. It differs from other panel cointegration tests in the literature, such as developed in Pedroni (1995), in that our framework is based on a large N , and the test is applied to \widehat{e}_{it} instead of X_{it} . While panel unit root tests based upon X_{it} is inappropriate if the data admit a factor structure, pooling of tests based upon \widehat{e}_{it} is asymptotically valid under the more plausible assumption that e_{it} is independent across i . It should be made clear that univariate tests proposed in Theorems 1 and 2 permit weak cross-section correlation of the idiosyncratic errors. It is only in developing pooled tests that independence of the idiosyncratic errors is necessary. The independence assumption can, in principle, be relaxed by restricting the number of cross-correlated errors to be finite so that as N increases, the p values are averaged over infinitely many units that are not cross-correlated.

2.4 Cointegration Tests

The proposed unit root test on \widehat{e}_{it} is valid whether F_t is $I(1)$ or $I(0)$. This convenient result holds whether the polynomial trend function is of order 0 or 1. However, when some or all of the factors are $I(1)$, a new test can also be obtained as follows. In the intercept only case, define \widehat{e}_{it}^1 as the estimated residual from the level regression

$$X_{it} = c_i + \lambda_i' \widehat{F}_t + e_{it}^1$$

where \widehat{F}_t is defined in the previous section. In the linear trend case, define \widehat{e}_{it}^1 as the estimated residual from the regression

$$X_{it} = c_i + \beta_i t + \lambda_i' \widehat{F}_t + e_{it}^1.$$

Denote the DF test on \widehat{e}^1 by $DF_{e_1}^c(i)$ and $DF_{e_1}^r(i)$, respectively.

Theorem 4 *Suppose some $\bar{r} \leq r$ of the factors are $I(1)$. Then $DF_{e_1}^c(i)$ and $DF_{e_1}^r(i)$ have the same limiting distributions as the residuals based cointegration test of Phillips and Ouliaris (1990) with \bar{r} integrated regressors, plus a constant and a time trend in the cointegrating regression, respectively.*

When \bar{r} common factors are assumed to be $I(1)$, testing the null hypothesis that e_{it} is $I(1)$ is the same as testing the null hypothesis that X_{it} does not cointegrate with \bar{r} integrated factors. The $DF_e(i)$ statistic then becomes the residuals based test for no stochastic cointegration with \bar{r} regressors. The asymptotic distributions are derived in Phillips and Ouliaris (1990). When an intercept is in the regression, the critical values for $DF_{e_1}^c(i)$ at the 5% level are -3.37 and -3.77 for $\bar{r} = 1$ and 2, respectively. If, in addition, the cointegrating regression has a time trend, the 5% critical values are -3.73 and -4.07 for $\bar{r} = 1, 2$, respectively.

An important difference between tests based on \widehat{e}_{it} and \widehat{e}_{it}^1 is that the limiting distributions of the former are asymptotically independent across i , but the latter are not. Essentially, $I(1)$ variables have non-vanishing effects on estimated residuals of a cointegrating regression. In consequence, the limiting distributions of residuals based cointegration tests are functions of the projection matrix of $I(1)$ variables. Since the $I(1)$ variables in the present context are the common trends, neither $DF_{e_1}^c(i)$ nor $DF_{e_1}^r(i)$ is asymptotically independent across i . More importantly, these statistics cannot be pooled.

3 Consistency of \widehat{F}_t

The asymptotic results stated in the previous section require consistent estimation of F_t and e_{it} when some, none, or all of these components are $I(1)$. In this section, we first state the assumptions underlying these results. We then discuss the properties of the factor estimates and provide two convergence results.

3.1 Assumptions

Let M be a generic positive number, not depending on T and N . Let f_t and z_{it} be defined by (5). *Assumption A:*

1. $E\|f_t\|^4 \leq M$,

2. $\frac{1}{T} \sum_{t=1}^T f_t f_t' \xrightarrow{p} \Sigma_f$, a $r \times r$ positive definite matrix;
3. $E(f_t) = 0$ and $\max_{1 \leq k \leq T} \frac{1}{\sqrt{T}} \|\sum_{t=1}^k f_t\| = O_p(1)$;
4. $\frac{1}{T^2} \sum_{t,s,u,v=1}^T |E(f_{jt} f_{js} f_{ju} f_{jv})| \leq M$, for $j = 1, 2, \dots, r$.

Assumption B: $E\|\lambda_i\|^4 \leq M$, and $\frac{1}{N} \sum_{i=1}^N \lambda_i \lambda_i' \xrightarrow{p} \Sigma_\Lambda$, a $r \times r$ positive definite matrix.

Assumption C:

1. $E(z_{it}) = 0$, $E|z_{it}|^8 \leq M$;
2. $E(N^{-1} \sum_{i=1}^N z_{is} z_{it}) = \gamma_N(s, t)$, $\sum_{s=1}^T |\gamma_N(s, t)| \leq M$ for all t .
3. $E(z_{it} z_{jt}) = \tau_{ij,t}$ with $|\tau_{ij,t}| \leq |\tau_{ij}|$ for some τ_{ij} and for all t , and $\sum_{i=1}^N |\tau_{ij}| \leq M$ for all j .
4. $E(\max_{1 \leq k \leq T} \frac{1}{\sqrt{TN}} \|\sum_{t=1}^k \sum_{i=1}^N \lambda_i z_{it}\|) \leq M$
5. $E\left|N^{-1/2} \sum_{i=1}^N [z_{is} z_{it} - E(z_{is} z_{it})]\right|^4 \leq M$, for every (t, s) ;
 $E\left(\max_{1 \leq k \leq T} \frac{1}{\sqrt{NT}} \left|\sum_{t=1}^k \sum_{i=1}^N (z_{is} z_{it} - E(z_{is} z_{it}))\right|\right)^2 \leq M$, for every s .

Assumption D: The idiosyncratic errors e_{it} , the common factors F_t , and the loadings λ_i are three mutually independent groups.

Assumption E: For every $m = 1, \dots, r$, $E\|F_{m0}\| \leq M$, and for every $i = 1, \dots, N$, $E\|e_{i0}\| \leq M$.

These conditions are sufficient for unit root inference to be unaffected by the replacement of F_t and e_{it} with their principal component estimates. Assumptions A-D have previously been used to analyze approximate factor models, see, for example, Stock and Watson (1998). However, Assumptions E, A3 and A4 are new to this paper. Assumption E is an initial condition assumption made commonly in unit root analysis. A3 is weaker than the requirement of functional central limit theorem for f_t . A4 is a weak serial-correlation condition imposed on the factors expressed in terms of the fourth moment. A2 together with B imply the existence of r factors. Assumption C2 allows some weak serial-correlation of z_{it} for every given i . Assumption C3 allows weak cross-section correlation for every given t . Under covariance stationarity, C3 implies a bound on the largest eigenvalue of the covariance matrix of e_{it} , making the model an ‘approximate factor model’ in the sense of Chamberlain and Rothschild (1983). For C4, let $\phi_{Nt} = \frac{1}{\sqrt{N}} \sum_{i=1}^N \lambda_i z_{it}$. The weak cross-section correlation together with the moment conditions on λ_i and z_{it} imply $E\|\phi_{Nt}\|^4 \leq M$ for all t . Thus if z_{it} are independent over t so that ϕ_{Nt} are also independent over t , then C4 can be shown to be true, say by the Doob’s inequality. Of course, independence is not necessary for C4. The explanation for the second requirement in C5 is similar. Assumptions A and C could also have been stated in terms of the fundamental innovations u_t and ϵ_{it} , but additional Lemmas would then be required to link them to f_t and z_{it} , resulting in repetitive statements.

3.2 Estimation of F_t

Existing proofs for consistency of the factor estimates assume that the idiosyncratic components are all $I(0)$. In particular, Bai and Ng (2002) considered estimation of r and showed that the squared deviation between the estimated factors and the true factors vanish, while Bai (2001a) derived the asymptotic distributions for the estimated F_t and λ_i . Both studies assume all variables are $I(0)$, but we need consistent estimates not just when e_{it} is $I(0)$, but also when it is $I(1)$.

The insight of the present analysis is that, by applying the method of principal components to the first differenced data, it is possible to obtain consistent estimates of F_t and e_{it} , regardless of the dynamic properties of F_t and e_{it} . To sketch the idea why this is the case, assume $\beta_i = 0$ and consider the factor model in differenced form:

$$\Delta X_{it} = \lambda_i' \Delta F_t + \Delta e_{it}.$$

Clearly, differencing removes the fixed effect c_i . This is desirable because a consistent estimate of it cannot be obtained when e_{it} is $I(1)$. Now if e_{it} is $I(1)$, $\Delta e_{it} = z_{it} \equiv \epsilon_{it}$ will be $I(0)$. If ϵ_{it} has weak cross-section and serial correlation, then provided they satisfy Assumption C, consistent estimates of ΔF_t can be obtained. If e_{it} is $I(0)$ but satisfy Assumption C, then Δe_{it} , although over-differenced, still satisfy those conditions, so consistent estimation of ΔF_t can once again be shown. We summarize these arguments in the following Lemma.

Lemma 1 *Let f_t and z_{it} be defined by (5). Consider estimation of (6) by the method of principal components and suppose Assumptions A to D hold. Then for $\delta_{NT} = \min\{\sqrt{N}, \sqrt{T}\}$, there exists an H with rank r such that as $N, T \rightarrow \infty$,*

- (a) $\frac{1}{T} \sum_{t=2}^T \|\hat{f}_t - H f_t\|^2 = O_p(\delta_{NT}^{-2})$,
- (b) $\min\{\sqrt{N}, T\}(\hat{f}_t - H f_t) = O_p(1)$, for each given t ,
- (c) $\min\{\sqrt{T}, N\}(\hat{\lambda}_i - H^{-1'} \lambda_i) = O_p(1)$, for each given i .

Although λ_i and f_t are not directly identifiable, there exists a $r \times r$ full rank matrix H such that \hat{f}_t is an estimate of $H f_t$ and $\hat{\lambda}_i$ is an estimate of $H^{-1'} \lambda_i$. Result (a) is proved in Bai and Ng (2002), while (b) and (c) are proved in Bai (2001a). It should be remarked that when e_{it} is $I(0)$, estimation using the data in level form will give a direct and consistent estimate on F_t . Although these estimates could be more efficient than the ones based upon first differencing, they are not consistent when e_{it} is $I(1)$.

In Pesaran and Smith (1995), it was shown that spurious correlations between two $I(1)$ variables do not arise in cross section regressions estimated with time averaged data under the assumption

of strictly exogenous regressors, iid errors, and T fixed. Phillips and Moon (1999) showed that an average long run relation, defined from long-run covariance matrices of a panel of $I(1)$ variables, can be identified when N and T are both large. Lemma 1 shows that the individual relations (not just the average) can be consistently estimated under a much wider range of conditions:- the regressors are unobserved, they can be $I(1)$ or $I(0)$, and the individual regressions may or may not be spurious.

Although consistent estimation of λ_i , f_t and z_{it} is implied by our previous works, the series we are interested in testing are $\widehat{F}_t = \sum_{s=2}^t \widehat{f}_s$ and $\widehat{e}_{it} = \sum_{s=2}^t \widehat{z}_{it}$. Thus, we need to show that given estimates of f_t and z_{it} , \widehat{F}_t and \widehat{e}_{it} are consistent for F_t and e_{it} , respectively.

Lemma 2 *Under the assumptions of Lemma 1,*

$$\max_{1 \leq k \leq T} \frac{1}{\sqrt{T}} \left\| \sum_{s=2}^t (\widehat{f}_s - Hf_s) \right\| = O_p(N^{-1/2}) + O_p(T^{-3/4}).$$

Loosely speaking, the lemma says that the cumulative sum of \widehat{f}_t is uniformly close to the cumulative sum of f_t provided $N, T \rightarrow \infty$.⁴ This result is instrumental in obtaining the limiting distributions of unit root tests for F_t . It would seem that for testing \widehat{e}_{it} , this result may not be sufficient since \widehat{e}_{it} also depends on $\widehat{\lambda}_i$. But as seen from the proof in the Appendix, we only require $(\widehat{\lambda}_i - H^{-1}\lambda_i)$ to be $o_p(1)$ for unit root tests on \widehat{e}_{it} to yield the same inference as testing e_{it} . But by Lemma 1c, this holds provided N and T tend to infinity. Thus, the conditions for valid testing of F_t and e_{it} using \widehat{F}_t and \widehat{e}_{it} are the same.

The above two Lemmas stated for the intercept only case extends to the linear trend model.

