DEAKIN UNIVERSITY AUSTRALIA

Faculty of Business and Law
School of Accounting, Economics and Finance

School Working Paper

Financial Econometrics Series

SWP 2014/05

The Local Power of the CADF and CIPS Panel Unit Root Tests

J. Westerlund, M. Hosseinkouchack, and M. Solberger

The Local Power of the CADF and CIPS Panel Unit Root Tests

Joakim Westerlund*
Deakin University
Australia

Mehdi Hosseinkouchack
Goethe University Frankfurt
Germany

Martin Solberger
Uppsala University
Sweden

August 1, 2013

Abstract

Very little is known about the local power of second generation panel unit root tests that are robust to cross-section dependence. This paper derives the local asymptotic power functions of the CADF and CIPS tests of Pesaran (A Simple Panel Unit Root Test in Presence of Cross-Section Dependence, *Journal of Applied Econometrics* **22**, 265–312, 2007), which are among the most popular tests around.

JEL Classification: C12; C13; C33.

Keywords: Panel unit root test; common factor model; cross-sectional averages; cross-sectional dependence; local asymptotic power.

1 Introduction

Cross-section dependence can pose serious problems when testing the null hypothesis of a unit root, and much effort has therefore gone into the development of so-called "second-generation" test procedures that are robust to such dependencies.[1] Two of the most popular tests are the cross-section augmented Dickey–Fuller (CADF) and Im et al. (2003) (CIPS) tests of Pesaran (2007), where the latter is just the cross-section average of the former when applied to each cross-section unit. In fact, these tests have in recent years become workhorses of the industry, with a very large number of applications. The popularity of the CADF and

*Corresponding author: Deakin University, Faculty of Business and Law, School of Accounting, Economics and Finance, Melbourne Burwood Campus, 221 Burwood Highway, VIC 3125, Australia. Telephone: +61 3 924 46973. Fax: +61 3 924 46283. E-mail address: j.westerlund@deakin.edu.au.

[1]See Breitung and Pesaran (2008) for a recent survey of the panel unit root and cointegration literature, and Westerlund and Breitung (2013, Section 5) for a detailed discussion of the effects of cross-section dependence.

CIPS tests does not stop with applications, but extends also to theoretical work, where they have been used as a basis for numerous extensions, including tests for a unit root with multiple common factors, and cointegration (see, for example, Pesaran et al., 2013; Banerjee and Carrion-i-Silvestre, 2011).

Yet, strangely enough, given their popularity, while much is known regarding the asymptotic properties of CADF and CIPS under the unit root null, as far as we are aware, nothing is known regarding their (local) power properties. The purpose of the current paper is to fill this gap in the literature. This is an important undertaking, because the small-sample powers of CADF and CIPS differ quite substantially from those of most existing tests. In order to illustrate this, let us begin by considering the simulated 5% size and local power reported in Table 1. The data generating process (DGP) is given by

$$y_{i,t} = \rho_i y_{i,t-1} + \gamma_i f_t + \epsilon_{i,t}, \tag{1}$$

where $t = 1, ..., T$ and $i = 1, ..., N$ index the time series and cross-section units, respectively, $y_{1,0} = ... = y_{N,0} = 0$, f_t is a common factor with loading γ_i, and $\epsilon_{i,t}$ is idiosyncratic. In the simulations, $(f_t, \epsilon_{i,t})' \sim N(0, I_2)$. The autoregressive parameter ρ_i is given the following local-to-unity specification:

$$\rho_i = \exp(T^{-1} c_i), \tag{2}$$

where c_i is the so-called "Pitman drift" measuring the deviation from the unit root ($c_i = 0$). In the simulations, $c_i \sim U(a, b)$, where a and b are calibrated such that the mean and variance of c_i are equal to μ_c and σ_c^2, respectively. The results are generated for 3,000 panels of size $N = T = 100$.[2]

A typical finding in the panel unit root literature is that theoretically the local power of tests should only depend on the mean of c_i, and therefore there should be no dependence on the variance, or any other moment for that matter (see, for example, Moon et al., 2007). This means that, from a theoretical point of view, one can just as well assume that $c_1 = ... = c_N = c$; there are no additional insights to be gained by allowing c_i to vary. These results have been confirmed in numerous simulation studies. In particular, while for small values of N power is typically decreasing in the variance c_i, as N increases this effect tends to go away (see, for example, Moon and Perron, 2008; Moon et al., 2007).[3] In contrast to these

[2]The appropriate critical values are taken from Pesaran (2007, Tables I (a) and II (a)).

[3]The fact that power decreases with increases in the variance of c_i is partly expected, as most panel unit root

results, according to Table 1, in a vast majority of cases power is actually increasing in σ_c^2, and this is true even if N is as large as 100. We also see that this effect is greater the larger is γ. This is also quite unexpected, because usually (factor model-based) tests of this kind have asymptotic distributions that are free of nuisance parameters even under the local alternative (see Moon and Perron, 2004; Westerlund and Larsson, 2012). Hence, power is increasing in the variance of c_i, and the extent of this effect is increasing in the relative importance of the common component, as measured by γ. The exception is if $\sigma_c^2 = 0$, in which case the effect of γ is no longer present. This is true for $N = T = 100$, but holds also for all other sample sizes attempted.

The above results suggest that the conventional wisdom that local power should only depend on μ_c might not apply in case of CADF and CIPS. The study of the local power of these tests is therefore interesting not only in its own right, but also for what it might imply for the local power of second-generation panel unit root tests in general. In fact, the only other second-generation local power studies that we are aware of are those of Moon and Perron (2004), and Westerlund and Larsson (2012), who consider similar DGP setups, although the test statistics are quite different from the ones considered here. The above results would also seem to have certain implications for empirical work. In particular, given the completely different power profile of CADF and CIPS when compared to those of other tests, they indicate that the most powerful test to use in practice will depend critically on the magnitude of γ and the heterogeneity of ρ_i, which are both estimable prior to testing.

The purpose of the present paper is to make the discussion in the last paragraph a little more precise. The paper is organized as follows. Section 2 presents the assumptions that we will be working under, which are chosen with an eye on simplicity and transparency. Section 3 reports the results of the analysis of the local asymptotic power, which are evaluated using Monte Carlo simulation in Section 4. Section 5 concludes. Proofs are collected in Appendix.

2 Assumptions

The DGP is similar to the one considered in Section 1, and is given by (1), (2) and Assumption 1.

Assumption 1.

tests are based on pooling, and pooling is only efficient if the autoregressive parameters are homogenous (see Westerlund and Breitung, 2013, Section 2).

(i) $\epsilon_{i,t}$ is independently and identically distributed (iid) across both i and t with $E(\epsilon_{i,t}) = 0$, $E(\epsilon_{i,t}^2) = \sigma_{\epsilon,i}^2 > 0$ and $E(\epsilon_{i,t}^4) < \infty$;

(ii) f_t is iid across t with $E(f_t) = 0$, $E(f_t^2) = 1$ and $E(f_t^4) < \infty$;

(iii) $c_i \leq 0$ is iid with $\mu_j = E(c_i^j)$ and $|\mu_j| < \infty$ for $j \geq 0$;

(iv) $\epsilon_{i,t}$, f_t and c_i are mutually independent;

(v) $y_{1,0} = ... = y_{N,0} = 0$;

(vi) γ_i is a fixed coefficient such that $\overline{\gamma} = N^{-1}\sum_{i=1}^N \gamma_i \to \gamma \neq 0$ and $\overline{\beta} = N^{-1}\sum_{i=1}^N \beta_i \to \beta \neq 0$ as $N \to \infty$, where $\beta_i = \gamma_i/\sigma_{\epsilon,i}$.

Note how $\beta_i^2 = \gamma_i^2/\sigma_{\epsilon,i}^2$ can be viewed as the ratio of the variance of $\gamma_i f_t$ relative to that of $\epsilon_{i,t}$. It can therefore be seen as a measure of the relative importance of the common component, which, as already alluded at in Section 1, is going to turn out to be an important determinant of the local power of the CADF and CIPS tests.

The above assumptions are roughly the same as those in Pesaran (2007). The only exceptions are; (a) the local-to-unity specification of ρ_i, (b) the fixed specification of γ_i, and (c) the absence of deterministic constant and trend terms. The local-to-unity specification in (2) is, of course, key in our paper and is similar to those considered by, for example, Moon and Perron (2008), and Moon et al. (2007). The main difference is that the rate of shrinking is not allowed to depend on N. A general formulation of the shrinking neighborhood is given by $N^{-\kappa}T^{-1}$, where $\kappa \geq 0$. The CADF test has non-negligible local power when $\kappa = 0$, but not when $\kappa > 0$, which is in agreement with other time series tests. This means that for any fixed alternative $\rho_i < 1$, the power of this test should be flat in N, which is just as expected given that it is a time series test. The fact that also the CIPS test, which is a pooled panel test, requires $\kappa = 0$ is, on the other hand, unexpected. It means that as N increases CIPS will tend to be dominated by other panel data tests having non-negligible local power for $\kappa > 0$ (see Sections 3 and 4 for a detailed discussion). The requirement that c_i has all its moments is less restrictive than the otherwise so common bounded support assumption (see Moon and Perron, 2008; Moon et al., 2007), which obviously implies finite moments. In terms of the notation of Section 1, we have $\mu_c = \mu_1$ and $\sigma_c^2 = \mu_2 - \mu_1^2$. As for the assumption placed on γ_i, while the proofs make use of the fact that γ_i is non-random, the results provided in

Section 3 continue to hold also under the random coefficient assumption of Pesaran (2007), provided that γ_i is independent of all the other random elements of the model. The effect of deterministic constant and trend terms is discussed Section 3.

The appropriate null hypothesis in case of CADF is that cross-section unit i has a unit root, which corresponds to the restriction that $c_i = 0$. The alternative hypothesis is that $c_i < 0$, in which case $y_{i,t}$ is said to be "locally stationary". In contrast to CADF, CIPS is a joint panel test statistic. The appropriate null in this case is therefore that $c_1 = ... = c_N = 0$ (all the units of the panel have a unit root; $\rho_1 = ... = \rho_N = 1$), while the alternative is that $c_i < 0$ ($\rho_i < 1$) for at least some i.

Remark 1. The assumption of serial correlation free errors is restrictive but can be relaxed in a relatively straightforward manner by means of the usual augmentation trick (see Pesaran, 2007, Section 5).

Remark 2. The assumption that $y_{1,0} = ... = y_{N,0} = 0$ is crucial in the above model with no deterministic constant or trend terms. It will, however, be unimportant for the models with a constant and/or linear trend considered in Section 3 (provided that $y_{i,0} = O_p(1)$ for all i). Hence, for expositional brevity we can just as well set $y_{i,0}$ to zero.

Remark 3. The requirement that f_t has unit variance is an identifying assumption, and as such it is not a restriction (if the variance is not one, then it is "absorbed" into γ_i).

3 Local asymptotic power

3.1 CADF

The CADF test statistic applied to unit i is given by

$$\text{CADF}_i = \frac{y'_{i,-1} M_x \Delta y_i}{\hat{\sigma}_{\epsilon,i} \sqrt{y'_{i,-1} M_x y_{i,-1}}},$$

where $M_x = I_{T-1} - x(x'x)^{-1}x'$, $x = (\Delta \overline{y}, \overline{y}_{-1})$, $\Delta y_i = (\Delta y_{i,2}, ..., \Delta y_{i,T})'$, $y_{i,-1} = (y_{i,1}, ..., y_{i,T-1})'$, $\Delta \overline{y} = N^{-1} \sum_{i=1}^{N} \Delta y_i$ with a similar definition of $\overline{y}_{-1}$, and $\hat{\sigma}^2_{\epsilon,i} = T^{-1}(\Delta y_i)' M_x \Delta y_i$.

Remark 4. Note how CADF_i is just the conventional Dickey-Fuller test augmented with x. The intuition for this is very simple. Assuming for simplicity that $\rho_1 = ... = \rho_N = \rho$, we have $\Delta \overline{y}_t = (\rho - 1)\overline{y}_{t-1} + \overline{\gamma} f_t + \overline{\epsilon}_t$, or $f_t = (\Delta \overline{y}_t - (\rho - 1)\overline{y}_{t-1} - \overline{\epsilon}_t)/\overline{\gamma} = (\Delta \overline{y}_t - (\rho - 1)\overline{y}_{t-1})/\overline{\gamma} +$

$o_p(1)$, where the last equality is due to the fact that $\epsilon_{i,t}$ is mean zero and independent across i. Hence, f_t may be approximated by a linear combination of $\Delta\overline{y}_t$ and $\overline{y}_{t-1}$, which is also the reason for the augmentation by x.

Remark 5. Since under the null $y_{i,t}$ and hence also $\overline{y}_t$ are unit root non-stationary, this means a regression of $y_{i,t}$ onto $\Delta\overline{y}_t$ and $\overline{y}_t$ is spurious. Therefore, $M_x y_i$, the defactored data, are going to behave just like the residuals from a standard time series spurious regression, whose limiting counterpart depends on the Brownian motion associated with the unit root regressor, here given by $\overline{y}_t$. This means that the dependence on the common component will not disappear, not even asymptotically, causing CADF$_i$ to be cross-section dependent (as $\overline{y}_t$ is common for all i). This is unlike most, if not all, other second-generation approaches where the defactored data are (asymptotically) cross-section independent (see, for example, Bai and Ng, 2004; 2010; Moon and Perron, 2004; Westerlund and Larsson, 2012). The reason for this difference is that in these other approaches the estimation is typically carried out while imposing the null. In the current context this means setting $\rho_1 = ... = \rho_N = 1$, and so, by using the same steps as in Remark 4 above, we have $f_t = \Delta\overline{y}_t/\overline{\gamma} + o_p(1)$. Since $\Delta\overline{y}_t$ is stationary, the resulting defactoring regression is no longer spurious, and therefore the dependence on (the Brownian motion associated with) $\overline{y}_t$ is eliminated. It is the cross-section dependence that gives the CADF and CIPS tests their unique power profile.