Lemma 3 *Suppose f_t and z_{it} defined in (5) satisfy Assumptions A-D. Then f_t and z_{it} defined in (9) also satisfy the same conditions. Consider estimating f and Λ from (8) by the method of principal components. Then Lemma 1 and Lemma 2 still hold.*

Uniform convergence of the factor estimates in large panels was proved in Stock and Watson (1998) under the assumption that $N \gg T^2$ and that F_t and e_{it} are stationary. However, a more general uniform consistency result can be obtained from Lemma 2. Because $\widehat{F}_t = \sum_{s=2}^t \widehat{f}_s$ and $H \sum_{s=2}^t f_t = H \sum_{s=2}^t \Delta F_s = HF_t - HF_1$, Lemma 2 implies (upon multiplying $T^{1/2}$ to each side)

$$\max_{1 \leq k \leq T} \left\| \widehat{F}_t - HF_t + HF_1 \right\| = O_p(T^{1/2}N^{-1/2}) + O_p(T^{-1/4}). \quad (10)$$

According to (10), \widehat{F}_t is uniformly consistent for HF_t (up to a shift factor HF_1), provided that $T/N \rightarrow 0$ when N and T go to infinity. This is obtained without knowing whether F_t or e_{it} are $I(0)$

⁴The $O_p(T^{-3/4})$ can be replaced by $O_p(\log T/T)$ if the moment generating function of f_t exists (i.e., if $Ee^{\tau \|f_t\|} \leq M$ for all t and for some $\tau > 0$).

or I(1), thus a very strong result. Since a location shift does not change the nonstationary property of a series, testing the demeaned process $\widehat{F}_t - \overline{\widehat{F}}$ is asymptotically the same as testing $H(F_t - \overline{F})$.

Result (10) is remarkable in that the common stochastic trends can be consistently estimated by the method of principal components, up to a rotation and a shift in level, regardless of the idiosyncratic errors being I(1) or I(0). This means that even if each cross-section equation is a spurious regression, the common stochastic trends are well defined and can be consistently estimated, if they exist. This is not possible within the framework of traditional time series analysis, in which N is fixed.

The result that when N and T are large, the entire vector of common factors can be consistently estimated under very general conditions is also of practical interest because it allows us to assess the relative importance of the common and the idiosyncratic components even when neither is observed. Because of its generality, (10) is stated as a result of independent interest. It should be made clear that uniform consistency is not necessary for PANIC, and thus we do not require $T/N \rightarrow 0$, though our results will hold under these stronger conditions. For PANIC to be valid, only Lemmas 1 and 2 are necessary.

Finally, a caveat about the number of non-stationary factors is in order. Because H is of full rank, each component of HF_t (and of \widehat{F}_t) will be I(1) if each component of F_t is I(1). Similarly, each component of HF_t (and of \widehat{F}_t) will be I(0) if each component of F_t is I(0). However, if only \bar{r} components ($0 < \bar{r} < r$) of F_t are I(1), more than \bar{r} components of HF_t could be I(1) since linear combinations of I(1) variables remain I(1). PANIC tests if each of the r estimated factors is non-stationary, which amounts to testing if the series formed by each row of HF_t is non-stationary. Thus the number of non-stationary \widehat{F}_t is an upper bound for the number of non-stationary F_t . If knowledge of \bar{r} is essential, methods that detect the number of cointegrating vectors in a small number of series, such as developed by Stock and Watson (1988) or Nyblom and Harvey (2000), can be applied to the vector series \widehat{F}_t . By Lemmas 1 and 2, such tests will also consistently determine \bar{r} .

4 Monte Carlo Simulations

We begin by using a model with one factor to show that \widehat{F}_t constructed as $\sum_{s=2}^t \widehat{f}_t$ is robust to different stationarity assumptions about e_{it} , where \widehat{f}_t is estimated from first differenced data. We generate F_t as an independent random walk of $N(0,1)$ errors with $T=100$. This is plotted in Figure 1 (a). Data are generated according to $X_{it} = \lambda_i F_t + e_{it}$, where λ_i is iid $N(1,1)$. We then construct \widehat{F}_t as discussed in Section 2 for the intercept only model. Since we can only identify F_t up to a location transformation and a scale shift, we consider a rotated estimate obtained as the fitted

value from the regression

$$F_t = a + b\widehat{F}_t + \text{error}.$$

An implication of Lemma 2 is that the rotated \widehat{F}_t should be an increasingly more precise estimate of F_t as N increases. On the other hand, estimation using the data in levels will not have this consistency property.

For the case where e_{it} is I(1), we simulate a random walk driven by iid $N(0,1)$ errors. Figures 1(b), (c) and (d) display the true factor process F_t along with (the rotated) \widehat{F}_t , estimated both by first differencing the data, and by applying the method of principal components to the data in level form, for $N=20, 50, 100$, respectively. Evidently, \widehat{F}_t gets closer to F_t as N increase if the data are differenced. In fact the rotated \widehat{F}_t is actually close to the true process even when $N = 20$. On the other hand, when the method of principal components is applied to levels of the same data, all the estimated series are far from the true series, showing that estimation using the data in levels is not consistent when e_{it} is I(1). We next assume the idiosyncratic errors are all I(0) by drawing e_{it} from an iid $N(0, 1)$ distribution. Figure 2 illustrates that even though the data are over differenced, the estimates are very precise. In this case, both the level and differenced methods give almost identical estimates.

Having shown that first differencing the data indeed yield precise estimates of the factors, we now use simulations to illustrate the finite sample properties of the proposed tests. We simulate data using Equations (1)-(3). The number of common factors r is set to 1. Data are generated with $\lambda_i \sim N(1, 1)$, $\epsilon_{it} \sim N(0, 1)$ and $u_t \sim N(0, \sigma_F^2)$. The following variations are considered:

- $\sigma_F^2=10, 1$, and $.5$.
- $(\rho, \alpha)=\{(.5,.8), (.8,.5), (0,.9), (.9,0), (1,0), (1,.5), (1,.8), (1,.9), (1,.95), (0,1), (.5,1), (.8,1), (.9,1), (.95,1)\}$;

We report results for $T=100$, and $N = 20, 100$. The number of replications is 1000. Table 1 presents results for the ADF test applied to $X_{it}, \widehat{F}_t, \widehat{e}_{it}$, and \widehat{e}_{it}^1 . The entries are to be interpreted as size if (i) F_t is tested and $\alpha = 1$, or (ii) \widehat{e}_{it} or \widehat{e}_{it}^1 is tested and $\rho = 1$, or (iii) X_{it} is tested and $\rho = 1$ or $\alpha = 1$. All other entries represent power. The augmented autoregressions have $M = 4(T/100)^{1/4}$ lags throughout. The column labelled \widehat{F} in Table 1 is the rejection rate of the ADF test applied to the estimated common factor. The remaining three columns are the average rejection rates, where the average is taken across N units over 1000 trials. Results for a particular i are similar.

The statistic for testing if X_{it} has a unit root should have a rejection rate that equals the nominal size of .05 when $\alpha = 1$ or $\rho = 1$. In finite samples, this is true only when $\rho = 1$ and σ^F is small. When $\sigma^F=10$ and $\alpha=.5$, for example, the ADF test rejects a unit root in X_{it} with probability

around .37 in the interecept model, and .4 in the linear trend model, even though $\rho = 1$. As noted earlier, testing for a unit root in X_{it} when it has components with different degrees of integration is difficult because of the negative moving average component in ΔX_{it} .

Turning now to unit root tests on \widehat{F}_t , the empirical size is close to the nominal size of .05 when α is 1 and exhibits little size distortion. At other values of α , it reflects power with rejection rates comparable to other unit root tests that are based on least squares detrending. This suggests that the error in estimating F_t is small. Thus, even when $N=20$, F_t can be treated as though it is known. Indeed, the results for $N = 20$ and $N = 100$ are very close except for small α . This supports our earlier claim that PANIC does not require a very large N . So the proposed approach can be used for many configurations of the data.

When $\rho = 1$, tests based on \widehat{e}_{it} or \widehat{e}_{it}^1 have rejection rates close to the nominal size of 5%. Although the tests have similar sizes, they have different power properties. Consider the power of the tests at $\rho=.9$. If $\alpha=0$, the ADF test using \widehat{e}_{it} rejects the null hypothesis with probability .46, much higher than the rejection rate of .1 when \widehat{e}_{it}^1 is used. For close alternatives, such as when $\alpha = 1$ and $.8 < \rho < 1$, power is always higher from testing \widehat{e}_{it} than \widehat{e}_{it}^1 . This is because the limiting distributions of tests for \widehat{e}_{it} coincide with tests that can be given an LM interpretation, and LM tests have better local power. Although the test based upon \widehat{e}_{it}^1 is more powerful for some parameter values, testing \widehat{e}_{it} does not require a priori judgement on whether F_t is I(1) or I(0), nor is it necessary to know how many of the common factors are non-stationary. Both features are desirable since the problem of testing if e_{it} has a unit root can be decoupled from whether the F_t s have unit roots. However, the more compelling argument for using \widehat{e}_{it} over \widehat{e}_{it}^1 is that tests based on \widehat{e}_{it} can be pooled.

Results for the pooled tests are reported in Table 2. The p -values required to construct the pooled tests are obtained as follows. We first simulate the asymptotic distributions reported in Theorems 1 and 2 by using partial sums of 500 $N(0,1)$ errors to approximate the standard Brownian motion in each of the 10,000 replications. A look-up table is then constructed to map 300 points on the asymptotic distributions to the corresponding p values. In particular, 100 points is used to approximate the upper tail, 100 to approximate the lower tail, and 100 points for the middle part of the asymptotic distributions.⁵

The results in Table 2a,b confirm that the standard method of pooling tests on X_{it} biases the tests toward rejecting the null hypothesis (see column 1). For example, when $\rho = .5$ and $\alpha = 1$, all N series are non-stationary in view of the common stochastic trend. The standard pooled test

⁵The p values match up very well Table 3 of MacKinnon (1994), whenever they are available. These look-up tables are available from the authors.

should have a rejection rate close to the nominal size. However, the rejection rate is .35. Consider also $\rho = 1$ and $\alpha = 0$. The common factor is iid. The pooled test has a size of .16 of .05 when σ^F is small and deteriorates to .96 when σ^F is large. These results are consistent with the findings of O’Connell (1998) that cross section correlation leads the standard pooled test to over-reject the null hypothesis.

In theory, pooling based on \widehat{e}_{it}^1 is invalid. In simulations, this pooled test has better properties than expected, though it has a distorted size when both the common and the idiosyncratic components have unit roots. They are also less powerful when ρ is close to one. On the other hand, the pooled test based on \widehat{e}_{it} is always very well-behaved even when N is as small as 20. It rejects the null hypothesis that every e_{it} has a unit root when they are in fact stationary with probability of almost one. When each of the e_{it} is non-stationary, the rejection rate roughly equals the size of the test. In view of the much documented size and power problem concerning unit root tests, these results are striking.

5 Application to Inflation

In this section, we examine whether the common factor and the idiosyncratic components of inflation are non-stationary. This exercise is interesting in its own right because there is no unique definition of inflation. For every available price series, an inflation rate can be computed. Yet, in most economic analysis, the object of interest is not the rate of growth of a particular price series, but that of the general price level. The growth rate of the general price level is sometimes referred to as ‘core inflation.’ The common factor in our analysis can be given this interpretation.

There exists a large body of research that attempts to measure core inflation⁶. In addition to the price indices used in the calculations, the various measures of core inflation also differ in the stationarity assumption. Quah and Vahey (1995) used a structural VAR to identify core inflation, assuming inflation is I(1). Bryan and Cecchetti (1993) extracted core inflation using a state space framework and assumed inflation rates are stationary. As discussed in Ng and Perron (1997), inflation data tend to have a negative moving average component, and this makes unit root testing imprecise. Indeed, empirical testing has failed to come to a consensus as to whether inflation is stationary. Some studies proceed assuming inflation is stationary, while others assume it is differenced stationary.⁷ Here, we provide a formal test of whether core inflation has a unit root without taking a stand on whether the data have unit roots. In fact, we use the dynamic properties of core inflation to help understand whether the observed inflation series are stationary.

⁶For a discussion on this and related issues, see Clark (2001) and Wynne (1999).

⁷See, for example Ireland (1999) and Stock and Watson (1999). Both rejected a unit root at the 10% level for some subsample but favored the unit root model in their empirical work.

We use a panel of 21 monthly price indices published by the Bureau of Labor Statistics continuously over the sample 1967-1997⁸ to construct a panel of monthly inflation. Thus, X_{it} is inflation, $i = 1, \dots, N = 21$ and $t = 1, \dots, T = 371$. All tests are performed using the intercept model, i.e. without a linear trend. The ADF test rejects a unit root in 8 inflation series. The more powerful DFGLS test of Elliott, Rothenberg and Stock (1996) rejects a unit root in 11 series.⁹

We then estimate one factor by applying the method of principal components to the standardized first difference of inflation. The standardization effectively downweights volatile series in calculating the principal components. Note that there is no need to demean the differenced data since the mean of differenced data is zero by assumption in the intercept model.¹⁰ The single factor explains roughly 10% of the total variation in the data. The last two columns of Tables 3 present estimates to gauge the relative importance of the common factor. The variance of $\Delta \hat{e}_{it}$ tends to be large relative to the variance of ΔX_{it} . Of the 21 series, only 8 have common factors explaining more than 5% of the variation in the series. This suggests that the idiosyncratic components in the inflation series are large and dominate the common components. This is confirmed by the fact that the variation of the common component relative to that of the idiosyncratic component is 10% or more in only six series. The ADF test for the common factor is -2.79, when the 5% critical value of -2.86. We cannot reject the null hypothesis that the common factor has a unit root at the 5% level. All inflation series are non-stationary in view of the presence of a unit root in the common factor.

We next apply the ADF test to \hat{e}_{it} . We can reject the unit root hypothesis in 7 series. In 6 of these cases, the unit root null is also rejected when \hat{e}_{it}^1 is tested. In each of the 7 cases that we reject a unit root in \hat{e}_{it} , the DFGLS also rejects a unit root in X_{it} even though there is a non-stationary common factor. In the 11 cases when the DFGLS cannot reject a unit root in X_{it} , the ADF also cannot reject a unit root in \hat{e}_{it} . The results are consistent with the presence of a common but small unit-root component. Direct testing of X_{it} becomes imprecise when the idiosyncratic component is stationary and large.

⁸Clark (2001) compared various measures of core inflation using 36 CPI series, but only 23 of them were available before 1980. The other two series, CPI and CPI ex-energy, are aggregates which we exclude from the analysis. This leaves 21 series. The sample ended in 1997 because of a change in the structure of the CPI data.

⁹The lag lengths selected by the MAIC for the DFGLS are between 8-12 (see Ng and Perron (2001)). The results reported in Table 3 are all based on 8 lags.

¹⁰The panel criterion developed in Bai and Ng (2002) was not used because N is too small for precise estimation of r .

6 Conclusion

This paper makes use of the observation that if a panel of data has a factor structure, then testing for the presence of a unit root in the common and the idiosyncratic terms separately should be more effective than unit root testing on the observed data. Accordingly, we first consider how the common factors can be consistently estimated irrespective of the stationarity property of the idiosyncratic errors. We then consider unit root tests based on the estimated data and show that inference about unit roots is not affected by the fact that the factors cannot be observed. While pooling is inappropriate when the observed data are cross-correlated, pooling over tests based on the idiosyncratic components are more likely to be valid. Simulations show that the proposed tests have good size and power properties even for panels with only 20 units.

The present analysis can be extended in several ways. The common-idiosyncratic decomposition enables inferential analysis in general. Thus, in addition to unit root tests, the null hypothesis of stationarity can also be considered. The deterministic terms in the factor model are estimated in the present paper by the method of least squares. As such, the unit root tests are implicitly based on least squares detrending. But as Elliott et al. (1996) showed, unit root tests based on GLS detrending are more powerful. The tests developed in this paper can potentially be improved along this dimension. In theory, the machinery developed in this paper can also be used to test the long memory, ARCH effects, and other time series features in the data. We leave this for future research.

Appendix

A: Proof of Lemma 2

For notational simplicity, we assume there are $T + 1$ observations ($t = 0, 1, \dots, T$) for this lemma. The differenced data have T observations so that x is $T \times N$. Let V_{NT} be the $r \times r$ diagonal matrix of the first r largest eigenvalues of $\frac{1}{TN}xx'$ in decreasing order. By the definition of eigenvectors and eigenvalues, we have $\frac{1}{TN}xx'\hat{f} = \hat{f}V_{NT}$ or $\frac{1}{NT}xx'\hat{f}V_{NT}^{-1} = \hat{f}$. Let $H = V_{NT}^{-1}(\hat{f}'f/T)(\Lambda'\Lambda/N)$ be a $r \times r$ matrix and $\delta_{NT} = \min\{\sqrt{N}, \sqrt{T}\}$. Then the following is a mathematical identity:

$$\hat{f}_t - Hf_t = V_{NT}^{-1} \left(\frac{1}{T} \sum_{s=1}^T \hat{f}_s \gamma_N(s, t) + \frac{1}{T} \sum_{s=1}^T \hat{f}_s \zeta_{st} + \frac{1}{T} \sum_{s=1}^T \hat{f}_s \eta_{st} + \frac{1}{T} \sum_{s=1}^T \hat{f}_s \xi_{st} \right), \quad (\text{A.1})$$

where for $z_t = (z_{1t}, z_{2t}, \dots, z_{Nt})'$,

$$\zeta_{st} = \frac{z_s' z_t}{N} - \gamma_N(s, t), \quad \eta_{st} = f_s' \Lambda' z_t / N, \quad \xi_{st} = f_t' \Lambda' z_s / N. \quad (\text{A.2})$$

Bai (2001a) showed that $\|V_{NT}^{-1}\| = O_p(1)$. Using $\hat{f}'\hat{f}/T = I_r$, together with Assumptions A and B, $\|H\| = O_p(1)$. Lemma 2 is implied by the following lemma:

Lemma A1 *Under Assumptions A-D, we have*

- (a). $T^{-3/2} \sup_{1 \leq k \leq T} \left\| \sum_{t=1}^k \sum_{s=1}^T \hat{f}_s \gamma_N(s, t) \right\| = O_p\left(\frac{1}{\sqrt{T}\delta_{NT}}\right) + O_p(T^{-3/4})$;
- (b). $T^{-3/2} \sup_{1 \leq k \leq T} \left\| \sum_{t=1}^k \sum_{s=1}^T \hat{f}_s \zeta_{st} \right\| = O_p\left(\frac{1}{\sqrt{N}}\right)$;
- (c). $T^{-3/2} \sup_{1 \leq k \leq T} \left\| \sum_{t=1}^k \sum_{s=1}^T \hat{f}_s \eta_{st} \right\| = O_p\left(\frac{1}{\sqrt{N}}\right)$;
- (d). $T^{-3/2} \sup_{1 \leq k \leq T} \left\| \sum_{t=1}^k \sum_{s=1}^T \hat{f}_s \xi_{st} \right\| = O_p\left(\frac{1}{\sqrt{N}\delta_{NT}}\right)$.

Proof: Consider part (a). By adding and subtracting terms,

$$\begin{aligned} \sum_{t=1}^k \sum_{s=1}^T \hat{f}_s \gamma_N(s, t) &= \sum_{t=1}^k \sum_{s=1}^T (\hat{f}_s - Hf_s + Hf_s) \gamma_N(s, t) \\ &= \sum_{s=1}^T (\hat{f}_s - Hf_s) \sum_{t=1}^k \gamma_N(s, t) + H \sum_{s=1}^T f_s \sum_{t=1}^k \gamma_N(s, t). \end{aligned}$$

Consider the first term:

$$\left\| \sum_{s=1}^T (\hat{f}_s - Hf_s) \sum_{t=1}^k \gamma_N(s, t) \right\| \leq \left(\sum_{s=1}^T \|\hat{f}_s - Hf_s\|^2 \right)^{1/2} \left(\sum_{s=1}^T \left| \sum_{t=1}^k \gamma_N(s, t) \right|^2 \right)^{1/2}$$

By Lemma 1(i), $\left(\sum_{s=1}^T \|\hat{f}_s - Hf_s\|^2 \right)^{1/2} = T^{1/2} O_p(\delta_{NT}^{-1})$. Because $|\sum_{t=1}^k \gamma_N(s, t)| \leq M$ for all k and s , $\left(\sum_{s=1}^T \left| \sum_{t=1}^k \gamma_N(s, t) \right|^2 \right)^{1/2} \leq M\sqrt{T}$. Thus $T^{-3/2} \left\| \sum_{s=1}^T (\hat{f}_s - Hf_s) \sum_{t=1}^k \gamma_N(s, t) \right\| = O_p\left(\frac{1}{\sqrt{T}\delta_{NT}}\right)$.

Consider the second term. We use the following fact: let X_1, X_2, \dots, X_T be an arbitrary sequence of random variables. If $\max_{1 \leq k \leq T} E|X_k|^\alpha \leq M$ ($\alpha > 0$), then $\max_{1 \leq k \leq T} |X_k| = O_p(T^{1/\alpha})$. Let $a_{sk} = \sum_{t=1}^k \gamma_N(s, t)$, then $E\|T^{-1/2} \sum_{s=1}^T f_s a_{sk}\|^4 \leq M$ by Assumptions A and C. This implies that (with $\alpha = 4$ and $X_k = T^{-1/2} \sum_{s=1}^T f_s a_{sk}$)

$$T^{-3/2} \sup_{1 \leq k \leq T} \left\| \sum_{s=1}^T f_s a_{sk} \right\| = O_p(T^{-3/4}),$$

proving (a). Consider part (b).

$$T^{-3/2} \sum_{t=1}^k \sum_{s=1}^T \widehat{f}_s \zeta_{st} = T^{-1} \sum_{s=1}^T (\widehat{f}_s - H f_s) \frac{1}{\sqrt{T}} \sum_{t=1}^k \zeta_{st} + HT^{-1} \sum_{s=1}^T f_s \frac{1}{\sqrt{T}} \sum_{t=1}^k \zeta_{st}.$$

For the first term,

$$\|T^{-1} \sum_{t=1}^k \sum_{s=1}^T (\widehat{f}_s - H f_s) \frac{1}{\sqrt{T}} \sum_{t=1}^k \zeta_{st}\| \leq \left(\frac{1}{T} \sum_{s=1}^T \|\widehat{f}_s - H f_s\|^2 \right)^{1/2} \left[\frac{1}{T} \sum_{s=1}^T \left(\frac{1}{\sqrt{T}} \sum_{t=1}^k \zeta_{st} \right)^2 \right]^{1/2}.$$

Furthermore,

$$\begin{aligned} \frac{1}{T} \sum_{s=1}^T \left(\frac{1}{\sqrt{T}} \sum_{t=1}^k \zeta_{st} \right)^2 &= \frac{1}{T} \sum_{s=1}^T \left[\frac{1}{\sqrt{T}} \sum_{t=1}^k \left(\frac{z'_s z_t}{N} - \gamma_N(s, t) \right) \right]^2 = \frac{1}{T} \sum_{s=1}^T \left[\frac{1}{\sqrt{T}} \sum_{t=1}^k \left(\frac{z'_s z_t}{N} - \frac{E(z'_s z_t)}{N} \right) \right]^2 \\ &= \frac{1}{N} \frac{1}{T} \sum_{s=1}^T \left[\frac{1}{\sqrt{NT}} \sum_{t=1}^k \sum_{i=1}^N (z_{is} z_{it} - E(z_{is} z_{it})) \right]^2 = O_p\left(\frac{1}{N}\right), \end{aligned}$$

uniformly in k by Assumption C5. Thus the first term is $O_p(\frac{1}{\delta_{NT}})O_p(\frac{1}{\sqrt{N}})$. Next,

$$T^{-3/2} \sum_{s=1}^T f_s \sum_{t=1}^k \zeta_{st} = \frac{1}{T\sqrt{N}} \sum_{s=1}^T f_s \frac{1}{\sqrt{TN}} \sum_{t=1}^k \sum_{i=1}^N (z_{is} z_{it} - E(z_{is} z_{it})) = \frac{1}{\sqrt{NT}} \sum_{s=1}^T f_s \phi_{k,s}$$

where $\phi_{k,s}$ is implicitly defined in the above expression. C5 implies that $E(\max_{1 \leq k \leq T} |\phi_{k,s}|) \leq M$. Thus $E(\max_{1 \leq k \leq T} \frac{1}{\sqrt{NT}} \|\sum_{s=1}^T f_s \phi_{k,s}\|) \leq \frac{1}{\sqrt{N}} \frac{1}{T} \sum_{s=1}^T E(\|f_s\| \max_{1 \leq k \leq T} |\phi_{k,s}|) = O(N^{-1/2})$ because $E(\|f_s\| \max_{1 \leq k \leq T} |\phi_{k,s}|) = E\|f_s\| \cdot E(\max_{1 \leq k \leq T} |\phi_{k,s}|) \leq M_1$ ($M_1 < \infty$) by the independence of f_s and the z'_{it} s. Thus, uniformly in k ,

$$T^{-3/2} \sum_{s=1}^T \widehat{f}_s \sum_{t=1}^k \zeta_{st} = O_p\left(\frac{1}{\delta_{NT}}\right) \cdot O_p\left(\frac{1}{\sqrt{N}}\right) + O_p\left(\frac{1}{\sqrt{N}}\right) = O_p\left(\frac{1}{\sqrt{N}}\right).$$

Consider part (c).

$$T^{-3/2} \sum_{s=1}^T \widehat{f}_s \eta_{st} = T^{-1} \sum_{s=1}^T (\widehat{f}_s - H f_s) \frac{1}{\sqrt{T}} \sum_{t=1}^k \eta_{st} + HT^{-1} \sum_{s=1}^T f_s \frac{1}{\sqrt{T}} \sum_{t=1}^k \eta_{st}.$$

But $T^{-1} \sum_{s=1}^T f_s \frac{1}{\sqrt{T}} \sum_{t=1}^k \eta_{st} = (\frac{1}{T} \sum_{s=1}^T f_s f_s') \frac{1}{N\sqrt{T}} \sum_{t=1}^k \sum_{i=1}^N \lambda_i e_{it} = O_p(\frac{1}{\sqrt{N}})$, uniformly in k by A(ii) and C4. Next,

$$\|T^{-1} \sum_{s=1}^T (\hat{f}_s - H f_s) \frac{1}{\sqrt{T}} \sum_{t=1}^k \eta_{st}\| \leq \left(\frac{1}{T} \sum_{s=1}^T \|\hat{f}_s - H f_s\|^2 \right)^{1/2} \cdot \left[\frac{1}{T} \sum_{s=1}^T \left(\frac{1}{\sqrt{T}} \sum_{t=1}^k \eta_{st}^2 \right) \right]^{1/2}.$$

The first expression is $O_p(1/\delta_{NT})$ by Lemma 1. For the second expression,

$$T^{-1} \sum_{s=1}^T \left(\frac{1}{\sqrt{T}} \sum_{t=1}^k \eta_{st} \right)^2 = \frac{1}{N} \frac{1}{T} \sum_{s=1}^T \left(f_s' \frac{1}{\sqrt{TN}} \sum_{t=1}^k \sum_{i=1}^N \lambda_i e_{it} \right)^2 = O_p\left(\frac{1}{N}\right),$$

uniformly in k . Thus, (c) is $O_p(\frac{1}{\sqrt{N}}) + O_p(\frac{1}{\sqrt{N}\delta_{NT}}) = O_p(\frac{1}{\sqrt{N}})$.