To succinctly express the limiting distribution of CADF$_i$ (also when deterministic terms are added to the test regression), it is useful to employ the following notation. Let $X(r)$ and $Y(r)$ denote two scalar processes on $r \in [0,1]$. Consider the continuous time regression of $Y(r)$ on $X(r)$, $Y(r) = \hat{\alpha}'X(r) + Q_X Y(r)$, where $\hat{\alpha}$ solves $\min_\alpha \int_0^1 ||Y(r) - \alpha' X(r)||^2 dr$. The continuous time regression residual, $Q_X Y(r)$, is defined as

$$Q_X Y(r) = Y(r) - \int_0^1 Y(v)X(v)dv \left(\int_0^1 X(v)^2 dv \right)^{-1} X(r).$$

Making use of this notation, we define

$$\begin{aligned} A_i &= \int_0^1 Q_{J_f} J_{y,i}(r)(dW_{\epsilon,i}(r) - \beta_i Q_{J_f} K_f(r)dr), \\ B_i &= \int_0^1 (Q_{J_f} J_{y,i}(r))^2 dr, \end{aligned}$$

with

$$
\begin{aligned}
J_{f,i}(r) &= \int_0^r \exp(c_i(r-v))dW_f(v), \\
J_{\epsilon,i}(r) &= \int_0^r \exp(c_i(r-v))dW_{\epsilon,i}(v), \\
J_f(r) &= \sum_{j=0}^{\infty} \mu_j \int_0^r \frac{(r-v)^j}{j!} dW_f(v), \\
J_{y,i}(r) &= \beta_i J_{f,i}(r) + J_{\epsilon,i}(r), \\
K_f(r) &= \sum_{j=0}^{\infty} (\mu_{j+1} - \mu_1 \mu_j) \int_0^r \frac{(r-v)^j}{j!} dW_f(v),
\end{aligned}
$$

where $W_f(r)$ and $W_{\epsilon,i}(r)$ are two independent standard Brownian motions, and $Q_{J_f} J_{y,i}(r)$ and $Q_{J_f} K_f(r)$ are the residuals from regressing $J_{y,i}(r)$ and $K_f(r)$, respectively, onto $J_f(r)$.

Theorem 1. *Under Assumption 1, as N, $T \to \infty$ with $T/N \to \tau < \infty$,*

$$
\text{CADF}_i \to_w c_i \sqrt{B_i} + \frac{A_i}{\sqrt{B_i}},
$$

where $\to_w$ signifies weak convergence.

We begin with a brief discussion of local power implications of Theorem 1; then we also have some general remarks. According to Theorem 1 the asymptotic distribution of CADF_i is a sum of two terms. Since $E(B_i) > 0$, the first term represents a drift in mean that is there only under the alternative that $c_i < 0$. Consider the mean of the second term, which is there even if $c_i = 0$. Since $J_{f,i}(r)$ and $J_{\epsilon,i}(r)$ are independent by assumption, and $E[\int_0^1 J_{\epsilon,i}(r) dW_{\epsilon,i}(r)] = 0$, we have

$$
E(A_i) = E\left(\int_0^1 Q_{J_f} J_{y,i}(r)(dW_{\epsilon,i}(r) - \beta_i Q_{J_f} K_f(r) dr) \right) = -\beta_i^2 \int_0^1 E[Q_{J_f} J_{f,i}(r) Q_{J_f} K_f(r)] dr.
$$

Let us denote by $\mathcal{F}$ the sigma-field generated by $\{W_f(r)\}_{r \geq 0}^1$. By using the results provided in the proof of Theorem 1 (see Appendix), it is not difficult to show that $E(Q_{J_f} J_{f,i}(r)|\mathcal{F}) = Q_{J_f} E(J_{f,i}(r)|\mathcal{F}) = Q_{J_f} J_f(r) = 0$. Hence,

$$
\int_0^1 E[Q_{J_f} J_{f,i}(r) Q_{J_f} K_f(r)] dr = \int_0^1 E[E(Q_{J_f} J_{f,i}(r)|\mathcal{F}) Q_{J_f} K_f(r)] dr = 0,
$$

giving $E(A_i) = 0$. The mean of the numerator of the second term is therefore (asymptotically) zero. However, while the mean is unaffected by c_i, the variance is not, as captured by the fact that if $c_i < 0$, then $J_{f,i}(r) \neq W_f(r)$, $J_{\epsilon,i}(r) \neq W_{\epsilon,i}(r)$ and $J_f(r) \neq W_f(r)$. Thus, the presence of c_i under the alternative exerts both a mean and a variance effect. The following predictions apply, which are all consistent with the Monte Carlo evidence reported in Table 1:

- While the mean of A_i is unaffected by the presence of β_i, B_i and the variance of A_i are not. Note in particular how the first- and second-order derivatives of B_i with respect to β_i are given by $2\int_0^1 Q_{J_f} J_{y,i}(r) Q_{J_f} J_{f,i}(r) dr$ and $2\int_0^1 (Q_{J_f} J_{f,i}(r))^2 dr > 0$. Therefore, power should tend to increase with β_i.

- If $c_1 = ... = c_N = c$, then $J_{f,i}(r) = J_f(r)$, and therefore, since $Q_{J_f} J_f(r) = 0$, we have $Q_{J_f} J_{y,i}(r) = Q_{J_f}(\beta_i J_{f,i}(r) + J_{\epsilon,i}(r)) = Q_{J_f}(\beta_i J_f(r) + J_{\epsilon,i}(r)) = Q_{J_f} J_{\epsilon,i}(r)$. Moreover, since in this case, $(\mu_{j+1} - \mu_1 \mu_j) = (c^{j+1} - c_1 c^j) = 0$ for all $j \geq 0$, we have $K_f(r) = 0$. The asymptotic distribution of CADF_i therefore simplifies to

$$c\sqrt{\int_0^1 (Q_{J_f} J_{\epsilon,i}(r))^2 dr} + \frac{\int_0^1 Q_{J_f} J_{\epsilon,i}(r) dW_{\epsilon,i}(r)}{\sqrt{\int_0^1 (Q_{J_f} J_{\epsilon,i}(r))^2 dr}},$$

  which is completely nuisance parameter free. In other words, if $\sigma_c^2 = 0$, then power should be independent of β_i.

- If $c_1 = ... = c_N = 0$ (so that the null is true), then the limiting distribution of CADF_i simplifies to

$$\text{CADF}_i \to_w \frac{\int_0^1 Q_{W_f} W_{\epsilon,i}(r) dW_{\epsilon,i}(r)}{\sqrt{\int_0^1 (Q_{W_f} W_{\epsilon,i}(r))^2 dr}},$$

  where $Q_{W_f} W_{\epsilon,i}(r)$ is the residual obtained by regressing $W_{\epsilon,i}(r)$ onto $W_f(r)$.[4] Consequently, while power will generally depend on β_i, size accuracy will not.

Hence, based on Theorem 1 we can explain most of the observed test behavior, including the dependence on β_i and σ_c^2.

Remark 6. Theorem 1 can be easily extended to cover also models with deterministic constant and/or trend terms, if properly demeaned and/or detrended data are used. If the data are (full-sample OLS) demeaned, then Theorem 1 applies with $J_{y,i}(r)$ and $K_f(r)$ replaced by the residuals obtained by regressing these processes on $D(r) = 1$, that is, $J_{y,i}(r)$ and $K_f(r)$ are

[4] Appropriate left-tail critical values for use with CADF_i can be found in Pesaran (2007, Tables I (a)–II (c)).

replaced by $Q_D J_{y,i}(r)$ and $Q_D K_f(r)$, respectively. If the data are detrended, then we simply set $D(r) = (1, r)'$.

Remark 7. Because of the dependence on $W_f(v)$, while free of nuisance parameters, the local asymptotic distribution of CADF_i is not cross-section independent. However, as remarked by Pesaran (2007, Remark 3.3), conditional on $\mathcal{F}$ the asymptotic distribution is in fact independent (and even iid).

3.2 CIPS

The CIPS test is simply the cross-sectional average of the individual CADF tests, that is, $\text{CIPS} = N^{-1} \sum_{i=1}^{N} \text{CADF}_i$. In view of Theorem 1, and the fact that conditional on $\mathcal{F}$ the asymptotic distribution of CADF_i is iid (see Remark 7 above), it is clear that as N, $T \to \infty$ with $T/N \to \tau < \infty$,

$$\text{CIPS} = \lim_{N \to \infty} N^{-1} \sum_{i=1}^{N} E\left(c_i \sqrt{B_i} + \frac{A_i}{\sqrt{B_i}} | \mathcal{F} \right) + o_p(1). \tag{3}$$

Hence, provided that CADF_i has finite moments, unconditionally CIPS converges in distribution. To ensure that the required moments exist, Pesaran (2007) used a truncated version of CIPS (which has finite moments by construction). However, since the truncation almost never took effect (see Pesaran, 2007, page 278), in what follows we focus on the untruncated test statistic. As already mentioned, the presence of c_i under the local alternative has two effects on the asymptotic distribution of CADF_i. The first, captured by $c_i \sqrt{B_i}$, is a drift in mean, while the second is an increase in variance, as captured by the fact that if $c_i \neq 0$, then $J_{f,i}(r) \neq W_f(r)$, $J_{\epsilon,i}(r) \neq W_{\epsilon,i}(r)$ and $J_f(r) \neq W_f(r)$. The first effect is typically the dominant one, as is clear from Table 1. In the analysis of the local power of CIPS, we therefore focus on this term.

Theorem 2. *It holds that*

$$E(c_i \sqrt{B_i}) \sim \mu_1 \Theta_i,$$

where

$$
\begin{aligned}
\Theta_i &= \theta_i - \frac{\omega_1^2(\beta_i^4+1)}{8\theta_i^3} - \frac{\beta_i^2\omega_3^2}{2\theta_i^3} - \frac{\beta_i^2\alpha_1(1+\beta_i^2)}{\alpha_4}\frac{\omega_4^2}{2\theta_i^3} - \frac{\alpha_1(1+\beta_i^2)}{\alpha_4}\frac{\omega_5^2}{2\theta_i^3} \\
&+ \frac{3\alpha_4\beta_i^4 - 4\alpha_1(\beta_i^4+\beta_i^2)}{4\alpha_4}\frac{\omega_6^2}{2\theta_i^3} - \frac{\beta_i^2}{4}\frac{\omega_{12}}{\theta_i^3} + \frac{\beta_i^4}{2}\frac{\omega_{14}}{\theta_i^3} - \frac{\beta_i^4}{4}\frac{\omega_{16}}{\theta_i^3} + \beta_i^2\frac{\omega_{35}}{\theta_i^3} \\
&- \frac{\alpha_4\beta_i^4 - 2\alpha_1(\beta_i^4+\beta_i^2)}{2\alpha_4}\frac{\omega_{46}}{\theta_i^3}, \\
\theta_i &= \sqrt{\alpha_1 + (\alpha_1 - \alpha_4)\beta_i^2},
\end{aligned}
$$

with $\sim$ signifying asymptotic equivalence, and α_1, α_4, ω_1^2, ω_3^2, ω_4^2, ω_5^2, ω_6^2, ω_{12}, ω_{14}, ω_{16}, ω_{35} and ω_{46} given in Appendix.

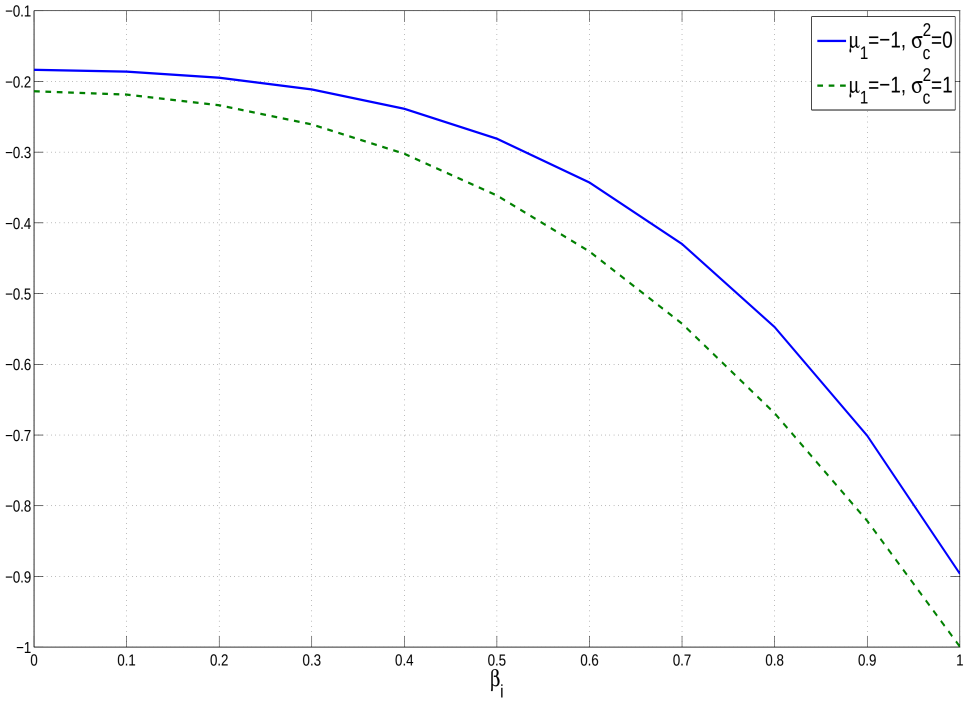
Denote by $M_c(x)$ the moment generating function of c_i. In Appendix we show that all constants in Theorem 2 are defined as integrals of $M_c(x)$ over $x \in [0,1]$. It therefore makes sense to approximate $M_c(x)$ by $1 + \mu_1 x + \mu_2 x^2/2$. $\mu_1\Theta_i$ will therefore be a function of μ_1, σ_c^2 and β_i, a function that can be illustrated graphically. Figure 1 shows that $|\mu_1\Theta_i|$ is increasing in β_i for different values of μ_1 and σ_c^2, that is, regardless of how c_i is drawn, higher values of γ_i relative to $\sigma_{\epsilon,i}$ result in a larger (in absolute value) drift term, which in turn implies higher local power for CIPS. This effect is clearly visible in Table 1. Figure 2 illustrates how μ_1 and σ_c^2 affect $\mu_1\Theta_i$ when $\beta_i = 1$ is kept fixed. It is seen that $|\mu_1\Theta_i|$ is generally increasing in both μ_1 and σ_c^2. The same is true for all values of β_i. Hence, regardless of the value of β_i, the larger (in absolute value) the values of μ_1 and σ_c^2, the larger the value of $|\mu_1\Theta_i|$, which is in turn suggestive of higher power. This is again consistent with the evidence reported in Table 1.