Finally for part (d),

$$T^{-3/2} \sum_{t=1}^k \sum_{s=1}^T \hat{f}_s \xi_{st} = T^{-3/2} \sum_{t=1}^k \sum_{s=1}^T \hat{f}_s f_t' \Lambda z_s / N = T^{-1} \sum_{s=1}^T (\hat{f}_s z_s' \Lambda / N) \frac{1}{\sqrt{T}} \sum_{t=1}^k f_t$$

It is proved in Bai (2001a) that $T^{-1} \sum_{s=1}^T (\hat{f}_s z_s' \Lambda / N) = O_p(\frac{1}{\sqrt{N}\delta_{NT}})$ (this can also be proved directly). By A(iii), $\frac{1}{\sqrt{T}} \sum_{t=1}^k f_t = O_p(1)$ uniformly in k . Thus (d) is equal to $O_p(\frac{1}{\sqrt{N}\delta_{NT}})$ uniformly in k . The proof of Lemma A1 is complete.

From equation (A.1) and Lemma A1

$$\max_{1 \leq k \leq T} \frac{1}{\sqrt{T}} \left\| \sum_{t=1}^k (\hat{f}_t - H f_t) \right\| = O_p\left(\frac{1}{\sqrt{T}\delta_{NT}}\right) + O_p\left(\frac{1}{T^{3/4}}\right) + O_p\left(\frac{1}{\sqrt{N}}\right) + O_p\left(\frac{1}{\sqrt{N}\delta_{NT}}\right) = O_p\left(\frac{1}{\sqrt{N}}\right) + O_p\left(\frac{1}{T^{3/4}}\right)$$

This proves Lemma 2.

B. Proof of Theorem 1

We first prove part 2 of Theorem 1. The objective is to show the weak convergence

$$DF_e^c(i) = \frac{\sum_{t=2}^T \hat{e}_{it-1} \Delta \hat{e}_{it}}{\left(s^2 \sum_{t=2}^T \hat{e}_{it-1}^2 \right)^{1/2}} \Rightarrow \frac{\frac{1}{2}(W_{ei}(1)^2 - 1)}{\left(\int_0^1 W_{ei}^2 dr \right)^{1/2}} \quad (\text{B.1})$$

where $s^2 = \frac{1}{T-1} \sum_{t=2}^T (\Delta \hat{e}_{it} - \hat{a}_i \hat{e}_{it-1})^2$ and \hat{a}_i is the OLS estimator in the regression of $\Delta \hat{e}_{it}$ on \hat{e}_{it-1} .

Because \hat{e}_{it} is defined as the cumulative sum of \hat{z}_{it} , i.e., $\hat{e}_{it} = \sum_{s=2}^t \hat{z}_{is}$, it follows that $\Delta \hat{e}_{it} = \hat{e}_{it} - \hat{e}_{it-1} = \hat{z}_{it}$. Noting

$$x_{it} = \lambda_i' f_t + z_{it} = \lambda_i' H^{-1} H f_t + z_{it},$$

$$x_{it} = \hat{\lambda}_i' \hat{f}_t + \hat{z}_{it},$$

and subtracting the first equation from the second, we obtain

$$\begin{aligned}\widehat{z}_{it} &= z_{it} + \lambda'_i H^{-1} H f_t - \widehat{\lambda}'_i \widehat{f}_t \\ &= z_{it} - \lambda'_i H^{-1} (\widehat{f}_t - H f_t) - (\widehat{\lambda}_i - H^{-1} \lambda_i)' \widehat{f}_t\end{aligned}$$

Because $\widehat{z}_{it} = \Delta \widehat{e}_{it}$ and $z_{it} = \Delta e_{it}$, we can rewrite the above as

$$\Delta \widehat{e}_{it} = \Delta e_{it} - \lambda'_i H^{-1} v_t - d'_i \widehat{f}_t \quad (\text{B.2})$$

where $v_t = \widehat{f}_t - H f_t$, and $d_i = \widehat{\lambda}_i - H^{-1} \lambda_i$. Note that $d_i = O_p(\max\{T^{-1/2}, N^{-1}\}) = o_p(1)$, as $N, T \rightarrow \infty$ by Lemma 1(iii). Lemma 2 implies

$$\sup_{2 \leq t \leq T} \left\| \frac{1}{\sqrt{T}} \sum_{s=2}^t v_s \right\| = o_p(1). \quad (\text{B.3})$$

Lemma B1 $\frac{1}{\sqrt{T}} \widehat{e}_{it} = \frac{1}{\sqrt{T}} e_{it} + o_p(1)$, where $o_p(1)$ is uniform in $t \in [2, T]$, as $N, T \rightarrow \infty$.

Proof: The cumulative sum of (B.2) leads to (define $\widehat{e}_{i1} = 0$)

$$\begin{aligned}\widehat{e}_{it} &= e_{it} - e_{i1} - \lambda'_i H^{-1} \sum_{s=2}^t v_s - d'_i \sum_{s=2}^t \widehat{f}_s, \text{ or} \\ \frac{\widehat{e}_{it}}{\sqrt{T}} &= \frac{e_{it}}{\sqrt{T}} - \frac{e_{i1}}{\sqrt{T}} - \lambda'_i H^{-1} \left(\frac{1}{\sqrt{T}} \sum_{s=2}^t v_s \right) - d'_i \frac{1}{\sqrt{T}} \sum_{s=2}^t \widehat{f}_s\end{aligned}$$

Now $\frac{e_{i1}}{\sqrt{T}} = O_p(\frac{1}{\sqrt{T}})$. The third term is $o_p(1)$ by (B.3). For the last term,

$$\begin{aligned}\|d'_i \frac{1}{\sqrt{T}} \sum_{s=2}^t \widehat{f}_s\| &\leq \|d_i\| \cdot \left\| \frac{1}{\sqrt{T}} \sum_{s=2}^t \widehat{f}_s \right\| \\ &\leq o_p(1) \left\| \frac{1}{\sqrt{T}} \sum_{s=2}^t (\widehat{f}_s - H f_s) \right\| + o_p(1) \left\| \frac{1}{\sqrt{T}} \sum_{s=2}^t f_s \right\| \cdot \|H\| \\ &= o_p(1) \left\| \frac{1}{\sqrt{T}} \sum_{s=2}^t v_s \right\| + o_p(1) O_p(1) \rightarrow 0\end{aligned}$$

because $\frac{1}{\sqrt{T}} \sum_{s=2}^t f_s = O_p(1)$ by Assumption A(iii) and $\left\| \frac{1}{\sqrt{T}} \sum_{s=2}^t v_s \right\| = o_p(1)$ by (B.3). (Note: the $o_p(1)$ in this lemma is in fact $O_p(\max\{T^{-1/2}, N^{-1/2}\})$.)

Lemma B2 Assume $N, T \rightarrow \infty$. Under the null hypothesis of $\rho_i = 1$,

- (i) $\frac{1}{T^2} \sum_{t=2}^T \widehat{e}_{it}^2 \Rightarrow \sigma_{ei}^2 \int_0^1 W_{ei}(r)^2 dr$.
 - (ii) $\frac{1}{T} \sum_{t=2}^T \widehat{e}_{it-1} \Delta \widehat{e}_{it} \Rightarrow \frac{\sigma_{ei}^2}{2} (W_{ei}(1)^2 - 1)$.
- where $\sigma_{ei}^2 = E(\epsilon_{it}^2)$.

Proof: (i) By Lemma B1, for $t = [Tr]$, $\frac{1}{\sqrt{T}}\widehat{e}_{it} = \frac{1}{\sqrt{T}}e_{it} + o_p(1) \Rightarrow \sigma_{ei}W_i(r)$. By the continuous mapping theorem, $\frac{1}{T^2}\sum_{t=2}^T\widehat{e}_{it}^2 \Rightarrow \sigma_{ei}^2\int_0^1W_i(r)^2dr$. Consider (ii). From the identity $\widehat{e}_{it}^2 = (\widehat{e}_{it-1} + \Delta\widehat{e}_{it})^2 = \widehat{e}_{it-1}^2 + (\Delta\widehat{e}_{it})^2 + 2\widehat{e}_{it-1}\Delta\widehat{e}_{it}$, we have

$$\frac{1}{T}\sum_{t=2}^T\widehat{e}_{it-1}\Delta\widehat{e}_{it} = \frac{\widehat{e}_{iT}^2}{2T} - \frac{\widehat{e}_{i1}^2}{2T} - \frac{1}{2T}\sum_{t=2}^T(\Delta\widehat{e}_{it})^2.$$

By Lemma B1, $\frac{1}{T}\widehat{e}_{iT}^2 \Rightarrow \sigma_{ei}^2W_i^2(1)$, and $\frac{1}{T}\widehat{e}_{i1}^2 = 0$ for $\widehat{e}_{i1} = 0$. From (B.2), $\Delta\widehat{e}_{it} = \Delta e_{it} - a_{it}$, where $a_{it} = \lambda'_i H^{-1}v_t + d'_i\Delta\widehat{F}_t$. Thus,

$$\frac{1}{T}\sum_{t=2}^T(\Delta\widehat{e}_{it})^2 = \frac{1}{T}\sum_{t=2}^T(\Delta e_{it})^2 - \frac{2}{T}\sum_{t=2}^T(\Delta e_{it})a_{it} + \frac{1}{T}\sum_{t=2}^T a_{it}^2.$$

Under the null hypothesis that $\rho_i = 1$, $\Delta e_{it} = \epsilon_{it}$, and $\frac{1}{T}\sum_{t=2}^T(\Delta e_{it})^2 \xrightarrow{p}\sigma_{ei}^2$. The middle term is $o_p(1)$ by the Cauchy-Schwartz inequality and $\frac{1}{T}\sum_{t=2}^T a_{it}^2 = o_p(1)$. The latter follows from

$$\begin{aligned} a_{it}^2 &\leq 2\|\lambda_i H^{-1}\|^2\|v_t\|^2 + 2\|d_i\|^2\|\widehat{f}_t\|^2 \\ \frac{1}{T}\sum_{t=2}^T a_{it}^2 &\leq 2\|\lambda_i H^{-1}\|^2 \cdot \frac{1}{T}\sum_{t=2}^T\|v_t\|^2 + 2\|d_i\|^2\frac{1}{T}\sum_{t=2}^T\|\widehat{f}_t\|^2 \\ &= O_p(1)O_p(\delta_{NT}^{-2}) + o_p(1)O_p(1) \rightarrow 0 \end{aligned}$$

by Lemma 1(i) and $\frac{1}{T}\sum_{t=2}^T\|\widehat{f}_t\|^2 = O_p(1)$. We obtained (ii) after combining results. Finally, (B.1) follows from $s^2 \xrightarrow{p}\sigma_{ei}^2$, Lemma B2, and the continuous mapping theorem.

We next consider part 1 of Theorem 1 (testing \widehat{F}_t with Demeaning). For simplicity, we consider the case of $k = 1$ and hence H is scalar. The objective is to show

$$DF_F^c = \frac{\sum_{t=2}^T(\widehat{F}_{t-1} - \overline{\widehat{F}})\Delta\widehat{F}_t}{\left(s^2\sum_{t=2}^T(\widehat{F}_{t-1} - \overline{\widehat{F}})^2\right)^{1/2}} \Rightarrow \frac{\int_0^1 W_u^c(r)dW_u(r)}{\left(\int_0^1 W_u^c(r)^2 dr\right)^{1/2}}. \quad (\text{B.4})$$

where $\overline{\widehat{F}} = \frac{1}{T-1}\sum_{t=2}^T\widehat{F}_t$, $s^2 = \frac{1}{T-1}\sum_{t=2}^T(\Delta\widehat{F}_t - \widehat{a}\widehat{F}_{t-1})^2$ with \widehat{a} being the OLS estimator when regressing $\Delta\widehat{F}_t$ on \widehat{F}_{t-1} , and $W_u^c(r) = W_u(r) - \int_0^1 W_u(r)dr$ is a demeaned Brownian motion.

Lemma B3 (i) $\frac{1}{\sqrt{T}}\widehat{F}_t = H\frac{1}{\sqrt{T}}F_t + o_p(1)$ and (ii) $\frac{1}{\sqrt{T}}\overline{\widehat{F}} = \frac{1}{\sqrt{T}}H\overline{F} + o_p(1)$ where $o_p(1)$ is uniform in $t \in [2, T]$ as $N, T \rightarrow \infty$.

Proof: Because $\widehat{F}_t = \sum_{s=2}^t \widehat{f}_s$, we have $\Delta\widehat{F}_t = \widehat{F}_t - \widehat{F}_{t-1} = \widehat{f}_t$. Thus $v_t = \widehat{f}_t - Hf_t = \Delta\widehat{F}_t - H\Delta F_t$, or $\Delta\widehat{F}_t = H\Delta F_t + v_t$. The cumulative sum implies (define $\widehat{F}_1 = 0$)

$$\widehat{F}_t = HF_t - HF_1 + \sum_{s=2}^t v_s, \text{ and} \quad (\text{B.5})$$

$$\frac{1}{\sqrt{T}}\widehat{F}_t = H\frac{1}{\sqrt{T}}F_t - H\frac{1}{\sqrt{T}}F_1 + \frac{1}{\sqrt{T}}\sum_{s=2}^t v_s.$$

The second on the right hand side is $O_p(T^{-1/2})$ and the third term is $o_p(1)$ uniformly in t by (B.3).

This proves (i). Averaging over (B.5), we obtain $\widehat{F} = H\bar{F} - HF_1 + \frac{1}{T-1}\sum_{t=2}^T\sum_{s=2}^t v_s$. Hence,

$$\frac{1}{\sqrt{T}}\widehat{F} = H\frac{1}{\sqrt{T}}\bar{F} - H\frac{F_1}{\sqrt{T}} + \frac{1}{(T-1)}\sum_{t=2}^T\left(\frac{1}{\sqrt{T}}\sum_{s=2}^t v_s\right).$$

The second term on the right is $O_p(T^{-1/2})$ and the last term is $o_p(1)$ because it is the average of $(\frac{1}{\sqrt{T}}\sum_{s=2}^t v_s)$ thus must be no larger than its maximum, which is $o_p(1)$ by (B.3).

Lemma B4 For $t = [Tr]$, as $N, T \rightarrow \infty$, under the null hypothesis that $\alpha = 1$,

(i) $\frac{1}{\sqrt{T}}\widehat{F}_t \Rightarrow H\sigma_u W_u(r)$, where $\sigma_u^2 = E(u_t^2) > 0$.

(ii) $\frac{1}{T^2}\sum_{t=2}^T(\widehat{F}_{t-1} - \widehat{F})^2 \Rightarrow H^2\sigma_u^2 \int_0^1 W_u^c(r)^2 dr$.

Proof: (i) follows from Lemma B3 and $T^{-1/2}F_t \Rightarrow \sigma_u W_u(r)$. For (ii), again by Lemma B3,

$$\frac{1}{\sqrt{T}}(\widehat{F}_t - \widehat{F}) = \frac{1}{\sqrt{T}}H(F_t - \bar{F}) + o_p(1) \Rightarrow H\sigma_u \left[W_u(r) - \int_0^1 W_u(r) dr \right] \equiv H\sigma_u W_u^c(r).$$

This together with the continuous mapping theorem leads to (ii).

Note that H depends on N and T . We simply denote its limit as H again. It is not necessary to explicitly study the limit of H because the test statistic in (B.4) can be rewritten in terms of $\{H^{-1}\widehat{F}_t\}$ (i.e., a scaling constant does not change the statistic). It is clear that $\frac{1}{\sqrt{T}}H^{-1}\widehat{F}_t \Rightarrow \sigma_u W_u(r)$. Other terms follow similarly. This remark also applies to the rest of the analysis.

Lemma B5 Under the null that $\alpha = 1$, $\frac{1}{T}\sum_{t=2}^T(\widehat{F}_{t-1} - \widehat{F})\Delta\widehat{F}_t \Rightarrow H^2\sigma_u^2 \int_0^1 W_u^c(r)dW_u(r)$.

Proof:

$$\frac{1}{T}\sum_{t=2}^T(\widehat{F}_{t-1} - \widehat{F})\Delta\widehat{F}_t = \frac{1}{T}\sum_{t=2}^T\widehat{F}_{t-1}\Delta\widehat{F}_t - \left(\frac{1}{\sqrt{T}}\widehat{F}\right)\left(\frac{1}{\sqrt{T}}\widehat{F}_T\right) \quad (\text{B.6})$$

(note $\sum_{t=2}^T\Delta\widehat{F}_t = \widehat{F}_T - \widehat{F}_1 = \widehat{F}_T$ because $\widehat{F}_1 = 0$). Consider the first term on the right:

$$\frac{1}{T}\sum_{t=2}^T\widehat{F}_{t-1}\Delta\widehat{F}_t = \frac{\widehat{F}_T^2}{2T} - \frac{\widehat{F}_1^2}{2T} - \frac{1}{2T}\sum_{t=2}^T(\Delta\widehat{F}_t)^2.$$

By Lemma B4(i), $\frac{1}{2}\frac{1}{T}\widehat{F}_T^2 \Rightarrow \frac{1}{2}H^2\sigma_u^2 W_u(1)^2$, and $\frac{1}{2T}\widehat{F}_1^2 = 0$. From $\Delta\widehat{F}_t = H\Delta F_t + v_t$,

$$\frac{1}{T}\sum_{t=2}^T(\Delta\widehat{F}_t)^2 = H^2\frac{1}{T}\sum_{t=2}^T(\Delta F_t)^2 + H\frac{1}{T}\sum_{t=2}^T(\Delta F_t)v_t + \frac{1}{T}\sum_{t=2}^T v_t^2 \xrightarrow{p} H^2\sigma_u^2 \quad (\text{B.7})$$

because $\frac{1}{T} \sum_{t=2}^T (\Delta F_t)^2 \xrightarrow{p} \sigma_u^2$ in view of $\Delta F_t = u_t$ when $\alpha = 1$. In addition, $\frac{1}{T} \sum_{t=2}^T v_t^2 = O_p(\delta_{NT}^{-2}) \rightarrow 0$ by Lemma 1(i), and the middle term is also $o_p(1)$ by the Cauchy-Schwarz inequality. Thus, the first term on right of (B.6) has a limit $H^2 \sigma_u^2 [W_u(1)^2 - 1]/2$.

Lemma B3 (ii) implies $\frac{1}{\sqrt{T}} \widehat{F} \Rightarrow H \sigma_u \int_0^1 W_u(r) dr$ because $T^{-3/2} \sum_{t=2}^T F_t \Rightarrow \sigma_u \int_0^1 W_u(r) dr$ when F_t is I(1). Lemma B4 (i) implies $\frac{1}{\sqrt{T}} \widehat{F}_T \Rightarrow H \sigma_u W_u(1)$. The second term of (B.6) has a limit $H^2 \sigma_u^2 W_u(1) \int_0^1 W_u(r) dr$. Summarizing results:

$$\frac{1}{T} \sum_{t=2}^T (\widehat{F}_{t-1} - \widehat{F}) \Delta \widehat{F}_t \Rightarrow H^2 \sigma_u^2 \left[\frac{1}{2} W_u(1)^2 - \frac{1}{2} - W_u(1) \int_0^1 W_u(r) dr \right] \equiv H^2 \sigma_u^2 \int_0^1 W_u^c(r) dW_u(r).$$

Finally, the weak convergence in (B.4) is implied by $s^2 \xrightarrow{p} H^2 \sigma_u^2$, Lemma B4(ii), Lemma B5, and the continuous mapping theorem.

C. Proof of Theorem 2

The model under differencing and then demeaning takes the form $x_{it} = \lambda'_i f_t + z_{it}$, where

$$\begin{aligned} x_{it} &= \Delta X_{it} - \overline{\Delta X_i} \\ f_t &= \Delta F_t - \overline{\Delta F} = \Delta F_t - \frac{F_T - F_1}{T-1}, \\ z_{it} &= \Delta e_{it} - \overline{\Delta e_i} = \Delta e_{it} - \frac{e_{iT} - e_{i1}}{T-1}. \end{aligned}$$

From $x_{it} = \widehat{\lambda}'_i \widehat{f}_t + \widehat{z}_{it}$, we have

$$\begin{aligned} \widehat{z}_{it} &= z_{it} + \lambda'_i f_t - \widehat{\lambda}'_i \widehat{f}_t \\ &= z_{it} + \lambda_i H^{-1} H f_t - \lambda'_i H^{-1} \widehat{f}_t + \lambda'_i H^{-1} \widehat{f}_t - \widehat{\lambda}'_i \widehat{f}_t \\ &= z_{it} - \lambda'_i H^{-1} (\widehat{f}_t - H f_t) - (\widehat{\lambda}_i - H^{-1'} \lambda_i)' \widehat{f}_t \\ &\equiv \Delta e_{it} - \overline{\Delta e_i} - \lambda'_i H^{-1} v_t - d'_i \widehat{f}_t \end{aligned}$$

where $v_t = \widehat{f}_t - H f_t$ and $d_i = \widehat{\lambda}_i - H^{-1'} \lambda_i$. Cumulative sum gives

$$\widehat{e}_{it} = \sum_{s=2}^t \widehat{z}_{is} = e_{it} - e_{i1} - \frac{e_{iT} - e_{i1}}{T-1} (t-1) - \lambda'_i H^{-1} \sum_{s=2}^t v_s - d_i \sum_{s=2}^t \widehat{f}_s \quad (\text{C.1})$$

By Lemma 3 in the main text,

$$\max_{2 \leq t \leq T} \frac{1}{\sqrt{T}} \left\| \sum_{s=2}^t v_s \right\| = o_p(1), \quad \frac{1}{T} \sum_{t=2}^T \|v_t\|^2 = O_p(\delta_{NT}^{-2}). \quad (\text{C.2})$$

Lemma C1 Under $\rho_i = 1$, for $t = [Tr]$, as $N, T \rightarrow \infty$, $\frac{1}{\sqrt{T}} \widehat{e}_{it} \Rightarrow \left[W_{ei}(r) - r W_{ei}(1) \right] \sigma_{ei} \equiv V_{ei}(r) \sigma_{ei}$.

Proof: From (C.1),

$$\frac{\widehat{e}_{it}}{\sqrt{T}} = \frac{e_{it}}{\sqrt{T}} - \left(\frac{e_{iT}}{\sqrt{T}}\right) \frac{(t-1)}{T-1} - \frac{e_{i1}}{\sqrt{T}} \left[1 - \frac{t-1}{T-1}\right] - \lambda'_i H^{-1} \left(\frac{1}{\sqrt{T}} \sum_{s=2}^t v_s\right) - \frac{d_i}{\sqrt{T}} \sum_{s=2}^t \widehat{f}_s.$$

By the invariance principle, $\frac{1}{\sqrt{T}} e_{it} \Rightarrow W_{ei}(r) \sigma_{ei}$, and $\frac{e_{iT}}{\sqrt{T}} \frac{t-1}{T-1} \Rightarrow r W_{ei}(1) \sigma_{ei}$. Furthermore, $\frac{e_{i1}}{\sqrt{T}} \rightarrow 0$. By (C.2), $\|\lambda'_i H^{-1} \frac{1}{\sqrt{T}} \sum_{s=2}^t v_s\| = o_p(1)$. Next $\|d'_i \frac{1}{\sqrt{T}} \sum_{s=2}^t \widehat{f}_s\| \leq \|d_i\| \cdot \|\frac{1}{\sqrt{T}} \sum_{s=2}^t \widehat{f}_t\|$, which is $o_p(1)$ uniformly in t . To see this, $\|d_i\| = o_p(1)$ by Lemma 1(iii). Also,

$$\begin{aligned} \frac{1}{\sqrt{T}} \sum_{s=2}^t \widehat{f}_s &= \frac{1}{\sqrt{T}} \sum_{s=2}^t (\widehat{f}_s - H f_s + H f_s) = \frac{1}{\sqrt{T}} \sum_{s=2}^t v_s + H \frac{1}{\sqrt{T}} \sum_{s=2}^t f_s \\ &= o_p(1) + \frac{1}{\sqrt{T}} \sum_{s=2}^t (\Delta F_s - \overline{\Delta F}) \\ &= o_p(1) + \frac{F_t - F_1}{\sqrt{T}} - \overline{\Delta F} \left(\frac{t-1}{\sqrt{T}}\right) \\ &= o_p(1) + O_p(1) - \left(\frac{F_T - F_1}{T-1}\right) \frac{t-1}{\sqrt{T}} = O_p(1) \end{aligned}$$

because $(F_T - F_1)/(T-1)^{1/2} = O_p(1)$ and $t/T = O(1)$ uniformly in t .

Lemma C2 Under $\rho_i = 1$, $\frac{1}{T} \sum_{t=2}^T (\Delta \widehat{e}_{it})^2 \xrightarrow{p} \sigma_{ei}^2$, as $N, T \rightarrow \infty$.

Proof: By definition, $\Delta \widehat{e}_{it} = \widehat{e}_{it} - \widehat{e}_{it-1} = \widehat{z}_{it}$. But

$$\widehat{z}_{it} = \Delta e_{it} - \overline{\Delta e_i} - \lambda'_i H^{-1} v_t - d'_i \widehat{f}_t \equiv \Delta e_{it} - a_{it},$$

where $a_{it} = \overline{\Delta e_i} + \lambda'_i H^{-1} v_t + d'_i \widehat{f}_t$. Thus,

$$\frac{1}{T} \sum_{t=2}^T (\Delta \widehat{e}_{it})^2 = \frac{1}{T} \sum_{t=2}^T (\Delta e_{it})^2 - \frac{2}{T} \sum_{t=2}^T (\Delta e_{it}) a_{it} + \frac{1}{T} \sum_{t=2}^T a_{it}^2. \quad (\text{C.3})$$

But $a_{it}^2 \leq 4(\overline{\Delta e_i})^2 + 4\|\lambda_i H^{-1}\|^2 \cdot \|v_t\|^2 + 4\|d'_i \widehat{f}_t\|^2$. Thus,

$$\begin{aligned} \frac{1}{T} \sum_{t=2}^T a_{it}^2 &\leq 4(\overline{\Delta e_i})^2 + 4\|\lambda_i H^{-1}\|^2 \frac{1}{T} \sum_{t=2}^T \|v_t\|^2 + 4\|d_i\|^2 \frac{1}{T} \sum_{t=2}^T \|\widehat{f}_t\|^2 \\ &= 4 \left(\frac{e_{iT} - e_{i1}}{T-1}\right)^2 + O_p(\delta_{NT}^{-2}) + o_p(1) \frac{1}{T} \sum_{t=2}^T \|\widehat{f}_t\|^2 \\ &= O_p\left(\frac{1}{T}\right) + O_p(\delta_{NT}^{-2}) + o_p(1) O_p(1) \rightarrow 0. \end{aligned}$$

since $\frac{1}{T} \sum_{t=2}^T \|\widehat{f}_t\|^2 = O_p(1)$. The second term of (C.3) is $o_p(1)$ by the Cauchy-Schwarz inequality, and the first term converges in probability to σ_{ei}^2 , proving Lemma C2.