Remark 8. The rate of shrinking of the local alternative, T^{-1}, is lower than the usual $N^{-1/2}T^{-1}$ rate in absence of deterministic constant and trend terms (see, for example, Moon et al., 2007). The reason for this is that the defactored data are not cross-section independent. The intuition for this is simple. Consider the asymptotic distribution of CIPS. With $\rho_i = \exp(N^{-\kappa}T^{-1}c_i)$, this distribution can be written as (for N and T large enough)

$$
\begin{aligned}
\sqrt{N}\text{CIPS} &= N^{-1/2}\sum_{i=1}^{N}\left(N^{-\kappa}c_i\sqrt{B_i} + \frac{A_i}{\sqrt{B_i}}\right) + o_p(1) \\
&= N^{1/2-\kappa}N^{-1}\sum_{i=1}^{N}c_i\sqrt{B_i} + N^{-1/2}\sum_{i=1}^{N}\frac{A_i}{\sqrt{B_i}} + o_p(1).
\end{aligned}
$$

Had A_i and B_i been independent across i, given that it has a zero mean (and four finite moments), a standard central limit-type argument would suggest that the second term is nor-

Figure 1: $\mu_1\Theta_i$ as a function of β_i for different values of μ_1 and σ_c^2.



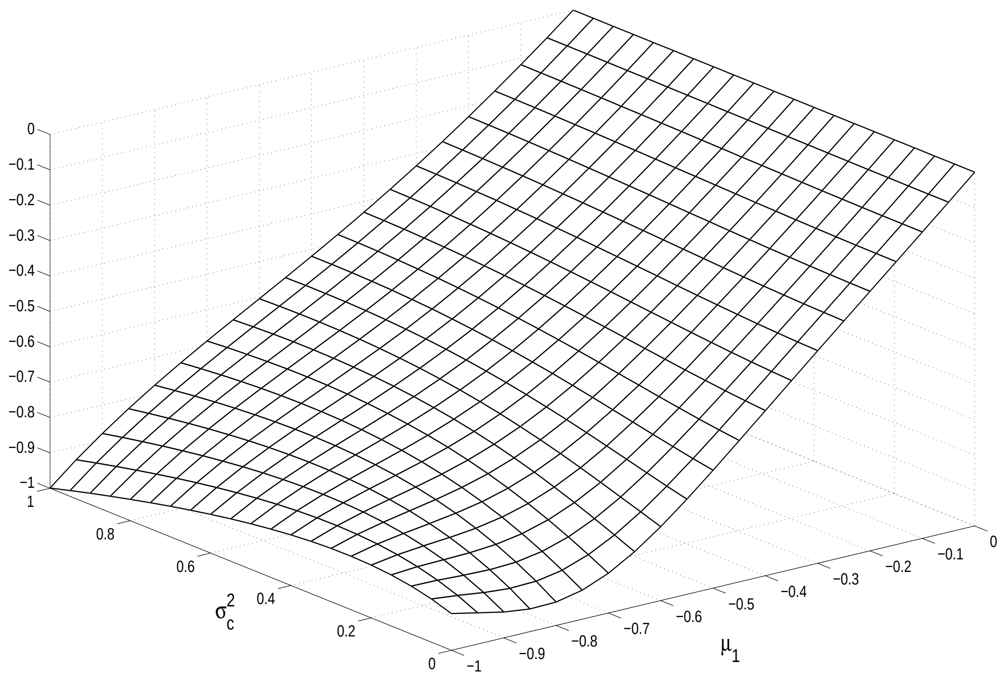
Note: The x- and y-axes display the value of β_i and $\mu_1\Theta_i$, respectively.

mal. Similarly, by a law of large numbers, $N^{-1}\sum_{i=1}^{N} c_i\sqrt{B_i} = \mu_1 N^{-1}\sum_{i=1}^{N} E(\sqrt{B_i}) + o_p(1)$, which is $O_p(1)$. Therefore, in case of cross-section independence, the appropriate test statistic to use is not CIPS but rather $\sqrt{N}$CIPS, which has non-negligible (and non-increasing) local power for $\kappa = 1/2$.

Remark 9. Most studies that consider local power in panels suppose that the data are cross-section independent, and a vast majority of these studies only consider tests based on within pooling, in which the numerator and denominator are summed over the cross-section before the ratio is taken (see, for example, Moon and Perron, 2008; Moon et al., 2007). In fact, the only other local power study of a between-type statistic, such as the one considered here, is that of Harris et al. (2009). They show more formally that the intuition given in Remark 8 above is correct, and that the normalized average of individual augmented Dickey–Fuller (ADF) test statistics has non-negligible local power within $N^{-1/2}T^{-1}$-neighborhoods of the null.

Remark 10. Not all factor-based second-generation test statistics have negligible local power

Figure 2: $\mu_1\Theta_i$ as a function of μ_1 and σ_c^2 when $\beta_i = 1$.



Note: The x-, y- and z-axes display the value of μ_1, σ_c^2 and $\mu_1\Theta_i$, respectively.

for $\kappa > 0$. In fact, provided that there are no deterministic constant and trend terms, most tests of this type have power within $N^{-1/2}T^{-1}$-neighborhoods of the null (see Moon and Perron, 2004). The reason for this difference lies with the way that the factors are estimated and removed. Here the defactored data are cross-section dependent, which implies a relatively low value of κ compatible with non-negligible local power when compared to approaches in which the defactored data are (asymptotically) cross-section independent (see Remarks 5 and 8).

4 Monte Carlo simulation

In this section we investigate briefly the relative performance of the CADF$_i$ and CIPS tests when compared to some existing second-generation unit root tests. Specifically, the CADF$_i$ and CIPS tests are compared with the $ADF_{\hat{e}}(i)$ and $P_{\hat{e}}$ tests of Bai and Ng (2004), which are among the most popular ones around with a huge number of applications. Unlike CADF$_i$ and CIPS, $ADF_{\hat{e}}(i)$ and $P_{\hat{e}}$ are based on using the principal components method to estimate

f_t. While $ADF_{\hat{e}}(i)$ is a time series test, $P_{\hat{e}}$ is a pooled panel test; therefore, the most relevant comparison here is between, on the one hand, $ADF_{\hat{e}}(i)$ and CADF$_i$, and, on the other hand, $P_{\hat{e}}$ and CIPS. Our purpose is to evaluate relative power as β_i and σ_c^2 varies. In particular, given the rather unique power profile of CADF$_i$ and CIPS, it seems reasonable to expect that no one test will be strictly preferred over the others, and that the choice of which test to use will depend to a large extent on the DGP. This is important because as far as we are aware there are no other Monte Carlo study that has considered the local power of CADF$_i$ and CIPS (see, for example, Gengenbach et al., 2009; De Silva et al., 2009).

The DGP is the same as in Section 1. Thus, since $\sigma_{\epsilon,i}^2 = 1$, we have $\beta_i = \beta = \gamma$. As before, all results are based on 3,000 replications. Unlike the tests considered here, $ADF_{\hat{e}}(i)$ and $P_{\hat{e}}$ can allow for more than one factor. Therefore, to make the results comparable, we assume that the researcher knows that there is just one factor present.

The results from the size and power comparisons are reported in Tables 2 and 3, respectively. They can be summarized as follows.

- Except for some minor distortions, all four tests are correctly sized.
- The local power of CADF$_i$ and CIPS is quite flat in N and T, which is as expected, because asymptotically there is no dependence on the sample size. In particular, power should not depend on N (see the discussion following Assumption 1).
- $P_{\hat{e}}$ is based on combining p-values, which are independent. It is therefore expected to have power in $N^{-1/2}T^{-1}$-neighborhoods of the null, which in turn means that the local power of this test in the current DGP with $\rho_i = \exp(T^{-1}c_i)$ should be increasing in N. This is true when $a = b$; however, whenever $a \neq b$ and/or β is not "small", then this need not be the case. For example, unreported results suggest that when $a = b = -1$, an increase in $N = T$ from 50 to 100 causes power to increase from 75% to 96%. By contrast, when $\beta \geq 10$, $a = -12$ and $b = 0$, an increase in $N = T$ from 50 to 100 actually leads to a decrease in power, which is partly expected given the previous literature (see Section 1).
- If $a = b = -3$, the level of power is unaffected by β, which is in accordance with our expectations (see the discussion following Theorem 1). In this case, CADF$_i$ (CIPS) is dominated by $ADF_{\hat{e}}(i)$ ($P_{\hat{e}}$). In fact, according to the results reported in Table 3 the power of $ADF_{\hat{e}}(i)$ ($P_{\hat{e}}$) is almost two times that of CADF$_i$ (CIPS).

- If $a \neq b$, then the relative power depends critically on the value taken by β. In particular, while the power of CADF$_i$ and CIPS is increasing in β, the power of $ADF_{\hat{e}}(i)$ and $P_{\hat{e}}$ is decreasing. The increase in relative power of CADF$_i$ and CIPS is largest among the smaller values of β. Therefore, any power disadvantage among the smaller values of β is quickly eliminated as β increases. Based on the results reported in Table 1, CADF$_i$ and CIPS dominate for all $\gamma \geq 5$ (unreported results suggest that the breakpoint is close to $\gamma = 2$).

The above results lead to the following very simple practical guideline: if β_i is "large" use CADF$_i$ (CIPS); otherwise, use $ADF_{\hat{e}}(i)$ ($P_{\hat{e}}$). Of course, in practice β_i is unknown; however, it can be easily estimated. Specifically, while γ_i and f_t are separately unidentifiable, we can estimate their product, $\gamma_i f_t$. In particular, letting $\hat{f}_t = \Delta \overline{y}_t$ (see Remark 5), $\hat{\gamma}_i$ can be obtained as the least squares slope in a time series regression of $\Delta y_{i,t}$ onto a constant and $\hat{f}_t$. Our proposed estimator of β_i is given by $\hat{\beta}_i = \hat{\gamma}_i \hat{\sigma}_f / \hat{\sigma}_{\epsilon,i}$, where $\hat{\sigma}_f^2 = T^{-1} \sum_{t=1}^{T} \hat{f}_t^2$. As a summary measure, we recommend using the average $\hat{\beta}_i$. Another important parameter here is σ_c^2, which can be inferred by looking at the variation of the estimated values of $\rho_1, ..., \rho_N$.

5 Conclusion

The CADF and CIPS tests of Pesaran (2007) are two of the most popular second-generation panel unit root tests around. Strangely enough, however, while the small-sample properties of these tests has been heavily scrutinized (see, for example, Gengenbach et al., 2009; De Silva et al., 2009; Pesaran, 2007), as far as we are aware nothing is known regarding local power. This is true in small samples, but also in theory, which presently only covers the behavior under the null hypothesis of a unit root (see Pesaran, 2007). As a response to this, the current paper offers an in-depth analysis of the local asymptotic power of CADF and CIPS. The new results are shown to deliver significant insights that go a long way towards explaining the observed test behavior.

References

Bai, J., and S. Ng (2004). A Panic Attack on Unit Roots and Cointegration. *Econometrica* **72**, 1127–1177.

Bai, J., and S. Ng (2010). Panel Unit Root Tests With Cross-Section Dependence: A Further Investigation. *Econometric Theory* **26**, 1088–1114.

Banerjee, A., and J. L. Carrion-i-Silvestre (2011). Testing for Structural Breaks Using Common Correlated Effects Estimators. Discussion Paper 11–16, Department of Economics, University of Birmingham.

Breitung J., and M. H. Pesaran (2008). Unit Roots and Cointegration in Panels. In Matyas, L., and P. Sevestre (Eds.), *The Econometrics of Panel Data*, Klüwer Academic Publishers, 279–322.

De Silva, S., K. Hadri and A. R. Tremayne (2009). Panel Unit Root Tests in the Presence of Cross-Sectional Dependence: Finite Sample Performance and an Application. *Econometrics Journal* **12**, 340–366.

Gengenbach, C., F. C. Palm and J.-P. Urbain (2009). Panel Unit Root Tests in the Presence of Cross-Sectional Dependencies: Comparison and Implications for Modelling. *Econometric Reviews* **29**, 111–145.

Harris, D., D. I. Harvey, S. J. Leybourne and N. D. Sakkas (2010). Local Asymptotic Power of the Im-Pesaran-Shin Panel Unit Root Test and the Impact of Initial Observations. *Econometric Theory* **26**, 311–324.

Im, K. S., M. H. Peseran and Y. Shin (2003). Testing for Unit Roots in Heterogeneous Panels. *Journal of Econometrics* **115**, 53–74.

Moon, H. R., and B. Perron (2004). Testing for Unit Root in Panels with Dynamic Factors. *Journal of Econometrics* **122**, 81–126.

Moon, H. R., and B. Perron (2008). Asymptotic Local Power of Pooled t-Ratio Tests for Unit Roots in Panels with Fixed Effects. *Econometrics Journal* **11**, 80–104.

Moon, H. R., B. Perron and P. C. B. Phillips (2007). Incidental Trends and the Power of Panel Unit Root Tests. *Journal of Econometrics* **141**, 416–459.

Pesaran, M. H. (2006). Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure. *Econometrica* **74**, 967–1012.

Pesaran, H. M. (2007). A Simple Panel Unit Root Test in the Presence of Cross Section Dependence. *Journal of Applied Econometrics* **22**, 265–312.

Pesaran, H. M., L. V. Smith, and T. Yamagata (2013). Panel Unit Root Tests in the Presence of a Multifactor Error Structure. *Journal of Econometrics* **175**, 94–115.

Westerlund, J., and R. Larsson (2012). Testing for Unit Roots in a Panel Random Coefficient Model. *Journal of Econometrics* **167**, 254–273.

Westerlund, J., and J. Breitung (2013). Lessons from a Decade of IPS and LLC. *Econometric Reviews* **32**, 547–591.

Appendix: Proofs

Proof of Theorem 1.

Let $x_t = (\Delta\overline{y}_t, \overline{y}_{t-1})'$, where $\Delta\overline{y}_t$ and $\overline{y}_{t-1}$ are as in the text, $\rho = E(\rho_i)$, $\theta = (1, -\rho)'/\overline{\gamma}$ and

$$\hat{f}_t = \frac{1}{\overline{\gamma}}(\Delta\overline{y}_t - (\rho - 1)\overline{y}_{t-1}) = \theta' x_t \tag{4}$$

such that

$$\hat{f}_t = f_t + \frac{1}{\overline{\gamma}}(\overline{\epsilon}_t + \overline{r}_{t-1}(\rho)), \tag{5}$$

where $r_{i,t}(\rho) = (\rho_i - \rho)y_{i,t}$, $\overline{r}_t(\rho) = N^{-1}\sum_{i=1}^{N} r_{i,t}(\rho)$, and similar definitions of $\overline{\gamma}$ and $\overline{\epsilon}_t$. It follows that

$$y_{i,t} = \rho_i y_{i,t-1} + \gamma_i f_t + \epsilon_{i,t} = \rho_i y_{i,t-1} + \gamma_i \hat{f}_t - \gamma_i d_t + \epsilon_{i,t} = \rho_i y_{i,t-1} + \gamma_i \theta' x_t - \gamma_i d_t + \epsilon_{i,t},$$

where $d_t = \hat{f}_t - f_t = (\overline{\epsilon}_t + \overline{r}_{t-1}(\rho))/\overline{\gamma}$, or, in vector notation,

$$\Delta y_i = (\rho_i - 1)y_{i,-1} + \gamma_i x\theta - \gamma_i d + \epsilon_i, \tag{6}$$

where Δy_i, $y_{i,-1}$ and x are as in the text, $\epsilon_i = (\epsilon_{i,2}, ..., \epsilon_{i,T})'$ and $d = (d_2, ..., d_T)' = (\overline{\epsilon} + \overline{r}_{-1}(\rho))/\overline{\gamma}$ with $\overline{\epsilon} = (\overline{\epsilon}_2, ..., \overline{\epsilon}_T)'$ and $\overline{r}_{-1}(\rho) = (\overline{r}_1(\rho), ..., \overline{r}_{T-1}(\rho))'$.

Consider the numerator of CADF$_i$, $y'_{i,-1}M_x\Delta y_i$. Making use of (6) and the fact that $M_x x = 0$, it is clear that the following must hold:

$$y'_{i,-1}M_x\Delta y_i = (\rho_i - 1)y'_{i,-1}M_x y_{i,-1} - \gamma_i y'_{i,-1}M_x d + y'_{i,-1}M_x\epsilon_i. \tag{7}$$

By further use of the definition of M_x, the first term on the right-hand side can be expanded as

$$y'_{i,-1}M_x y_{i,-1} = y'_{i,-1}y_{i,-1} - y'_{i,-1}x(x'x)^{-1}x'y_{i,-1},$$

where, letting $D_T = \text{diag}(\sqrt{T}, T)$,

$$\begin{aligned}
D_T^{-1}x'xD_T^{-1} &= \begin{bmatrix} T^{-1}(\Delta\overline{y})'\Delta\overline{y} & T^{-3/2}(\Delta\overline{y})'\overline{y}_{-1} \\ T^{-3/2}\overline{y}'_{-1}\Delta\overline{y} & T^{-2}\overline{y}'_{-1}\overline{y}_{-1} \end{bmatrix} \\
&= \begin{bmatrix} T^{-1}(\Delta\overline{y})'\Delta\overline{y} & 0 \\ 0 & T^{-2}\overline{y}'_{-1}\overline{y}_{-1} \end{bmatrix} + O_p(T^{-1/2}), \\
T^{-1}y'_{i,-1}xD_T^{-1} &= (T^{-3/2}y'_{i,-1}\Delta\overline{y}, T^{-2}y'_{i,-1}\overline{y}_{-1}) = (0, T^{-2}y'_{i,-1}\overline{y}_{-1}) + O_p(T^{-1/2}).
\end{aligned}$$

It follows that, by Taylor expansion,

$$\begin{aligned} T^{-2}y'_{i,-1}M_x y_{i,-1} &= T^{-2}y'_{i,-1}y_{i,-1} - T^{-1}y'_{i,-1}xD_T^{-1}(D_T^{-1}x'xD_T^{-1})^{-1}T^{-1}D_T^{-1}x'y_{i,-1} \\ &= T^{-2}(y'_{i,-1}y_{i,-1} - y'_{i,-1}\overline{y}_{-1}(\overline{y}'_{-1}\overline{y}_{-1})^{-1}\overline{y}'_{-1}y_{i,-1}) + O_p(T^{-1/2}). \end{aligned}$$

Consider $T^{-1/2}\overline{y}_t$. Define $J_{f,i,t} = T^{-1/2}\sum_{s=1}^t \rho_i^{t-s} f_s$ and $J_{\epsilon,i,t} = T^{-1/2}\sum_{s=1}^t \rho_i^{t-s}\epsilon_{i,s}$. Hence, since $y_{i,0} = 0$, we can use repeated substitution to obtain

$$T^{-1/2}y_{i,t} = \gamma_i J_{f,i,t} + J_{\epsilon,i,t}, \tag{8}$$

which in turn implies

$$T^{-1/2}\overline{y}_t = N^{-1}\sum_{i=1}^N \gamma_i J_{f,i,t} + \overline{J}_{\epsilon t}.$$

Clearly, $\sqrt{N}\overline{J}_{\epsilon t} = N^{-1/2}\sum_{i=1}^N J_{\epsilon,i,t} = O_p(1)$. Moreover, by substitution of $\rho_i = \exp(c_i T^{-1})$ and then Taylor expansion of the type $\exp(x) = \sum_{j=0}^\infty x^j/j!$,

$$\begin{aligned} N^{-1}\sum_{i=1}^N \gamma_i J_{f,i,t} &= N^{-1}\sum_{i=1}^N \gamma_i T^{-1/2}\sum_{s=1}^t \rho_i^{t-s} f_s \\ &= N^{-1}\sum_{i=1}^N \gamma_i T^{-1/2}\sum_{s=1}^t\sum_{j=0}^\infty \frac{(c_i T^{-1}(t-s))^j}{j!} f_s \\ &= \overline{\gamma}T^{-1/2}\sum_{s=1}^t E\left(\sum_{j=0}^\infty \frac{(c_i T^{-1}(t-s))^j}{j!}\right) f_s + O_p(N^{-1/2}) \\ &= \overline{\gamma}J_{ft} + O_p(N^{-1/2}), \end{aligned}$$

where $J_{ft} = T^{-1/2}\sum_{s=1}^t E(\rho_i^{t-s})f_s$. The third equality holds, because, defining $\mu_j = E(c_i^j)$ for $j \geq 0$, then $N^{-1}\sum_{i=1}^N \gamma_i c_i^j = \overline{\gamma}\mu_j + N^{-1}\sum_{i=1}^N \gamma_i(c_i^j - \mu_j) = \overline{\gamma}\mu_j + O_p(N^{-1/2})$. It follows that

$$T^{-1/2}\overline{y}_t = N^{-1}\sum_{i=1}^N \gamma_i J_{f,i,t} + \overline{J}_{\epsilon t} = \overline{\gamma}J_{ft} + O_p(N^{-1/2}). \tag{9}$$

Thus, defining $M_J = I_{T-1} - J_{f,-1}(J'_{f,-1}J_{f,-1})^{-1}J'_{f,-1}$, where $J_{f,-1} = (J_{f1}, ..., J_{fT-1})'$ stacks J_{ft-1}, we can show that

$$\begin{aligned} & T^{-2}y'_{i,-1}M_x y_{i,-1} \\ &= T^{-2}(y'_{i,-1}y_{i,-1} - y'_{i,-1}\overline{y}_{-1}(\overline{y}'_{-1}\overline{y}_{-1})^{-1}\overline{y}'_{-1}y_{i,-1}) + O_p(T^{-1/2}) \\ &= T^{-2}(y'_{i,-1}y_{i,-1} - y'_{i,-1}\overline{\gamma}J_{f,-1}(\overline{\gamma}J'_{f,-1}\overline{\gamma}J_{f,-1})^{-1}\overline{\gamma}J'_{f,-1}y_{i,-1}) + O_p(T^{-1/2}) + O_p(N^{-1/2}) \\ &= T^{-2}y'_{i,-1}M_J y_{i,-1} + O_p(T^{-1/2}) + O_p(N^{-1/2}). \end{aligned} \tag{10}$$

Next, consider the second term on the right-hand side of (7). Again, by the definition of M_x,

$$y_{i,-1}'M_xd = y_{i,-1}'d - y_{i,-1}'x(x'x)^{-1}x'd,$$

where $D_T^{-1}d'x = (T^{-1/2}d'\Delta\overline{y}, T^{-1}d'\overline{y}_{-1})$. Consider

$$T^{-1/2}d'\Delta\overline{y} = \frac{1}{\overline{\gamma}}\sum_{t=2}^{T} T^{-1/2}\Delta\overline{y}_t(\overline{\epsilon}_t + \overline{r}_{t-1}(\rho)).$$

By using Taylor expansion of the type $(\rho_i - \rho_0) = O_p(T^{-1})$, where $\rho_0 \in \{1, \rho\}$, we have that $\sqrt{T}\overline{r}_{t-1}(\rho_0) = N^{-1}\sum_{i=1}^{N} T(\rho_i - \rho_0)T^{-1/2}y_{i,t}$ is of the same order of magnitude as $T^{-1/2}y_{i,t}$ (for $T \to \infty$ and a fixed N). In fact, it is not difficult to show that with $\rho_0 = \rho$ and $(\rho_i - \rho) = T^{-1}(c_i - \mu_1) + O_p(T^{-2})$,

$$\begin{aligned}
\sqrt{T}\overline{r}_{t-1}(\rho) &= \sqrt{T}N^{-1}\sum_{i=1}^{N}(\rho_i - \rho)y_{i,t-1} \\
&= T^{-1/2}\sum_{s=1}^{t-1} N^{-1}\sum_{i=1}^{N}(c_i - \mu_1)\rho_i^{t-s}(\gamma_i f_s + \epsilon_{i,s}) + o_p(1) \\
&= \sum_{j=0}^{\infty} T^{-1/2}\sum_{s=1}^{t-1}\frac{(T^{-1}(t-s))^j}{j!}N^{-1}\sum_{i=1}^{N}(c_i - \mu_1)c_i^j(\gamma_i f_s + \epsilon_{i,s}) + o_p(1) \\
&= \sum_{j=0}^{\infty} T^{-1/2}\sum_{s=1}^{t-1}\frac{(T^{-1}(t-s))^j}{j!}N^{-1}\sum_{i=1}^{N}(c_i - \mu_1)c_i^j\gamma_i f_s + O_p(N^{-1/2}), \qquad (11)
\end{aligned}$$

which is $O_p(1)$. By using this, the fact that $\Delta\overline{y}_t = \overline{r}_{t-1}(1) + \overline{\gamma}f_t + \overline{\epsilon}_t$, and the assumed independence of f_t and $\epsilon_{i,t}$, we can show that

$$\begin{aligned}
T^{-1/2}\sum_{t=2}^{T}\Delta\overline{y}_t\overline{\epsilon}_t &= T^{-1}\sum_{t=2}^{T}\sqrt{T}\overline{r}_{t-1}(1)\overline{\epsilon}_t + T^{-1/2}\sum_{t=2}^{T}\overline{\gamma}f_t\overline{\epsilon}_t + \sqrt{T}N^{-1}T^{-1}\sum_{t=2}^{T}(\sqrt{N}\overline{\epsilon}_t)^2 \\
&= O_p(N^{-1/2}) + O_p(\sqrt{T}N^{-1}), \\
T^{-1/2}\sum_{t=2}^{T}\Delta\overline{y}_t\overline{r}_{t-1}(\rho) &= T^{-3/2}\sum_{t=2}^{T}\sqrt{T}\overline{r}_{t-1}(1)\sqrt{T}\overline{r}_{t-1}(\rho) + T^{-1}\sum_{t=2}^{T}\overline{\gamma}f_t\sqrt{T}\overline{r}_{t-1}(\rho) \\
&+ T^{-1}\sum_{t=2}^{T}\overline{\epsilon}_t\sqrt{T}\overline{r}_{t-1}(\rho) \\
&= O_p(T^{-1/2}) + O_p((NT)^{-1/2}).
\end{aligned}$$