The DF statistic is

$$DF_e^\tau(i) = \frac{T^{-1} \sum_{t=2}^T \widehat{e}_{it-1} \Delta \widehat{e}_{it}}{\left(s^2 T^{-2} \sum_{t=2}^T \widehat{e}_{it-1}^2 \right)^{1/2}},$$

where $s^2 = T^{-1} \sum_{t=2}^T (\Delta \widehat{e}_{it} - \widehat{a}_i \widehat{e}_{it-1})^2$, which can be shown to converge to σ_{ei}^2 . For the numerator,

$$\frac{1}{T} \sum_{t=2}^T \widehat{e}_{it-1} \Delta \widehat{e}_{it} = \frac{1}{2T} (\widehat{e}_{iT})^2 - \frac{1}{2T} \widehat{e}_{i1}^2 - \frac{1}{2T} \sum_{t=2}^T (\Delta \widehat{e}_{it})^2$$

Now $\frac{1}{T} \widehat{e}_{iT}^2 \Rightarrow \sigma_{ei}^2 V_{ei}^2(1) = 0$ by Lemma C1, and $\frac{1}{T} \widehat{e}_{i1}^2 = 0$. Together with Lemma C2,

$$\frac{1}{T} \sum_{t=2}^T \widehat{e}_{it-1} \Delta \widehat{e}_{it} \Rightarrow \frac{-\sigma_{ei}^2}{2}.$$

The denominator $T^{-1} \sum_{t=2}^T \widehat{e}_{it-1}^2 \Rightarrow \sigma_{ei}^2 \int_0^1 V_{i\epsilon}(r)^2 dr$ by Lemma C1. Collecting results,

$$DF_e^\tau(i) \Rightarrow -\frac{1}{2} \left(\int_0^1 V_{i\epsilon}(r) \right)^{-1/2}.$$

We next consider part 1 of Theorem 2 (Testing \widehat{F}_t with demeaning and detrending). For simplicity, we consider the case of $k = 1$ so that H is a scalar. From $\widehat{f}_t = H f_t + v_t$,

$$\begin{aligned} \widehat{F}_t &= \sum_{s=2}^t \widehat{f}_s = H \sum_{s=2}^t f_s + \sum_{s=2}^t v_s = H \sum_{s=2}^t (\Delta F_s - \overline{\Delta F}) + \sum_{s=2}^t v_s \\ &= H \left[F_t - F_1 - \frac{F_T - F_1}{T-1} (t-1) \right] + V_t, \end{aligned} \quad (\text{C.4})$$

where $V_t = \sum_{s=2}^t v_s$. For a sequence y_t , let y_t^τ denote the residual from regressing $\{y_t\}$ on $[1, t]$ ($t = 2, \dots, T$). That is, y_t^τ is the demeaned and detrended series. Then from (C.4),

$$\widehat{F}_t^\tau = H F_t^\tau + V_t^\tau \quad (\text{C.5})$$

[clearly, demeaning and detrending will remove $F_1 + \frac{F_T - F_1}{T-1} (t-1)$]. The DF statistic for the series \widehat{F}_t with demeaning and detrending is numerically equal to (see, e.g., Hayashi (2000), page 608)

$$DF_F^\tau = \frac{T^{-1} \sum_{t=2}^T \widehat{F}_{t-1}^\tau \Delta \widehat{F}_t}{\sqrt{s^2 T^{-2} \sum_{t=2}^T (\widehat{F}_{t-1}^\tau)^2}}, \quad (\text{C.6})$$

where $s^2 = \frac{1}{T-2} \sum_{t=2}^T (\Delta \widehat{F}_t - \widehat{a} - \widehat{b}t - \widehat{c}\widehat{F}_{t-1})^2$ with $(\widehat{a}, \widehat{b}, \widehat{c})$ being the OLS estimates when regressing $\Delta \widehat{F}_t$ on $[1, t, \widehat{F}_{t-1}]$. It is easy to show that $s^2 \xrightarrow{p} H^2 \sigma_u^2$.

Lemma C3 $\max_{2 \leq t \leq T} \frac{1}{\sqrt{T}} |V_t^\tau| = o_p(1)$.

Proof: This simply follows from $\frac{1}{\sqrt{T}}|V_t| = \frac{1}{\sqrt{T}}|\sum_{s=2}^t v_s| = o_p(1)$ uniformly in t . To see this, let $V = (V_2, V_3, \dots, V_T)'$, $V^\tau = (V_2^\tau, V_3^\tau, \dots, V_T^\tau)'$, and let $Z = (Z_2, Z_3, \dots, Z_T)'$, where $Z_t = (1, t)$. Then $V^\tau = M_z V$, where $M_z = I - Z(Z'Z)^{-1}Z'$. Then, it is easy to show that V_t^τ can be written as

$$V_t^\tau = V_t + a_T \frac{1}{T} \sum_{j=2}^T V_j + b_T(t/T) \frac{1}{T^2} \sum_{j=2}^T j V_j$$

where a_T and b_T are bounded numbers. It follows that

$$\frac{1}{\sqrt{T}} V_t^\tau = \frac{1}{\sqrt{T}} V_t + a_T \frac{1}{T^{3/2}} \sum_{j=2}^T V_j + b_T(t/T) \frac{1}{T^{5/2}} \sum_{j=2}^T j V_j.$$

However, $|\frac{1}{T^{3/2}} \sum_{j=2}^T V_j| \leq \max_t \frac{1}{\sqrt{T}} |V_t| = o_p(1)$, by equation (C.2). Similarly, $\frac{1}{T^{5/2}} |\sum_{j=2}^T j V_j| \leq \max_t \frac{1}{\sqrt{T}} |V_t| = o_p(1)$, proving the lemma.

Lemma C4 As $N, T \rightarrow \infty$,

$$(i) \frac{1}{\sqrt{T}} \widehat{F}_t^\tau = H \frac{1}{\sqrt{T}} F_t^\tau + o_p(1) \Rightarrow H \sigma_u W_u^\tau(r)$$

$$(ii) T^{-2} \sum_{t=2}^T (\widehat{F}_{t-1}^\tau)^2 = H^2 T^{-2} \sum_{t=2}^T (F_{t-1}^\tau)^2 + o_p(1) \Rightarrow H^2 \sigma_u^2 \int_0^1 W_u^\tau(r)^2 dr.$$

Proof: (i) follows from (C.5), Lemma C3, and $\frac{1}{\sqrt{T}} F_t^\tau \Rightarrow \sigma_u W_u^\tau(r)$. The equality in (ii) follows from (C.5) and Lemma C3. The weak convergence in (ii) follows from the continuous mapping theorem and (i).

Lemma C5 $\frac{1}{T} \sum_{t=2}^T \widehat{F}_{t-1}^\tau \Delta \widehat{F}_t \Rightarrow H^2 \sigma_u^2 \int_0^1 W_u^\tau(r) dW_u(r)$

Proof: First note that

$$\frac{1}{T} \sum_{t=2}^T \widehat{F}_{t-1}^\tau \Delta \widehat{F}_t = \frac{1}{T} \sum_{t=2}^T \widehat{F}_{t-1}^\tau \Delta \widehat{F}_t^\tau \quad (C.7)$$

where $\Delta \widehat{F}_t^\tau = \widehat{F}_t^\tau - \widehat{F}_{t-1}^\tau$. To see this, write $\widehat{F}_t^\tau = \widehat{F}_t - \widehat{a} - \widehat{b}t$ for some \widehat{a} and \widehat{b} . This is possible because \widehat{F}_t^τ is the projection residual of \widehat{F}_t . Thus, $\Delta \widehat{F}_t^\tau = \Delta \widehat{F}_t - \widehat{b}$. Equation (C.7) is proved by noting $\sum_{t=2}^T \widehat{F}_{t-1}^\tau = 0$ (normal equation). Next use the identity

$$\frac{1}{T} \sum_{t=2}^T \widehat{F}_{t-1}^\tau \Delta \widehat{F}_t^\tau = \frac{1}{2T} (\widehat{F}_T^\tau)^2 - \frac{1}{2T} (\widehat{F}_1^\tau)^2 - \frac{1}{2T} \sum_{t=2}^T (\Delta \widehat{F}_t^\tau)^2 \quad (C.8)$$

By Lemma C4, $\frac{1}{2T} (\widehat{F}_T^\tau)^2 \Rightarrow \frac{1}{2} H^2 \sigma_u^2 W_u^\tau(1)^2$ and $\frac{1}{2T} (\widehat{F}_1^\tau)^2 \Rightarrow \frac{1}{2} H^2 \sigma_u^2 W_u^\tau(0)^2$. Next, from $\Delta \widehat{F}_t^\tau = \Delta \widehat{F}_t - \widehat{b}$ and $\widehat{b} = o_p(1)$ (the slope coefficient has this property), we have $\frac{1}{T} \sum_{t=2}^T (\Delta \widehat{F}_t^\tau)^2 = \frac{1}{T} \sum_{t=2}^T (\Delta \widehat{F}_t)^2 + o_p(1) \xrightarrow{p} H^2 \sigma_u^2$, see (B.7) for an identical proof. In summary,

$$\frac{1}{T} \sum_{t=2}^T \widehat{F}_{t-1}^\tau \Delta \widehat{F}_t \Rightarrow \frac{1}{2} H^2 \sigma_u^2 [W_u^\tau(1)^2 - W_u^\tau(0)^2 - 1] = H^2 \sigma_u^2 \int_0^1 W_u^\tau(r) dW_u(r).$$

This proves Lemma C5. Finally, the limiting distribution of (C.6), as stated in Theorem 2, is a consequence of Lemma C4, Lemma C5, and the continuous mapping theorem.

References

- Bai, J. and Ng, S. (2002), Determining the Number of Factors in Approximate Factor Models, *forthcoming in Econometrica*.
- Bai, J. S. (2001), Inference on Factor Models of Large Dimensions, mimeo, Boston College.
- Baltagi, B. and Kao, C. (2001), Nonstationary Panels, Cointegration in Panels, and Dynamic Panels, Center for Policy Research, Syracuse University.
- Banerjee, A., Marcellino, M. and Osbat, C. (2001), Testing for PPP: Should we use Panel Methods?, mimeo, European University Institute.
- Bryan, M. and Cecchetti, S. G. (1993), The Consumer Price Index as a Measure of Inflation, *Economic Review of the Federal Reserve Bank of Cleveland* **29**, 15–24.
- Chamberlain, G. and Rothschild, M. (1983), Arbitrage, Factor Structure and Mean-Variance Analysis in Large Asset Markets, *Econometrica* **51**, 1305–1324.
- Choi, I. (2001), Unit Root Tests for Panel Data, *Journal of International Money and Finance* **20**, 249–272.
- Clark, T. (2001), Comparing Measures of Core Inflation, *Federal Reserve Bank of Kansas City Economic Review* **2**, 5–31.
- Dickey, D. A. and Fuller, W. A. (1979), Distribution of the Estimators for Autoregressive Time Series with a Unit Root, *Journal of the American Statistical Association* **74**, 427–431.
- Elliott, G., Rothenberg, T. J. and Stock, J. H. (1996), Efficient Tests for an Autoregressive Unit Root, *Econometrica* **64**, 813–836.
- Engle, R. F. and Kozicki, S. (1993), Testing for Common Features, *Journal of Business and Economic Statistics* **11**, 369–379.
- Forni, M., Hallin, M., Lippi, M. and Reichlin, L. (2000), The Generalized Dynamic Factor Model: Identification and Estimation, *Review of Economics and Statistics* **82:4**, 540–554.
- Fuller, W. A. (1976), *Introduction to Statistical Time Series*, John Wiley, New York.
- Gonzalo, J. and Granger, C. (1995), Estimation of Common Long-Memory Components in Cointegrated Systems, *Journal of Business and Economic Statistics* **13**, 27–35.
- Hayashi, F. (2000), *Econometrics*, Princeton University Press, Princeton, N.J.
- Im, K., Pesaran, M. and Shin, Y. (1997), Testing for Unit Roots in Heterogeneous Panels, mimeo, Cambridge University.
- Ireland, P. (1999), Does the Time-Consistency Problem Explain the Behavior of Inflation in the United States, *Journal of Monetary Economics* **44:2**, 279–293.
- King, R. G., Plosser, C., Stock, J. and Watson, M. (1991), Stochastic Trends and Economic Fluctuations, *American Economic Review* **81:4**, 819–40.
- Levin, A. and Lin, C. F. (1993), Unit Root Tests in Panel Data: Asymptotic and Finite Sample Properties. UCSD Discussion Paper 93-56.