Hence, provided that $\max\{N^{-1/2}, T^{-1/2}\} < \sqrt{T}N^{-1}$,

$$T^{-1/2}d'\Delta\overline{y} = \frac{1}{\overline{\gamma}}\sum_{t=2}^{T} T^{-1/2}\Delta\overline{y}_t(\overline{\epsilon}_t + \overline{r}_{t-1}(\rho)) = O_p(\sqrt{T}N^{-1}),$$

which is $o_p(1)$ provided that $\sqrt{T}/N = o(1)$. Making use of this assumption, the previously obtained result for $D_T^{-1}x'xD_T^{-1}$, and then $T^{-1}y'_{i,-1}xD_T^{-1} = (0, T^{-2}y'_{i,-1}\overline{y}_{-1}) + O_p(T^{-1/2})$, we obtain

$$\begin{aligned}
&T^{-1}y'_{i,-1}x(x'x)^{-1}x'd \\
&= T^{-1}y'_{i,-1}xD_T^{-1}(D_T^{-1}x'xD_T^{-1})^{-1}D_T^{-1}x'd \\
&= (0, T^{-2}y'_{i,-1}\overline{y}_{-1}) \begin{bmatrix} T^{-1}(\Delta\overline{y})'\Delta\overline{y} & 0 \\ 0 & T^{-2}\overline{y}'_{-1}\overline{y}_{-1} \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ T^{-1}\overline{y}'_{-1}d \end{bmatrix} + o_p(1) \\
&= T^{-2}y'_{i,-1}\overline{y}_{-1}(T^{-2}\overline{y}'_{-1}\overline{y}_{-1})^{-1}T^{-1}\overline{y}'_{-1}d + o_p(1).
\end{aligned}$$

Moreover, since

$$T^{-1}y'_{i,-1}M_J\overline{\epsilon} = T^{-1}y'_{i,-1}\overline{\epsilon} - T^{-2}y_{i,-1}J_{f,-1}(T^{-2}J'_{f,-1}J_{f,-1})^{-1}T^{-1}J'_{f,-1}\overline{\epsilon} = O_p(N^{-1/2}),$$

we can show that $T^{-1}y'_{i,-1}M_Jd = T^{-1}y'_{i,-1}M_J\overline{r}_{-1}(\rho)/\overline{\gamma} + o_p(1)$, and therefore

$$\begin{aligned}
T^{-1}y'_{i,-1}M_xd &= T^{-1}y'_{i,-1}d - T^{-1}y'_{i,-1}xD_T^{-1}(D_T^{-1}x'xD_T^{-1})^{-1}D_T^{-1}x'd \\
&= T^{-1}y'_{i,-1}d - T^{-2}y'_{i,-1}\overline{y}_{-1}(T^{-2}\overline{y}'_{-1}\overline{y}_{-1})^{-1}T^{-1}\overline{y}'_{-1}d + o_p(1) \\
&= T^{-1}y'_{i,-1}M_Jd + o_p(1) \\
&= \frac{1}{\overline{\gamma}}T^{-1}y'_{i,-1}M_J\overline{r}_{-1}(\rho) + o_p(1). \qquad (12)
\end{aligned}$$

It remains to consider $y'_{i,-1}M_x\epsilon_i$, the third term on the right-hand side of (7). We have

$$\begin{aligned}
T^{-1/2}\epsilon'_i\Delta\overline{y} &= T^{-1/2}\sum_{t=2}^{T}\overline{r}_{t-1}(1)\epsilon_{i,t} + T^{-1/2}\sum_{t=2}^{T}\overline{\gamma}f_t\epsilon_{i,t} + (NT)^{-1/2}\sum_{t=2}^{T}\sqrt{N}\overline{\epsilon}_t\epsilon_{i,t} \\
&= O_p(1) + O_p(\sqrt{T}N^{-1/2}),
\end{aligned}$$

which is $O_p(1)$ as long as $T/N = O(1)$, which in turn implies $\sqrt{T}/N = o(1)$. Thus, provided that $T/N = O(1)$,

$$\begin{aligned}
&T^{-1}y'_{i,-1}x(x'x)^{-1}x'\epsilon_i \\
&= T^{-1}y'_{i,-1}xD_T^{-1}(D_T^{-1}x'xD_T^{-1})^{-1}D_T^{-1}x'\epsilon_i \\
&= (0, T^{-2}y'_{i,-1}\overline{y}_{-1}) \begin{bmatrix} T^{-1}(\Delta\overline{y})'\Delta\overline{y} & 0 \\ 0 & T^{-2}\overline{y}'_{-1}\overline{y}_{-1} \end{bmatrix}^{-1} \begin{bmatrix} T^{-1/2}(\Delta\overline{y})'\epsilon_i \\ T^{-1}\overline{y}'_{-1}\epsilon_i \end{bmatrix} + o_p(1) \\
&= T^{-2}y'_{i,-1}\overline{y}_{-1}(T^{-2}\overline{y}'_{-1}\overline{y}_{-1})^{-1}T^{-1}\overline{y}'_{-1}\epsilon_i + o_p(1),
\end{aligned}$$

from which we obtain

$$\begin{aligned}
T^{-1}y'_{i,-1}M_x\epsilon_i &= T^{-1}y'_{i,-1}\epsilon_i - T^{-1}y'_{i,-1}xD_T^{-1}(D_T^{-1}x'xD_T^{-1})^{-1}D_T^{-1}x'\epsilon_i \\
&= T^{-1}(y'_{i,-1}\epsilon_i - y'_{i,-1}\overline{y}_{-1}(\overline{y}'_{-1}\overline{y}_{-1})^{-1}\overline{y}'_{-1}\epsilon_i) + o_p(1) \\
&= T^{-1}y'_{i,-1}M_J\epsilon_i + o_p(1). \qquad (13)
\end{aligned}$$

Hence, by inserting (10)–(13) into (7),

$$\begin{aligned} & T^{-1}y'_{i,-1}M_x\Delta y_i \\ &= T^{-1}(\rho_i-1)y'_{i,-1}M_x y_{i,-1} - \gamma_i T^{-1}y'_{i,-1}M_x d + T^{-1}y'_{i,-1}M_x\epsilon_i \\ &= T^{-1}(\rho_i-1)y'_{i,-1}M_J y_{i,-1} + T^{-1}y'_{i,-1}M_J(\epsilon_i - \gamma_i\overline{r}_{-1}(\rho)/\overline{\gamma}) + o_p(1). \quad (14) \end{aligned}$$

Note that $J'_{f,i,-1}M_J = J'_{f,i,-1} - J'_{f,i,-1}J_{f,-1}(J'_{f,-1}J_{f,-1})^{-1}J'_{f,-1}$, where

$$T^{-1}J'_{f,i,-1}J_{f,-1} = T^{-1}\sum_{t=2}^{T} J_{f,i,t-1}J_{ft-1} = T^{-2}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\sum_{k=1}^{t-1}\rho_i^{t-s}\rho^{t-k}f_s f_k,$$

which is generally different from $T^{-1}J'_{f,-1}J_{f,-1}$, and so $J'_{f,i,-1}M_J \neq 0$. The exception is when $c_1 = ... = c_N = c$, in which case $y'_{i,-1}M_J = (\gamma_i J_{f,-1} + J_{\epsilon,i,-1})'M_J = J'_{\epsilon,i,-1}M_J$, suggesting that

$$T^{-1}y'_{i,-1}M_x\Delta y_i = T^{-1}(\rho_i-1)J'_{\epsilon,i,-1}M_J J_{\epsilon,i,-1} + T^{-1}J'_{\epsilon,i,-1}M_J(\epsilon_i - \gamma_i\overline{r}_{-1}(\rho)/\overline{\gamma}) + o_p(1).$$

Moreover, since in this case, $\overline{r}_{t-1}(\rho) = N^{-1}\sum_{i=1}^{N}(\rho_i-\rho)y_{i,t-1} = 0$, and therefore

$$T^{-1}y'_{i,-1}M_x\Delta y_i = T^{-1}(\rho_i-1)J'_{\epsilon,i,-1}M_J J_{\epsilon,i,-1} + T^{-1}J'_{\epsilon,i,-1}M_J\epsilon_i + o_p(1).$$

However, this is only true if c_i is equal for all i. If c_i is not equal, then the appropriate expansion for $T^{-1}y'_{i,-1}M_x\Delta y_i$ is again given by (14).

Let us now consider $\hat{\sigma}^2_{\epsilon,i}$. Again, since $T^{-1/2}(\Delta y_i)'xD_T^{-1} = (T^{-1}(\Delta y_i)'\Delta\overline{y}, 0) + O_p(T^{-1/2})$ and $D_T^{-1}x'xD_T^{-1}$ is asymptotically diagonal (provided that $\sqrt{T}/N = o(1)$),

$$\begin{aligned} \hat{\sigma}^2_{\epsilon,i} &= T^{-1}(\Delta y_i)'M_x\Delta y_i \\ &= T^{-1}(\Delta y_i)'\Delta y_i - T^{-1/2}(\Delta y_i)'xD_T^{-1}(D_T^{-1}x'xD_T^{-1})^{-1}T^{-1/2}D_T^{-1}x'\Delta y_i \\ &= T^{-1}((\Delta y_i)'\Delta y_i - (\Delta y_i)'\Delta\overline{y}((\Delta\overline{y})'\Delta\overline{y})^{-1}(\Delta\overline{y})'\Delta y_i) + o_p(1) \\ &= T^{-1}(\Delta y_i)'M_f\Delta y_i + o_p(1), \end{aligned}$$

where $M_f = I_{T-1} - f(f'f)^{-1}f'$ and $f = (f_2, ..., f_T)'$. The last equality holds because

$$\begin{aligned} T^{-1}(\Delta y_i)'\Delta\overline{y} &= (TN)^{-1}\sum_{t=2}^{T}\sum_{j=1}^{N}(\rho_j-1)y_{j,t-1}\Delta y_{i,t} + \overline{\gamma}T^{-1}\sum_{t=2}^{T}f_t\Delta y_{i,t} + T^{-1}\sum_{t=2}^{T}\overline{\epsilon}_t\Delta y_{i,t} \\ &= \overline{\gamma}T^{-1}\sum_{t=2}^{T}f_t\Delta y_{i,t} + O_p(T^{-1/2}) + O_p(N^{-1/2}), \end{aligned}$$

and by the same arguments,

$$T^{-1}(\Delta\overline{y})'\Delta\overline{y} = \overline{\gamma}^2 T^{-1}\sum_{t=2}^{T}f_t^2 + O_p(T^{-1}) + O_p(N^{-1}) + O_p((NT)^{-1/2}).$$

Moreover, since by construction, $fM_f = 0$ and $\epsilon_i' f = O_p(\sqrt{T})$, we obtain

$$
\begin{aligned}
\hat{\sigma}^2_{\epsilon,i} &= T^{-1}(\Delta y_i)' M_f \Delta y_i + o_p(1) \\
&= T^{-1}((\rho_i - 1) y_{i,-1} + \epsilon_i)' M_f ((\rho_i - 1) y_{i,-1} + \epsilon_i) + o_p(1) \\
&= T^{-1}\epsilon_i' M_f \epsilon_i + o_p(1) = T^{-1}\epsilon_i'\epsilon_i + o_p(1) = \sigma^2_{\epsilon,i} + o_p(1). \qquad (15)
\end{aligned}
$$

We are now ready to consider CADF_i. By direct insertion of (10), (14) and (15), and then Taylor expansion of the type $(\rho_i - 1) = T^{-1}c_i + O(T^{-2})$,

$$
\begin{aligned}
\text{CADF}_i &= \frac{T^{-1} y_{i,-1}' M_x (\Delta y_i)}{\hat{\sigma}_{\epsilon,i}\sqrt{T^{-2} y_{i,-1}' M_x y_{i,-1}}} \\
&= c_i \frac{\sqrt{T^{-2} y_{i,-1}' M_J y_{i,-1}}}{\sigma_{\epsilon,i}} + \frac{T^{-1} y_{i,-1}' M_J (\epsilon_i - \gamma_i \overline{r}_{t-1}(\rho)/\overline{\gamma})}{\sigma_{\epsilon,i}\sqrt{T^{-2} y_{i,-1}' M_J y_{i,-1}}} + o_p(1), \qquad (16)
\end{aligned}
$$

which holds provided that $T/N = O(1)$. As for the limiting distribution of this test statistic, given (8), (9) and (11), it is not difficult to show that, using $\lfloor x \rfloor$ to denote the integer part of x,

$$
\begin{aligned}
T^{-1/2} y_{i,\lfloor rT \rfloor} &= \gamma_i J_{if,\lfloor rT \rfloor} + J_{\epsilon,i,\lfloor rT \rfloor} \to_w \gamma_i J_{f,i}(r) + \sigma_{\epsilon,i} J_{\epsilon,i}(r) = \sigma_{\epsilon,i} J_{y,i}(r), \qquad (17) \\
T^{-1/2} \overline{y}_{\lfloor rT \rfloor} &= \overline{\gamma} J_{f,\lfloor rT \rfloor} + o_p(1) \to_w \gamma J_f(r), \qquad (18) \\
\sqrt{T}\overline{r}_{\lfloor rT \rfloor}(\rho) &= \sum_{j=0}^{\infty} T^{-1/2} \sum_{s=1}^{\lfloor rT \rfloor} \frac{(T^{-1}(\lfloor rT \rfloor - s))^j}{j!} N^{-1} \sum_{i=1}^{N} (c_i - \mu_1) c_i^j \gamma_i f_s + o_p(1) \\
&\to_w \gamma \sum_{j=0}^{\infty} (\mu_{j+1} - \mu_1 \mu_j) \int_0^r \frac{(r-v)^j}{j!} dW_f(v) = \gamma K_f(r) \qquad (19)
\end{aligned}
$$

as $N,\ T \to \infty$, where $J_{y,i}(r) = \beta_i J_{f,i}(r) + J_{\epsilon,i}(r)$, $\beta_i = \gamma_i/\sigma_{\epsilon,i}$, $J_{f,i}(r) = \int_0^r \exp(c_i(r-s)) dW_f(s)$, $J_{\epsilon,i}(r) = \int_0^r \exp(c_i(r-s)) dW_{\epsilon,i}(s)$, $\gamma = \lim_{N\to\infty} \overline{\gamma}$, and an obvious definition of $K_f(r)$. Also, by using $\exp(x) = \sum_{j=0}^{\infty} x^j/j!$, we have

$$
J_f(r) = \int_0^r E(\exp(c_i(r-s))) dW_f(s) = \sum_{j=0}^{\infty} \mu_j \int_0^r \frac{(r-s)^j}{j!} dW_f(s).
$$

Let Q_{J_f} be a projection residual operator such that

$$
Q_{J_f} J_{y,i}(r) = J_{y,i}(r) - \int_0^1 J_{y,i}(v) J_f(v) dv \left(\int_0^1 J_f(v)^2 dv \right)^{-1} J_f(r)
$$

is the residual from a continuous time regression of $J_{y,i}(r)$ onto $J_f(r)$. It follows that, by the continuous mapping theorem,

$$\sigma_{\epsilon,i}^{-2}T^{-2}y_{i,-1}'M_J y_{i,-1} \rightarrow_w \int_0^1 (Q_{J_f}J_{y,i}(r))^2 dr, \quad (20)$$

$$\sigma_{\epsilon,i}^{-2}T^{-1}y_{i,-1}'M_J(\epsilon_i - \gamma_i \overline{r}_{t-1}(\rho)/\overline{\gamma}) \rightarrow_w \int_0^1 Q_{J_f}J_{y,i}(r)(dW_{\epsilon,i}(r) - \beta_i Q_{J_f}K_f(r)dr), \quad (21)$$

which holds as N, $T \to \infty$ with $T/N = O(1)$. The limiting distribution of CADF$_i$ is therefore given by

$$\text{CADF}_i \rightarrow_w c_i\sqrt{\int_0^1 (Q_{J_f}J_{y,i}(r))^2 dr} + \frac{\int_0^1 Q_{J_f}J_{y,i}(r)(dW_{\epsilon,i}(r) - \beta_i Q_{J_f}K_f(r)dr)}{\sqrt{\int_0^1 (Q_{J_f}J_{y,i}(r))^2 dr}},$$

as required for the proof. ■

Proof of Theorem 2.

Let

$$g_i(v) = \sqrt{\beta_i^2 v_1 + v_2 + 2\beta_i v_3 - \frac{(\beta_i v_4 + v_5)^2}{v_6}},$$

where $v = (v_1, ..., v_6)'$ with

$$\begin{aligned} v_1 &= \int_0^1 (J_{f,i}(r))^2 dr, \\ v_2 &= \int_0^1 (J_{\epsilon,i}(r))^2 dr, \\ v_3 &= \int_0^1 J_{f,i}(r)J_{\epsilon,i}(r)dr, \\ v_4 &= \int_0^1 J_{f,i}(r)J_f(r)dr, \\ v_5 &= \int_0^1 J_{\epsilon,i}(r)J_f(r)dr, \\ v_6 &= \int_0^1 (J_f(r))^2 dr. \end{aligned}$$

Let $\alpha = E(v) = (\alpha_1, ..., \alpha_6)'$. Approximating $g_i(v)$ by a second-order Taylor series (around

$v = \alpha$), and then taking expectations we get

$$
\begin{aligned}
E[g_i(v)] &\sim g_i(\alpha) + \frac{1}{2}E\left[\left(\sum_{k=1}^{6}(v_k - \alpha_k)\frac{\partial}{\partial v_k}\right)^2 g(v)|_{v=\alpha}\right] \\
&= \theta_i - \frac{\beta_i^4}{4}\frac{\omega_1^2}{2\theta_i^3} - \frac{1}{4}\frac{\omega_2^2}{2\theta_i^3} + \frac{-\beta_i^2\omega_3^2}{2\theta_i^3} - \frac{\beta_i^2(\mu_2 + \beta_i(2\alpha_3 + \beta_i\alpha_1))}{\alpha_6}\frac{\omega_4^2}{2\theta_i^3} \\
&- \frac{\alpha_2 + 2\beta_i\alpha_3 + \beta_i^2\alpha_1}{\alpha_6}\frac{\omega_5^2}{2\theta_i^3} \\
&+ \frac{(\alpha_5 + \beta_i\alpha_4)^2(3\mu_5^2 - 4\alpha_2\alpha_6 + 6\alpha_4\alpha_5\beta_i + \beta_i(-8\alpha_3\alpha_6 + 3\alpha_4^2\beta_i - 4\alpha_1\alpha_6\beta_i))}{4\alpha_6^4}\frac{\omega_6^2}{2\theta_i^3} \\
&- \frac{\beta_i^2}{4}\frac{\omega_{12}}{\theta_i^3} - \frac{\beta_i^3}{2}\frac{\omega_{13}}{\theta_i^3} + \frac{\beta_i^3(\mu_5 + \beta_i\alpha_4)}{2\alpha_6}\frac{\omega_{14}}{\theta_i^3} + \frac{\beta_i^2(\alpha_5 + \beta_i\mu_4)}{2\alpha_6}\frac{\omega_{15}}{\theta_i^3} \\
&- \frac{\beta_i^2(\alpha_5 + \beta_i\alpha_4)^2}{4\mu_6^2}\frac{\omega_{16}}{\theta_i^3} - \frac{\beta_i}{2}\frac{\omega_{23}}{\theta_i^3} \\
&+ \frac{\beta_i(\alpha_5 + \beta_i\alpha_4)}{2\alpha_6}\frac{\omega_{24}}{\theta_i^3} + \frac{\alpha_5 + \beta_i\alpha_4}{2\mu_6}\frac{\omega_{25}}{\theta_i^3} - \frac{(\alpha_5 + \beta_i\alpha_4)^2}{4\alpha_6^2}\frac{\omega_{26}}{g_\mu^3} \\
&+ \frac{\beta_i^2(\alpha_5 + \beta_i\alpha_4)}{\alpha_6}\frac{\omega_{34}}{\theta_i^3} + \frac{\beta_i(\mu_5 + \beta_i\alpha_4)}{\alpha_6}\frac{\omega_{35}}{\theta_i^3} - \frac{\beta_i(\alpha_5 + \beta_i\mu_4)^2}{2\alpha_6^2}\frac{\omega_{36}}{\theta_i^3} \\
&- \frac{\beta_i(\alpha_5 + \alpha_4\beta_i)(\mu_5^2 - 2\alpha_2\alpha_6 + 2\alpha_4\alpha_5\beta_i + \beta_i(-\alpha_3\alpha_6 + \alpha_4^2\beta_i - 2\alpha_1\alpha_6\beta_i))}{2\alpha_6^2}\frac{\omega_{46}}{\theta_i^3} \\
&+ \left(\omega_{45}\frac{\partial^2}{\partial\alpha_4\partial\alpha_5} + \omega_{56}\frac{\partial^2}{\partial\alpha_5\partial\alpha_6}\right)g(v)|_{v=\alpha}.
\end{aligned}
\tag{22}
$$

In what remains we calculate α_j, $\omega_j^2 = \text{var}(v_j)$ and $\omega_{jk} = \text{cov}(v_j, v_k)$ for $j, k = 1, ..., 6$.

Consider v_1. Clearly, letting $e^x = \exp(x)$,

$$
\begin{aligned}
\alpha_1 &= E\int_0^1 (J_{f,i}(r))^2 dr = E\left(\int_0^1 E[(J_{f,i}(r))^2|c_i]dr\right) = E\left(\int_0^1\int_0^r e^{2c_i(r-u)}dudr\right) \\
&= \int_0^1\int_0^r M_c(2r - 2u)dudr.
\end{aligned}
\tag{23}
$$

ω_1^2 can be written as

$$
\begin{aligned}
\omega_1^2 &= E[\text{var}(v_1|c_i)] + \text{var}[E(v_1|c_i)] \\
&= E\left(\int_0^1\int_0^1 \text{cov}[(J_{f,i}(r))^2, (J_{f,i}(s))^2|c_i]drds\right) + \text{cov}[E(v_1|c_i), E(v_1|c_i)] \\
&= \int_0^1\int_0^1 (E[(J_{f,i}(r))^2(J_{f,i}(s))^2] - E[(J_{f,i}(r))^2]E[(J_{f,i}(s))^2])drds \\
&+ \text{var}\left(\int_0^1\int_0^r e^{2c_i(r-u)}dudr\right).
\end{aligned}
$$

Conditional on c_i, $J_{f,i}(r)$ and $J_{f,i}(v)$ are jointly normally distributed. Specifically,

$$
\begin{bmatrix} J_{f,i}(r) \\ J_{f,i}(s) \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \int_0^r e^{2c_i(r-u)}du & \int_0^{r\wedge s} e^{c_i(r+s-2u)}du \\ . & \int_0^s e^{2c_i(s-u)}du \end{bmatrix}\right),
$$

where $r \wedge s = \min\{r, s\}$. It follows that

$$\begin{aligned}
E[(J_{f,i}(r))^2(J_{f,i}(s))^2|c_i] &= 2\left(\int_0^{r\wedge s} e^{c_i(r+s-2u)}du\right)^2 + \left(\int_0^r e^{2c_i(r-u)}du\right)\left(\int_0^s e^{2c_i(s-u)}du\right) \\
&= 2\int_0^{r\wedge s}\int_0^{r\wedge s} e^{c_i(2r+2s-2u-2s)}duds + \int_0^r\int_0^s e^{c_i(2r+2s-2u-2s)}duds,
\end{aligned}$$

giving

$$\begin{aligned}
\omega_1^2 &= 2\int_0^1\int_0^1\int_0^{r\wedge s}\int_0^{r\wedge v} M_c(2r+2v-2u-2s)dudsdrdv \\
&+ \int_0^1\int_0^1\int_0^r\int_0^v M_c(2r+2v-2u-2s)dsdudrdv \\
&- \int_0^1\int_0^1\int_0^r\int_0^v M_c(2r+2v-2s-2x)dsdxdrdv + \mathrm{var}\left(\int_0^1\int_0^r e^{2c_i(r-u)}dudr\right).
\end{aligned}$$

The last term on the right is

$$\begin{aligned}
&\mathrm{var}\left(\int_0^1\int_0^r e^{2c_i(r-u)}dudr\right) \\
&= E\left[\left(\int_0^1\int_0^r e^{2c_i(r-u)}dudr\right)^2\right] - \left[E\left(\int_0^1\int_0^r e^{2c_i(r-u)}dudr\right)\right]^2 \\
&= \int_0^1\int_0^1\int_0^r\int_0^x M_c\left(2r-2u+2x-2s\right)dsdudxdr - \left(\int_0^1\int_0^r M_c\left(2r-2u\right)dudr\right)^2.
\end{aligned}$$

Putting all these terms together and simplifying we get

$$\begin{aligned}
\omega_1^2 &= 2\int_0^1\int_0^1\int_0^{r\wedge v}\int_0^{r\wedge v} M_c\left(2r+2v-2u-2s\right)dudsdrdv \\
&+ \int_0^1\int_0^1\int_0^r\int_0^v M_c(2r+2v-2u-2s)dsdudrdv \\
&- \left(\int_0^1\int_0^r M_c(2r-2u)dudr\right)^2. \qquad (24)
\end{aligned}$$

ω_{12} can be computed as

$$\begin{aligned}
\omega_{12} &= E[\mathrm{cov}(v_1, v_2|c_i)] + \mathrm{cov}[E(v_1|c_i), E(v_2|c_i)] \\
&= \mathrm{cov}[E(v_1|c_i), E(v_2|c_i)] \\
&= \mathrm{cov}\left(\int_0^1\int_0^r e^{2c_i(r-u)}dudr, \int_0^1\int_0^x e^{2c_i(x-s)}dsdx\right) \\
&= \int_0^1\int_0^1\int_0^r\int_0^x M_c\left(2r-2u+2x-2s\right)dsdudxdr \\
&- \left(\int_0^1\int_0^r M_c\left(2r-2u\right)dudr\right)^2. \qquad (25)
\end{aligned}$$

It is easy to see that $\omega_{13} = 0$. For ω_{14},

$$\begin{aligned}
\omega_{14} &= \text{cov}(v_1, v_4) \\
&= E\left[\text{cov}\left(v_1, v_4|c_i\right)\right] + \text{cov}\left[E\left(v_1|c_i\right), E\left(v_4|c_i\right)\right] \\
&= E\int_0^1\int_0^1 \text{cov}[J_{fi}^2(v), J_{fi}(r)J_f(r)|c_i]drdv \\
&+ \text{cov}\left(\int_0^1\int_0^r e^{2c_i(r-v)}dvdr, \int_0^1\int_0^r e^{c_i(r-v)}M_c\left(r-v\right)dvdr|c_i\right) \\
&= T_1 + T_2,
\end{aligned}$$

with obvious definitions of T_1 and T_2. T_1 can be written as

$$T_1 = E\int_0^1\int_0^1 E[J_{fi}^3(v)J_f(r)|c_i]drdv - E\int_0^1\int_0^1 E(J_{fi}^2(v)|c_i)E(J_{fi}(r)|c_i)E(J_f(r)|c_i)drdv,$$

where $[J_{fi}(r), J_f(v)]'|c_i \sim N(0, \Sigma)$ with

$$\Sigma = \begin{bmatrix} \int_0^v e^{2c_i(v-u)}du & \int_0^{r\wedge v} e^{c_i(v-u)}M_c\left(r-u\right)du \\ \cdot & \int_0^r M_c\left(r-u\right)^2 du \end{bmatrix} = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \cdot & \Sigma_{22} \end{bmatrix}.$$