- MacKinnon, J. (1994), Approximate Asymptotic Distribution Functions for Unit-Root and Cointegration Tests, *Journal of Business and Economic Statistics* **12**, 167–177.
- Maddala, G. S. and Wu, S. (1999), A Comparative Study of Unit Root Tests with Panel data and a New Simple Test, *Oxford Bulletin of Economics and Statistics* **Special Issue**, 631–652.
- Ng, S. and Perron, P. (1997), Estimation and Inference in Nearly Unbalanced and Nearly Cointegrated Systems, *Journal of Econometrics* **79**, 54–81.
- Ng, S. and Perron, P. (2001), Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power, *Econometrica* **69:6**, 1519–1554.
- Nyblom, J. and Harvey, A. (2000), Tests of Common Stochastic Trends, *Econometric Theory* **16:2**, 176–199.
- O’Connell, P. (1998), The Overvaluation of Purchasing Power Parity, *Journal of International Economics* **44**, 1–19.
- Pantula, S. G. (1991), Asymptotic Distributions of Unit Root Tests When the Process is Nearly Stationary, *Journal of Business and Economic Statistics* **9**, 63–71.
- Pedroni, P. (1995), Panel Cointegration: Asymptotic and Finite Sample Properties of Pooled Time Series Tests with an Application to the PPP Hypothesis: New Results, Indiana University Working Paper 95-013.
- Pesaran, H. and Smith, R. (1995), Estimating Long Run Relationships from Dynamic Heterogeneous Panels, *Journal of Econometrics* **68**, 79–113.
- Phillips, P. and Moon, R. (1999), Linear Regression Limit Theory for Nonstationary Panel Data, *Econometrica* **67**, 1057–1111.
- Phillips, P. C. B. and Ouliaris, S. (1990), Asymptotic Properties of Residual Based Tests for Cointegration, *Econometrica* **58**, 165–193.
- Quah, D. (1994), Exploiting Cross-Section Variations for Unit Root Inference in Dynamic Panels, *Economics Letters* **44**, 1–9.
- Quah, D. and Vahey, S. P. (1995), Measuring Core Inflation, *Economic Journal* **105**, 1130–1144.
- Said, S. E. and Dickey, D. A. (1984), Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order, *Biometrika* **71**, 599–607.
- Schmidt, P. and Lee, J. (1991), A Modification of Schmidt-Phillips Unit Root Test, *Economics Letters* **36**, 285–289.
- Schmidt, P. and Phillips, P. (1992), LM Tests for a Unit Root in the Presence of Deterministic Trends, *Oxford Bulletin of Economics and Statistics* **54**, 257–287.
- Schwert, G. W. (1989), Tests for Unit Roots: A Monte Carlo Investigation, *Journal of Business and Economic Statistics* **7**, 147–160.
- Stock, J. H. and Watson, M. W. (1988), Testing for Common Trends, *Journal of the American Statistical Association* **83**, 1097–1107.
- Stock, J. H. and Watson, M. W. (1998), Diffusion Indexes, NBER Working Paper 6702.

Stock, J. H. and Watson, M. W. (1999), Forecasting Inflation, *Journal of Monetary Economics* **44**, 293–335.

Wynne, M. (1999), Core Inflation: A Review of Some Conceptual Issues, Federal Reserve of Dallas WP 9903.

Table 1a: Rejection rates for the null hypothesis of a unit root, $p = 0, k = 1$

				$\sigma^F = \sqrt{10}$				$\sigma^F = 1$				$\sigma^F = \sqrt{5}$			
T	N	ρ	α	X	\widehat{F}	\widehat{e}	\widehat{e}^1	X	\widehat{F}	\widehat{e}	e^1	X	\widehat{F}	\widehat{e}	\widehat{e}^1
100	20	1.00	0.00	0.28	0.96	0.05	0.02	0.08	0.52	0.06	0.02	0.07	0.29	0.06	0.02
100	20	1.00	0.50	0.37	0.93	0.05	0.02	0.11	0.66	0.06	0.02	0.09	0.47	0.06	0.02
100	20	1.00	0.80	0.30	0.56	0.05	0.02	0.13	0.49	0.06	0.03	0.10	0.40	0.05	0.03
100	20	1.00	0.90	0.16	0.24	0.06	0.03	0.10	0.24	0.06	0.03	0.09	0.21	0.05	0.03
100	20	1.00	0.95	0.11	0.15	0.06	0.04	0.08	0.13	0.05	0.04	0.08	0.12	0.06	0.04
100	20	0.00	1.00	0.10	0.08	0.44	0.98	0.16	0.07	0.43	0.98	0.19	0.06	0.43	0.98
100	20	0.50	1.00	0.10	0.06	0.58	0.86	0.18	0.06	0.58	0.86	0.23	0.07	0.58	0.86
100	20	0.80	1.00	0.08	0.05	0.59	0.41	0.17	0.07	0.58	0.41	0.21	0.07	0.57	0.42
100	20	0.90	1.00	0.08	0.06	0.43	0.17	0.12	0.06	0.43	0.18	0.14	0.06	0.44	0.18
100	20	0.95	1.00	0.07	0.06	0.25	0.09	0.09	0.07	0.25	0.09	0.10	0.07	0.25	0.09
100	20	1.00	1.00	0.06	0.05	0.06	0.05	0.06	0.07	0.06	0.05	0.07	0.07	0.06	0.05
100	20	0.50	0.80	0.66	0.60	0.68	0.83	0.75	0.62	0.67	0.83	0.79	0.59	0.67	0.83
100	20	0.80	0.50	0.85	0.94	0.64	0.31	0.73	0.94	0.64	0.32	0.69	0.93	0.64	0.32
100	20	0.00	0.90	0.33	0.28	0.57	0.98	0.42	0.28	0.57	0.98	0.47	0.27	0.57	0.98
100	20	0.90	0.00	0.66	1.00	0.46	0.10	0.35	0.93	0.47	0.10	0.31	0.82	0.46	0.11
100	100	1.00	0.00	0.28	0.99	0.06	0.02	0.08	0.90	0.06	0.02	0.07	0.77	0.05	0.02
100	100	1.00	0.50	0.36	0.96	0.05	0.02	0.11	0.90	0.05	0.02	0.09	0.84	0.06	0.02
100	100	1.00	0.80	0.29	0.57	0.05	0.02	0.13	0.54	0.06	0.02	0.10	0.53	0.06	0.02
100	100	1.00	0.90	0.18	0.26	0.05	0.03	0.11	0.26	0.06	0.03	0.09	0.26	0.06	0.03
100	100	1.00	0.95	0.10	0.13	0.05	0.04	0.09	0.13	0.06	0.04	0.08	0.14	0.06	0.04
100	100	0.00	1.00	0.10	0.07	0.43	0.98	0.15	0.05	0.44	0.98	0.19	0.06	0.44	0.98
100	100	0.50	1.00	0.10	0.07	0.58	0.86	0.17	0.05	0.58	0.86	0.22	0.06	0.58	0.86
100	100	0.80	1.00	0.09	0.06	0.59	0.41	0.16	0.07	0.58	0.42	0.20	0.06	0.59	0.42
100	100	0.90	1.00	0.08	0.06	0.43	0.18	0.12	0.06	0.43	0.18	0.15	0.07	0.43	0.18
100	100	0.95	1.00	0.08	0.07	0.25	0.09	0.08	0.05	0.25	0.09	0.09	0.08	0.25	0.09
100	100	1.00	1.00	0.07	0.07	0.06	0.05	0.07	0.07	0.06	0.05	0.07	0.06	0.06	0.05
100	100	0.50	0.80	0.66	0.60	0.67	0.84	0.75	0.60	0.67	0.83	0.80	0.62	0.68	0.83
100	100	0.80	0.50	0.86	0.95	0.64	0.32	0.73	0.95	0.64	0.32	0.69	0.95	0.64	0.32
100	100	0.00	0.90	0.33	0.28	0.57	0.98	0.44	0.29	0.57	0.98	0.45	0.24	0.57	0.98
100	100	0.90	0.00	0.65	0.99	0.47	0.10	0.35	0.99	0.46	0.10	0.31	0.98	0.46	0.10

Table 1b: Rejection rates for the null hypothesis of a unit root, $p = 1, k = 1$

				$\sigma^F = \sqrt{10}$				$\sigma^F = 1$				$\sigma^F = \sqrt{5}$			
T	N	ρ	α	X	\widehat{F}	\widehat{e}	\widehat{e}^1	X	\widehat{F}	\widehat{e}	e^1	X	\widehat{F}	\widehat{e}	\widehat{e}^1
100	20	1.00	0.00	0.33	0.95	0.05	0.02	0.10	0.63	0.06	0.03	0.08	0.41	0.05	0.03
100	20	1.00	0.50	0.40	0.82	0.05	0.02	0.14	0.66	0.05	0.03	0.10	0.51	0.06	0.03
100	20	1.00	0.80	0.25	0.38	0.05	0.03	0.14	0.39	0.05	0.03	0.12	0.33	0.06	0.04
100	20	1.00	0.90	0.13	0.15	0.05	0.04	0.10	0.17	0.05	0.04	0.09	0.14	0.05	0.04
100	20	1.00	0.95	0.10	0.11	0.05	0.05	0.08	0.10	0.05	0.05	0.08	0.10	0.05	0.05
100	20	0.00	1.00	0.10	0.07	0.35	0.94	0.17	0.07	0.34	0.94	0.20	0.06	0.34	0.94
100	20	0.50	1.00	0.11	0.07	0.48	0.75	0.19	0.06	0.48	0.74	0.25	0.07	0.48	0.74
100	20	0.80	1.00	0.10	0.07	0.36	0.30	0.15	0.07	0.36	0.30	0.18	0.08	0.37	0.30
100	20	0.90	1.00	0.08	0.07	0.20	0.14	0.11	0.08	0.20	0.13	0.12	0.07	0.20	0.13
100	20	0.95	1.00	0.08	0.08	0.10	0.08	0.09	0.09	0.10	0.08	0.09	0.07	0.10	0.08
100	20	1.00	1.00	0.08	0.08	0.05	0.05	0.07	0.07	0.05	0.05	0.07	0.08	0.05	0.05
100	20	0.50	0.80	0.47	0.41	0.53	0.71	0.56	0.41	0.52	0.72	0.59	0.39	0.53	0.71
100	20	0.80	0.50	0.69	0.82	0.38	0.23	0.53	0.80	0.38	0.23	0.48	0.79	0.38	0.23
100	20	0.00	0.90	0.22	0.17	0.42	0.94	0.32	0.18	0.41	0.94	0.36	0.19	0.40	0.94
100	20	0.90	0.00	0.55	0.97	0.21	0.08	0.23	0.82	0.20	0.09	0.20	0.68	0.20	0.08
100	100	1.00	0.00	0.33	0.97	0.05	0.02	0.10	0.91	0.05	0.02	0.08	0.85	0.05	0.02
100	100	1.00	0.50	0.39	0.84	0.05	0.02	0.14	0.82	0.05	0.03	0.10	0.76	0.05	0.03
100	100	1.00	0.80	0.25	0.39	0.05	0.03	0.14	0.40	0.05	0.03	0.11	0.36	0.05	0.03
100	100	1.00	0.90	0.14	0.17	0.05	0.04	0.11	0.18	0.05	0.04	0.09	0.17	0.05	0.04
100	100	1.00	0.95	0.09	0.10	0.05	0.05	0.09	0.11	0.05	0.05	0.08	0.10	0.05	0.05
100	100	0.00	1.00	0.10	0.07	0.35	0.94	0.18	0.08	0.35	0.94	0.21	0.06	0.35	0.94
100	100	0.50	1.00	0.11	0.07	0.48	0.74	0.19	0.06	0.48	0.74	0.24	0.06	0.47	0.74
100	100	0.80	1.00	0.10	0.07	0.37	0.30	0.15	0.07	0.36	0.30	0.18	0.07	0.37	0.31
100	100	0.90	1.00	0.08	0.07	0.20	0.13	0.11	0.07	0.20	0.13	0.12	0.08	0.20	0.13
100	100	0.95	1.00	0.08	0.06	0.10	0.08	0.09	0.08	0.10	0.08	0.09	0.07	0.10	0.08
100	100	1.00	1.00	0.07	0.08	0.05	0.05	0.07	0.06	0.05	0.05	0.07	0.06	0.05	0.05
100	100	0.50	0.80	0.46	0.40	0.53	0.71	0.56	0.40	0.53	0.71	0.61	0.42	0.53	0.71
100	100	0.80	0.50	0.70	0.83	0.38	0.23	0.53	0.83	0.38	0.23	0.49	0.83	0.39	0.23
100	100	0.00	0.90	0.22	0.17	0.41	0.94	0.32	0.19	0.41	0.94	0.36	0.17	0.41	0.94
100	100	0.90	0.00	0.54	0.96	0.20	0.08	0.24	0.95	0.20	0.08	0.20	0.93	0.20	0.08

Table 2a: Pooled Tests: Rejection rates for the null hypothesis of a unit root , $p = 0, k = 1$