Note how the dependence on i has been suppressed in Σ. From now on Σ will be used generically to denote the (conditional) covariance matrix between random processes on $[0,1]$. Now, since $E[(J_{fi}(v))^3 J_f(r)] = 3\Sigma_{11}\Sigma_{12}$, we have

$$\begin{aligned}
T_1 &= 3E\int_0^1\int_0^1\left(\int_0^v e^{2c_i(v-u)}du\right)\left(\int_0^{r\wedge v} e^{c_i(v-x)}M_c\left(r-x\right)dx\right)drdv \\
&- E\int_0^1\int_0^1\left(\int_0^v e^{2c_i(v-u)}du\right)\left(\int_0^{r\wedge v} e^{c_i(v-x)}M_c\left(r-x\right)dx\right)drdv \\
&= 2E\int_0^1\int_0^1\int_0^v\int_0^{r\wedge v} e^{c_i(2v-2u+v-x)}M_c\left(r-x\right)dxdudrdv \\
&= 2\int_0^1\int_0^1\int_0^v\int_0^{r\wedge v} M\left(3v-2u-x\right)M_c\left(r-x\right)dxdudrdv,
\end{aligned}$$

and

$$\begin{aligned}
T_2 &= \text{cov}\left(\int_0^1\int_0^r e^{2c_i(r-v)}dvdr, \int_0^1\int_0^r e^{c_i(r-v)}M_c\left(r-v\right)dvdr|c_i\right) \\
&= E\left(\int_0^1\int_0^x\int_0^1\int_0^r e^{2c_i(r-v)}e^{c_i(x-u)}M_c\left(x-u\right)dvdrdudx|c_i\right) \\
&- E\left(\int_0^1\int_0^r e^{2c_i(r-v)}dvdr|c_i\right)E\left(\int_0^1\int_0^r e^{c_i(r-v)}M_c\left(r-v\right)dvdr|c_i\right) \\
&= \int_0^1\int_0^x\int_0^1\int_0^r M_c\left(2r-2v+x-u\right)M_c\left(x-u\right)dvdrdudx \\
&- \int_0^1\int_0^r M_c\left(2r-2v\right)dvdr\int_0^1\int_0^r M_c^2\left(r-v\right)dvdr.
\end{aligned}$$

Putting everything together,

$$\begin{aligned}
\omega_{14} &= 2\int_0^1\int_0^1\int_0^v\int_0^{r\wedge v} M\left(3v-2u-x\right)M_c\left(r-x\right)dxdudrdv \\
&+ \int_0^1\int_0^x\int_0^1\int_0^r M_c\left(2r-2v+x-u\right)M_c\left(x-u\right)dvdrdudx \\
&- \int_0^1\int_0^r M_c\left(2r-2v\right)dvdr\int_0^1\int_0^r M_c\left(r-v\right)^2 dvdr. \qquad (26)
\end{aligned}$$

It is easily seen that $\omega_{15}=0$. For ω_{16}, we have

$$\omega_{16} = \text{cov}\left(v_1, v_6\right) = E\left[\text{cov}\left(v_1, v_6|c_i\right)\right] + \text{cov}\left[E\left(v_1|c_i\right), E\left(v_6|c_i\right)\right].$$

where

$$\begin{aligned}
\text{cov}\left(v_1, v_6|c_i\right) &= \text{cov}\left(\int_0^1 (J_{fi}(r))^2 dr, \int_0^1 (J_f(r))^2 dr|c_i\right) \\
&= \int_0^1\int_0^1 (E[(J_{fi}(r))^2(J_f(v))^2] - E[(J_{fi}(r))^2]E[(J_f(v))^2])drdv.
\end{aligned}$$

Since $[J_{fi}(r), J_f(v)]'|c_i$ is normal with covariance matrix

$$\Sigma = \begin{bmatrix} \int_0^r e^{2c_i(r-u)}du & \int_0^{r\wedge v} e^{c_i(r-u)}M_c\left(v-u\right)du \\ . & \int_0^v M_c^2\left(v-u\right)du \end{bmatrix},$$

we can show that $E[(J_{fi}(r))^2(J_f(v))^2] = 2\Sigma_{12}^2 + \Sigma_{11}\Sigma_{22}$, where Σ_{12}, Σ_{11} and Σ_{22} are the elements of Σ. This result in turn implies

$$\begin{aligned}
&E[\text{cov}\left(v_1, v_6|c_i\right)] \\
&= \int_0^1\int_0^1 E\left[2\left(\int_0^{r\wedge v} e^{c_i(r-u)}M_c\left(v-u\right)du\right)^2 + \int_0^r e^{2c_i(r-u)}du\int_0^v M_c\left(v-u\right)^2 du\right]drdv \\
&- E\left(\int_0^1\int_0^r e^{2c_i(r-u)}dudr\int_0^1\int_0^v M_c\left(v-u\right)^2 dudv\right), \\
&= R_1 + R_2 - R_3,
\end{aligned}$$

where

$$\begin{aligned}
R_1 &= 2\int_0^1\int_0^1 E\left[\left(\int_0^{r\wedge v} e^{c_i(r-u)}M_c\left(v-u\right)du\right)^2\right]drdv \\
&= 2\int_0^1\int_0^1\int_0^{r\wedge v}\int_0^{r\wedge v} E[e^{c_i(2r-u-x)}M_c\left(v-u\right)M_c\left(v-x\right)]dudxdrdv \\
&= 2\int_0^1\int_0^1\int_0^{r\wedge v}\int_0^{r\wedge v} M_c\left(2r-u-x\right)M_c\left(v-u\right)M_c\left(v-x\right)dudxdrdv,
\end{aligned}$$

$$\begin{aligned}
R_2 &= E\left(\int_0^1\int_0^r e^{2c_i(r-u)}dudr\int_0^1\int_0^v M_c^2\left(v-u\right)dudv\right) \\
&= \int_0^1\int_0^r M_c\left(2r-2u\right)dudr\int_0^1\int_0^v M_c\left(v-u\right)^2 dudv,
\end{aligned}$$

and

$$R_3 = \int_0^1 \int_0^r M_c\left(2r - 2u\right) dudr \int_0^1 \int_0^v M_c\left(v - u\right)^2 dudv.$$

Finally,

$$\text{cov}[E\left(v_1|c_i\right), E\left(v_6|c_i\right)] = \text{cov}\left(\int_0^1 \int_0^r e^{2c_i(r-u)}dudr, \int_0^1 \int_0^r M_c^2\left(r-u\right) dudr\right) = 0.$$

Putting everything together we have

$$\begin{aligned}
\omega_{16} &= 2\int_0^1 \int_0^1 \int_0^{r\wedge v} \int_0^{r\wedge v} M_c\left(2r - u - x\right) M_c\left(v - u\right) M_c\left(v - x\right) dudxdrdv \\
&+ \int_0^1 \int_0^1 \int_0^v \int_0^r M_c\left(2r - 2x\right) M_c^2\left(v - u\right) dxdudrdv \\
&- \int_0^1 \int_0^r M_c\left(2r - 2u\right) dudr \int_0^1 \int_0^v M_c^2\left(v - u\right) dudv \\
&= 2\int_0^1 \int_0^1 \int_0^{r\wedge v} \int_0^{r\wedge v} M_c\left(2r - u - x\right) M_c\left(v - u\right) M_c\left(v - x\right) dudxdrdv \qquad (27)
\end{aligned}$$

Consider v_2. It is easy to show that $\alpha_2 = \alpha_1$, $\omega_2^2 = \omega_1^2$ and $\omega_{2j} = 0$ for $j = 3, ..., 6$. We therefore move on to v_3. Clearly, $\alpha_3 = 0$. As for ω_3^2, we have

$$\begin{aligned}
&\text{var}\left(v_3\right) \\
&= E[\text{var}\left(v_3|c_i\right)] + \text{var}[E\left(v_3|c_i\right)] \\
&= E\left[\text{var}\left(\int_0^1 J_{fi}(r)J_{\epsilon i}(r)dr|c_i\right)\right] \\
&= E\left(\int_0^1 \int_0^1 \text{cov}\left[J_{fi}(r)J_{\epsilon i}(r), J_{fi}(u)J_{\epsilon i}(u)|c_i\right] drdu\right) \\
&= E\left(\int_0^1 \int_0^1 \left[E\left(J_{fi}(r)J_{\epsilon i}(r)J_{fi}(u)J_{\epsilon i}(u)|c_i\right) - E\left(J_{fi}(r)J_{\epsilon i}(r)|c_i\right) E\left(J_{fi}(u)J_{\epsilon i}(u)|c_i\right)\right] drdu\right) \\
&= E\left(\int_0^1 \int_0^1 E[J_{fi}(r)J_{fi}(u)|c_i]E[J_{\epsilon i}(r)J_{\epsilon i}(u)|c_i]drdu\right) \\
&= E\left[\int_0^1 \int_0^1 \left(\int_0^{r\wedge u} e^{c_i(r+u-2v)}dv\right)^2 drdu\right] \\
&= E\left(\int_0^1 \int_0^1 \int_0^{r\wedge u} \int_0^{r\wedge u} e^{c_i(2r+2u-2v-2x)}dxdvdrdu\right) \\
&= \int_0^1 \int_0^1 \int_0^{r\wedge u} \int_0^{r\wedge u} M_c\left(2r + 2u - 2v - 2x\right) dxdvdrdu. \qquad (28)
\end{aligned}$$

It is easily established that $\omega_{34} = \omega_{36} = 0$. For ω_{35}, we have

$$
\begin{aligned}
& \operatorname{cov}(v_3, v_5) \\
= \; & E\left[\operatorname{cov}(v_3, v_5|c_i)\right] + \operatorname{cov}\left[E(v_3|c_i), E(v_5|c_i)\right] \\
= \; & E\left[\operatorname{cov}\left(\int_0^1 J_{fi}(r) J_{\epsilon i}(r) dr, \int_0^1 J_{\epsilon i}(u) J_f(u) du | c_i\right)\right] \\
+ \; & \operatorname{cov}\left[E\left(\int_0^1 J_{fi}(r) J_{\epsilon i}(r) dr | c_i\right), E\left(\int_0^1 J_{\epsilon i}(r) J_f(r) dr | c_i\right)\right] \\
= \; & E\left(\int_0^1 \int_0^1 (E[J_{fi}(r) J_f(u)|c_i] E[J_{\epsilon i}(r) J_{\epsilon i}(u)|c_i] - E[J_{fi}(r) J_{\epsilon i}(r)|c_i] E[J_{\epsilon i}(u) J_f(u)|c_i]) du dr\right) \\
= \; & E\left(\int_0^1 \int_0^1 E[J_{fi}(r) J_f(u)|c_i] E[J_{\epsilon i}(r) J_{\epsilon i}(u)|c_i] du dr\right) \\
= \; & E\left[\int_0^1 \int_0^1 \left(\int_0^{r \wedge u} e^{c_i(r-v)} M(u-v) dv\right) \left(\int_0^{r \wedge u} e^{c_i(r+u-2v)} dv\right) du dr\right] \\
= \; & E\left(\int_0^1 \int_0^1 \int_0^{r \wedge u} \int_0^{r \wedge u} e^{c_i(2r-v+u-2x)} M(u-v) dv dx du dr\right) \\
= \; & \int_0^1 \int_0^1 \int_0^{r \wedge u} \int_0^{r \wedge u} M_c(2r+u-v-2x) M(u-v) dv dx du dr. \qquad (29)
\end{aligned}
$$

We now turn to v_4. Note that $\alpha_4 = \int_0^1 \int_0^r M_c(r-v)^2 dv dr$. ω_4^2 can be expanded as

$$
\begin{aligned}
& \operatorname{var}(v_4) \\
= \; & E\left[\operatorname{var}(v_4|c_i)\right] + \operatorname{var}\left[E(v_4|c_i)\right] \\
= \; & E\left[\operatorname{var}\left(\int_0^1 J_{fi}(r) J_f(r) dr | c_i\right)\right] + \operatorname{var}\left[E\left(\int_0^1 J_{fi}(r) J_f(r) dr | c_i\right)\right] \\
= \; & E\left(\int_0^1 \int_0^1 \operatorname{cov}[J_{fi}(r) J_f(r), J_{fi}(u) J_f(u)|c_i] dr du\right) + \operatorname{var}\left(\int_0^1 E[J_{fi}(r) J_f(r)|c_i] dr\right) \\
= \; & E\left(\int_0^1 \int_0^1 (E[J_{fi}(r) J_f(r) J_{fi}(u) J_f(u)|c_i] - E[J_{fi}(r) J_f(r)|c_i] E[J_{fi}(u) J_f(u)|c_i]) dr du\right) \\
+ \; & \operatorname{var}\left(\int_0^1 E[J_{fi}(r) J_f(r)|c_i] dr\right)
\end{aligned}
$$