				$\sigma^F = \sqrt{10}$			$\sigma^F = 1$			$\sigma^F = \sqrt{.5}$		
T	N	ρ	α	X	\hat{e}	\hat{e}^\perp	X	\hat{e}	\hat{e}^\perp	X	\hat{e}	\hat{e}^\perp
100	20	1.00	0.00	0.96	0.06	0.00	0.25	0.07	0.01	0.16	0.07	0.02
100	20	1.00	0.50	0.97	0.07	0.00	0.46	0.07	0.02	0.29	0.06	0.03
100	20	1.00	0.80	0.90	0.06	0.01	0.58	0.07	0.03	0.42	0.06	0.04
100	20	1.00	0.90	0.69	0.07	0.04	0.45	0.07	0.04	0.34	0.07	0.04
100	20	1.00	0.95	0.49	0.09	0.08	0.35	0.07	0.06	0.30	0.07	0.09
100	20	0.00	1.00	0.36	1.00	1.00	0.51	0.99	1.00	0.60	0.99	1.00
100	20	0.50	1.00	0.35	1.00	1.00	0.56	1.00	1.00	0.69	1.00	1.00
100	20	0.80	1.00	0.36	1.00	1.00	0.59	1.00	1.00	0.67	1.00	1.00
100	20	0.90	1.00	0.36	1.00	0.96	0.48	1.00	0.96	0.61	1.00	0.96
100	20	0.95	1.00	0.31	1.00	0.50	0.40	1.00	0.51	0.47	1.00	0.50
100	20	1.00	1.00	0.27	0.07	0.15	0.24	0.07	0.15	0.19	0.07	0.16
100	20	0.50	0.80	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
100	20	0.80	0.50	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
100	20	0.00	0.90	0.83	1.00	1.00	0.94	1.00	1.00	0.97	1.00	1.00
100	20	0.90	0.00	1.00	1.00	0.68	1.00	1.00	0.72	1.00	1.00	0.70
100	100	1.00	0.00	1.00	0.07	0.00	0.57	0.06	0.00	0.35	0.07	0.00
100	100	1.00	0.50	1.00	0.06	0.00	0.83	0.07	0.00	0.63	0.07	0.00
100	100	1.00	0.80	0.96	0.05	0.01	0.88	0.07	0.01	0.77	0.07	0.02
100	100	1.00	0.90	0.82	0.05	0.05	0.75	0.06	0.07	0.67	0.07	0.09
100	100	1.00	0.95	0.57	0.06	0.13	0.59	0.08	0.13	0.54	0.06	0.13
100	100	0.00	1.00	0.41	1.00	1.00	0.60	1.00	1.00	0.69	1.00	1.00
100	100	0.50	1.00	0.41	1.00	1.00	0.66	1.00	1.00	0.75	1.00	1.00
100	100	0.80	1.00	0.41	1.00	1.00	0.64	1.00	1.00	0.77	1.00	1.00
100	100	0.90	1.00	0.40	1.00	1.00	0.59	1.00	1.00	0.73	1.00	1.00
100	100	0.95	1.00	0.38	1.00	0.86	0.50	1.00	0.86	0.59	1.00	0.85
100	100	1.00	1.00	0.36	0.06	0.30	0.39	0.05	0.30	0.36	0.07	0.30
100	100	0.50	0.80	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
100	100	0.80	0.50	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
100	100	0.00	0.90	0.89	1.00	1.00	0.98	1.00	1.00	1.00	1.00	1.00
100	100	0.90	0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table 2b: Pooled Tests: Rejection rates for the null hypothesis of a unit root , $p = 1, k = 1$

				$\sigma^F = \sqrt{10}$			$\sigma^F = 1$			$\sigma^F = \sqrt{.5}$		
T	N	ρ	α	X	\hat{e}	\hat{e}^\perp	X	\hat{e}	\hat{e}^\perp	X	\hat{e}	\hat{e}^\perp
100	20	1.00	0.00	0.97	0.07	0.00	0.36	0.09	0.00	0.26	0.08	0.01
100	20	1.00	0.50	0.96	0.07	0.00	0.62	0.06	0.01	0.40	0.07	0.01
100	20	1.00	0.80	0.78	0.07	0.01	0.64	0.06	0.02	0.50	0.08	0.03
100	20	1.00	0.90	0.53	0.07	0.04	0.47	0.07	0.04	0.38	0.07	0.04
100	20	1.00	0.95	0.42	0.08	0.07	0.33	0.06	0.06	0.31	0.08	0.08
100	20	0.00	1.00	0.35	0.99	1.00	0.54	0.99	1.00	0.65	0.99	1.00
100	20	0.50	1.00	0.37	1.00	1.00	0.58	1.00	1.00	0.74	1.00	1.00
100	20	0.80	1.00	0.38	1.00	1.00	0.58	1.00	1.00	0.68	1.00	1.00
100	20	0.90	1.00	0.34	0.99	0.79	0.47	0.98	0.78	0.57	0.98	0.80
100	20	0.95	1.00	0.31	0.60	0.32	0.40	0.62	0.33	0.39	0.60	0.30
100	20	1.00	1.00	0.32	0.07	0.12	0.29	0.07	0.10	0.26	0.09	0.11
100	20	0.50	0.80	0.92	1.00	1.00	0.99	1.00	1.00	1.00	1.00	1.00
100	20	0.80	0.50	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99
100	20	0.00	0.90	0.70	1.00	1.00	0.85	1.00	1.00	0.93	1.00	1.00
100	20	0.90	0.00	1.00	0.98	0.38	0.98	0.98	0.40	0.97	0.98	0.38
100	100	1.00	0.00	0.99	0.07	0.00	0.68	0.07	0.00	0.49	0.07	0.00
100	100	1.00	0.50	0.99	0.08	0.00	0.89	0.07	0.00	0.78	0.07	0.00
100	100	1.00	0.80	0.86	0.06	0.01	0.85	0.07	0.01	0.79	0.07	0.01
100	100	1.00	0.90	0.68	0.07	0.06	0.68	0.06	0.07	0.68	0.07	0.06
100	100	1.00	0.95	0.48	0.07	0.14	0.54	0.07	0.14	0.54	0.06	0.13
100	100	0.00	1.00	0.43	1.00	1.00	0.65	1.00	1.00	0.76	1.00	1.00
100	100	0.50	1.00	0.47	1.00	1.00	0.71	1.00	1.00	0.81	1.00	1.00
100	100	0.80	1.00	0.47	1.00	1.00	0.67	1.00	1.00	0.82	1.00	1.00
100	100	0.90	1.00	0.41	1.00	1.00	0.61	1.00	0.99	0.71	1.00	1.00
100	100	0.95	1.00	0.38	0.99	0.63	0.54	0.99	0.63	0.59	1.00	0.64
100	100	1.00	1.00	0.38	0.06	0.22	0.43	0.07	0.21	0.42	0.07	0.21
100	100	0.50	0.80	0.96	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
100	100	0.80	0.50	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
100	100	0.00	0.90	0.75	1.00	1.00	0.93	1.00	1.00	0.98	1.00	1.00
100	100	0.90	0.00	1.00	1.00	0.88	1.00	1.00	0.88	1.00	1.00	0.88

Table 3: Results for Inflation

		DFGLS	ADF	ADF			
series		X		\hat{e}	\hat{e}^1	$\frac{\text{var}(\widehat{\Delta e})}{\text{var}(\Delta X)}$	$\frac{\sigma(\lambda_i F_t)}{\sigma(\hat{e})}$
1	meats	-3.373 ⁺	-6.267 ⁺	-3.092 ⁺	-7.174 ⁺	0.857	0.406
2	fruits	-1.335	-4.018 ⁺	-1.305	-4.077 ⁺	0.995	0.088
3	usedcars	-2.059 ⁺	-4.163 ⁺	-1.801	-4.178 ⁺	0.994	0.063
4	schbooks	-1.640	-2.529	-1.533	-2.449	0.997	0.078
5	othfdhome	-3.324 ⁺	-4.225 ⁺	-2.815 ⁺	-5.224 ⁺	0.958	0.117
6	gaselec	-1.487	-1.887	-1.426	-1.989	0.991	0.089
7	fmaleapp	-1.117	-3.612 ⁺	-1.547	-3.440 ⁺	0.711	0.735
8	newveh	-1.743	-2.664	-1.742	-2.677	0.966	0.172
9	hskpsupp	-1.981 ⁺	-2.758	-1.708	-3.224	0.987	0.089
10	persedsvs	-1.076	-2.354	-0.933	-2.147	0.887	0.389
11	footwear	-2.144 ⁺	-2.860 ⁺	-3.169 ⁺	-3.898 ⁺	0.731	0.743
12	alcohol	-3.018 ⁺	-3.679 ⁺	-3.020 ⁺	-3.847 ⁺	0.964	0.186
13	shelter	-2.171 ⁺	-2.679	-2.307 ⁺	-2.776	0.983	0.111
14	maleapp	-2.494 ⁺	-3.421 ⁺	-3.675 ⁺	-4.607 ⁺	0.673	0.825
15	etaingds	-2.442 ⁺	-3.062	-2.799 ⁺	-3.495 ⁺	0.998	0.049
16	furnishings	-1.761	-2.110	-1.553	-2.749	0.920	0.289
17	appsvs	-1.421	-1.573	-1.086	-1.536	0.976	0.121
18	automaint	-1.830	-1.948	-1.763	-2.352	0.980	0.109
19	medgds	-0.923	-2.113	-0.673	-2.110	0.934	0.166
20	medsvs	-1.996 ⁺	-2.182	-1.407	-2.329	0.994	0.064
21	fdnothome	-1.615	-2.277	-1.413	-2.049	0.904	0.217
5% CV		-1.91	-2.82	-1.91	-3.31		
10% CV		-1.60	-2.52	-1.60	-3.02		

The ADF is based on Said and Dickey (1984) with 4 lags.

Figure 1: True and Estimated F_t when e_{it} is $I(1)$

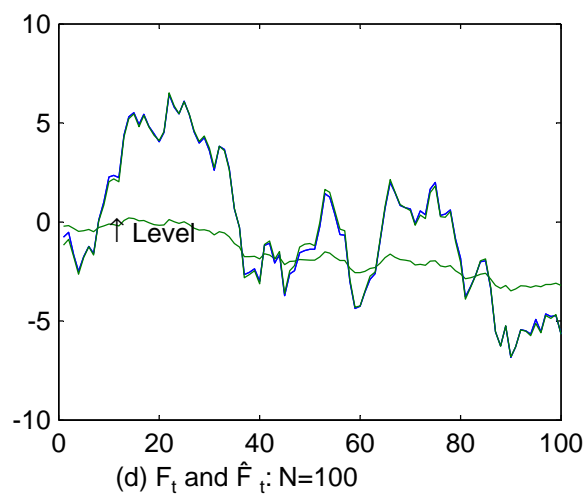
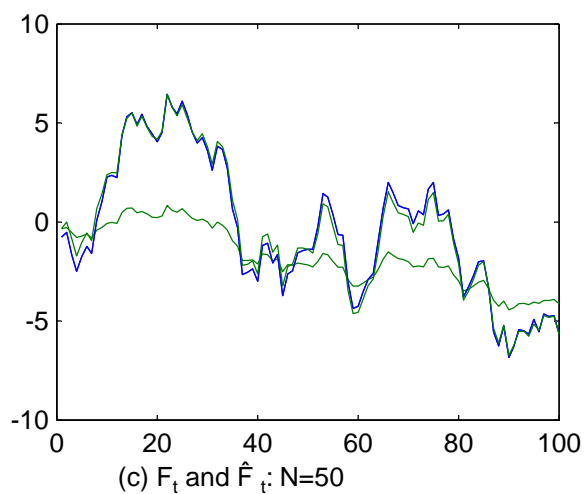
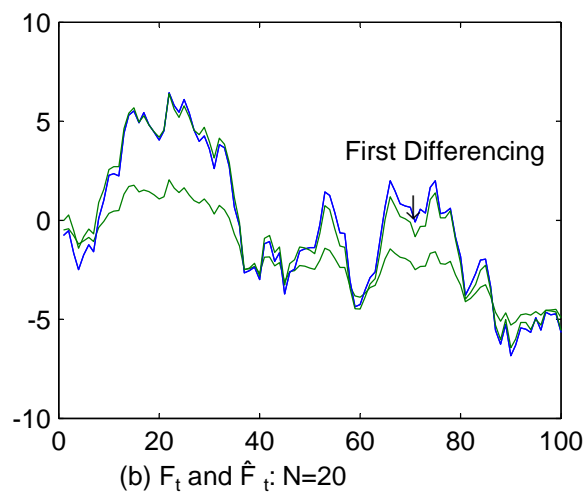
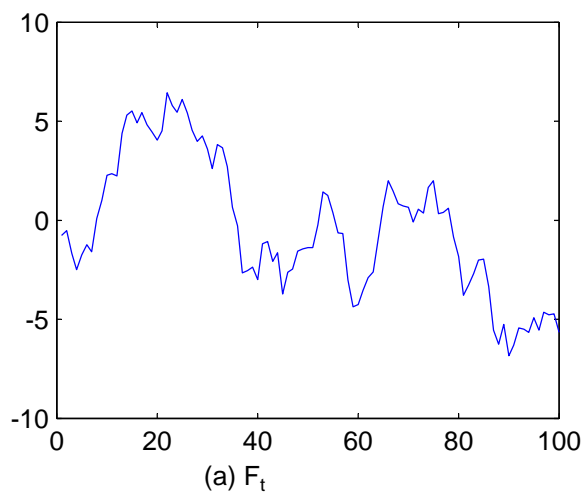


Figure 2: True and Estimated F_t when e_{it} is $I(0)$