The covariance matrix of $[J_{fi}(r), J_f(r), J_{fi}(u), J_f(u)]'|c_i$ has the following unique elements:

$$\begin{aligned}
\Sigma_{11} &= \int_0^r e^{2c_i(r-x)}dx, \\
\Sigma_{12} &= \int_0^r e^{c_i(r-x)}M_c\,(r-x)\,dx, \\
\Sigma_{13} &= \int_0^{r\wedge u} e^{c_i(2r-2x)}dx, \\
\Sigma_{14} &= \int_0^{r\wedge u} e^{c_i(r-x)}M_c\,(u-x)\,dx, \\
\Sigma_{22} &= \int_0^r M_c^2\,(r-x)\,dx, \\
\Sigma_{23} &= \int_0^{r\wedge u} e^{c_i(u-x)}M_c\,(r-x)\,dx, \\
\Sigma_{24} &= \int_0^{r\wedge u} M_c\,(r-x)\,M_c\,(u-x)\,dx, \\
\Sigma_{33} &= \int_0^u e^{2c_i(u-x)}dx, \\
\Sigma_{34} &= \int_0^u e^{c_i(u-x)}M_c\,(u-x)\,dx, \\
\Sigma_{44} &= \int_0^u M_c^2\,(u-x)\,dx,
\end{aligned}$$

suggesting $E[J_{fi}(r)J_f(r)J_{fi}(u)J_f(u)|c_i] = \Sigma_{14}\Sigma_{23} + \Sigma_{13}\Sigma_{24} + \Sigma_{12}\Sigma_{34}$, which in turn implies

$$\begin{aligned}
&\operatorname{var}(v_4) \\
&= E\left(\int_0^1\int_0^1 (\Sigma_{14}\Sigma_{23} + \Sigma_{13}\Sigma_{24} + \Sigma_{12}\Sigma_{34} - \Sigma_{12}\Sigma_{34}) drdu\right) + \operatorname{var}\left(\int_0^1 \Sigma_{12} dr\right) \\
&= E\left[\int_0^1\int_0^1 \left(\int_0^{r\wedge u} e^{c_i(r-x)} M_c(u-x)\,dx\right)\left(\int_0^{r\wedge u} e^{c_i(u-x)} M_c(r-x)\,dx\right) drdu\right] \\
&+ E\left[\int_0^1\int_0^1 \left(\int_0^{r\wedge u} e^{c_i(2r-2x)} dx\right)\left(\int_0^{r\wedge u} M_c(r-x)\,M_c(u-x)\,dx\right) drdu\right] \\
&+ \operatorname{var}\left[\int_0^1 \left(\int_0^r e^{c_i(r-x)} M_c(r-x)\,dx\right) dr\right] \\
&= E\left(\int_0^1\int_0^1\int_0^{r\wedge u}\int_0^{r\wedge u} e^{c_i(r+u-x-s)} M_c(u-x)\,M_c(r-s)\,dsdxdrdu\right) \\
&+ E\left(\int_0^1\int_0^1\int_0^{r\wedge u}\int_0^{r\wedge u} e^{c_i(2r-2s)} M_c(r-x)\,M_c(u-x)\,dsdxdrdu\right) \\
&+ \operatorname{var}\left(\int_0^1\int_0^r e^{c_i(r-x)} M_c(r-x)\,dxdr\right) \\
&= \int_0^1\int_0^1\int_0^{r\wedge u}\int_0^{r\wedge u} M_c(r+u-x-s)\,M_c(u-x)\,M_c(r-s)\,dsdxdrdu \\
&+ \int_0^1\int_0^1\int_0^{r\wedge u}\int_0^{r\wedge u} M_c(2r-2s)\,M_c(r-x)\,M_c(u-x)\,dsdxdrdu \\
&+ E\left[\left(\int_0^1\int_0^r e^{c_i(r-x)} M_c(r-x)\,dxdr\right)^2\right] - \left[E\left(\int_0^1\int_0^r e^{c_i(r-x)} M_c(r-x)\,dxdr\right)\right]^2.
\end{aligned}$$

Hence,

$$\begin{aligned}
\operatorname{var}(v_4) &= \int_0^1\int_0^1\int_0^{r\wedge u}\int_0^{r\wedge u} M_c(r+u-x-s)\,M_c(u-x)\,M_c(r-s)\,dsdxdrdu \\
&+ \int_0^1\int_0^1\int_0^{r\wedge u}\int_0^{r\wedge u} M_c(2r-2s)\,M_c(r-x)\,M_c(u-x)\,dsdxdrdu \\
&+ \int_0^1\int_0^r\int_0^1\int_0^u M_c(r+u-x-s)\,M_c(r-x)\,M_c(u-s)\,dsdudxdr \\
&- \left(\int_0^1\int_0^r M_c^2(r-x)\,dxdr\right)^2. \qquad (30)
\end{aligned}$$

It is easy to verify that $\omega_{45} = 0$. As for ω_{46}, we have

$$
\begin{aligned}
\operatorname{cov}(v_4, v_6) &= E\left[\operatorname{cov}(v_4, v_6|c_i)\right] + \operatorname{cov}\left[E(v_4|c_i), E(v_6|c_i)\right] \\
&= E\left[\operatorname{cov}\left(\int_0^1 J_{fi}(r)J_f(r)dr, \int_0^1 J_f^2(r)dr|c_i\right)\right] \\
&+ \operatorname{cov}\left[E\left(\int_0^1 J_{fi}(r)J_f(r)dr|c_i\right), E\left(\int_0^1 J_f^2(r)dr|c_i\right)\right] \\
&= E\left(\int_0^1\int_0^1 \operatorname{cov}(J_{fi}(r)J_f(r), J_f^2(u)|c_i)drdu\right) \\
&= E\left(\int_0^1\int_0^1 (E[J_{fi}(r)J_f(r)J_f^2(u)|c_i] - E[J_{fi}(r)J_f(r)|c_i]E[J_f^2(u)|c_i])drdu\right).
\end{aligned}
$$

The covariance matrix of $[J_{fi}(r), J_f(r), J_f(u)]'|c_i$ is given by

$$
\Sigma = \begin{bmatrix}
\int_0^r e^{2c_i(r-x)}dx & \int_0^r e^{c_i(r-x)}M_c(r-x)\,dx & \int_0^{r\wedge u} e^{c_i(r-x)}M_c(u-x)\,dx \\
\cdot & \int_0^r M_c(r-x)^2\,dx & \int_0^{r\wedge u} M_c(u-x)\,M_c(r-x)\,dx \\
\cdot & \cdot & \int_0^u M_c(u-x)^2\,dx
\end{bmatrix},
$$

giving $E[J_{fi}(r)J_f(r)J_f^2(u)|c_i] = 2\Sigma_{13}\Sigma_{23} + \Sigma_{12}\Sigma_{33}$, from which we obtain

$$
\begin{aligned}
&\operatorname{cov}(v_4, v_6) \\
&= E\left[\int_0^1\int_0^1\left(2\Sigma_{13}\Sigma_{23} + \Sigma_{12}\Sigma_{33} - \int_0^r e^{c_i(r-x)}M_c(r-x)\,dx\int_0^u M_c(u-x)^2\,dx\right)drdu\right] \\
&= 2E\left[\int_0^1\int_0^1\left(\int_0^{r\wedge u} e^{c_i(r-x)}M_c(u-x)\,dx\right)\left(\int_0^{r\wedge u} M_c(u-x)\,M_c(r-x)\,dx\right)drdu\right] \\
&+ E\left[\int_0^1\int_0^1\left(\int_0^r e^{c_i(r-x)}M_c(r-x)\,dx\right)\left(\int_0^u M_c(u-x)^2\,dx\right)drdu\right] \\
&- E\left[\int_0^1\int_0^1\left(\int_0^r e^{c_i(r-x)}M_c(r-x)\,dx\right)\left(\int_0^u M_c(u-x)^2\right)dxdrdu\right] \\
&= 2\int_0^1\int_0^1\left(\int_0^{r\wedge u} M_c(r-x)\,M_c(u-x)\,dx\right)^2 drdu. \qquad (31)
\end{aligned}
$$

Next, consider v_5. We begin by noting that $\alpha_5 = 0$. For ω_5^2,

$$
\begin{aligned}
&\text{var}\,(v_5) \\
&= E[\text{var}\,(v_5|c_i)] + \text{var}[E\,(v_5|c_i)] \\
&= E\left[\text{var}\left(\int_0^1 J_{\epsilon i}(r)J_f(r)dr|c_i\right)\right] + \text{var}\left[E\left(\int_0^1 J_{\epsilon i}(r)J_f(r)dr|c_i\right)\right] \\
&= E\left(\int_0^1\int_0^1 \text{cov}[J_{\epsilon i}(r)J_f(r), J_{\epsilon i}(u)J_f(u)|c_i]dudr\right) \\
&= E\left(\int_0^1\int_0^1 (E[J_{\epsilon i}(r)J_f(r)J_{\epsilon i}(u)J_f(u)|c_i] - E[J_{\epsilon i}(r)J_f(r)|c_i]E[J_{\epsilon i}(u)J_f(u)|c_i])dudr\right) \\
&= E\left(\int_0^1\int_0^1 E[J_{\epsilon i}(r)J_{\epsilon i}(u)EJ_f(r)J_f(u)|c_i]dudr\right) \\
&= E\left[\int_0^1\int_0^1 \left(\int_0^{r\wedge u} e^{c_i(r-x)}e^{c_i(u-x)}dx\right)\left(\int_0^{r\wedge u} M_c\,(r-x)\,M_c\,(u-x)\,dx\right)dudr\right] \\
&= \int_0^1\int_0^1 \left(\int_0^{r\wedge u} M_c\,(r+u-2x)\,dx\right)\left(\int_0^{r\wedge u} M_c\,(r-x)\,M_c\,(u-x)\,dx\right)dudr. \qquad (32)
\end{aligned}
$$

Finally, consider v_6. It is not difficult to show that $\alpha_6 = \alpha_4$ and $\omega_{56} = 0$. We therefore focus on ω_6^2, which can be evaluated in the following fashion:

$$
\begin{aligned}
\text{var}\,(v_6) &= \text{var}\left(\int_0^1 J_f^2(r)dr\right) \\
&= \int_0^1\int_0^1 \text{cov}(J_f^2(r), J_f^2(u))dudr \\
&= \int_0^1\int_0^1 (E[(J_f(r))^2(J_f(u))^2] - E[(J_f(r))^2]E[(J_f(u))^2])dudr \\
&= 2\int_0^1\int_0^1 \left(\int_0^{r\wedge u} M_c\,(u-x)\,M_c\,(r-x)\,dx\right)^2 dudr \qquad (33)
\end{aligned}
$$

where we have used the fact that $J_f\,(r)$ and $J_f\,(u)$ are jointly normally distributed.

The required result is obtained by direct substitution of the above moments into (22). ■

Table 1: 5% size and local power when $N = T = 100$.

	$\mu_c = -3$		$\mu_c = -6$		$\mu_c = -12$	
γ	CADF	CIPS	CADF	CIPS	CADF	CIPS
			$\sigma_c^2 = 0$			
1	8.4	63.8	16.1	100.0	45.4	100.0
5	8.3	63.9	16.1	100.0	45.3	100.0
10	8.3	63.7	16.1	100.0	45.3	100.0
20	8.3	63.8	16.1	100.0	45.3	100.0
			$\sigma_c^2 = \mu_1^2/12$			
1	9.4	53.2	20.2	100.0	47.5	100.0
5	15.5	74.6	38.6	100.0	64.4	100.0
10	32.7	94.0	59.9	100.0	73.6	100.0
20	56.9	99.5	73.2	100.0	78.8	100.0

Notes: The DGP is given by $y_{i,t} = \rho_i y_{i,t-1} + \gamma f_t + \epsilon_{i,t}$, where $y_{i,0} = 0$, $\rho_i = \exp(T^{-1}c_i)$ and $(f_t, \epsilon_{i,t}) \sim N(0, I_2)$. Here $c_i \sim U(a, b)$, where a and b are calibrated such that c_i has mean μ_c and variance σ_c^2. The rejection frequencies for CADF are averaged across the cross-section.

Table 2: 5% size.

β	CADF_i	$ADF_{\hat{e}}(i)$	CIPS	$P_{\hat{e}}$
		$N = T = 50$		
0.1	4.5	4.4	6.0	5.6
1.0	4.5	4.4	5.9	5.3
5.0	4.5	4.4	5.9	5.5
10.0	4.5	4.4	5.9	5.5
15.0	4.5	4.4	6.0	5.5
20.0	4.5	4.4	5.9	5.5
30.0	4.5	4.4	6.0	5.5
50.0	4.5	4.4	6.0	5.5
		$N = T = 100$		
0.1	4.7	4.6	5.2	5.8
1.0	4.7	4.7	5.7	5.7
5.0	4.7	4.6	5.8	5.7
10.0	4.7	4.6	5.7	5.9
15.0	4.7	4.6	5.7	5.9
20.0	4.7	4.7	5.7	5.9
30.0	4.7	4.6	5.7	5.9
50.0	4.7	4.6	5.7	5.9

Notes: See Table 1 for an explanation.

Table 3: 5% local power.

	$N = T = 50$				$N = T = 100$			
β	CADF_i	$ADF_{\hat{e}}(i)$	CIPS	$P_{\hat{e}}$	CADF_i	$ADF_{\hat{e}}(i)$	CIPS	$P_{\hat{e}}$
				$a = b = -3$				
0.1	7.7	13.3	57.3	100.0	8.4	15.0	63.5	100.0
1.0	7.7	13.3	57.3	100.0	8.4	15.0	63.8	100.0
5.0	7.7	13.4	57.7	100.0	8.3	15.0	63.9	100.0
10.0	7.8	13.3	57.3	100.0	8.3	15.0	63.7	100.0
15.0	7.8	13.3	57.3	100.0	8.3	15.0	63.8	100.0
20.0	7.8	13.3	57.3	100.0	8.3	15.0	63.8	100.0
30.0	7.8	13.3	57.2	100.0	8.3	15.0	63.8	100.0
50.0	7.8	13.3	57.2	100.0	8.3	15.0	63.8	100.0
				$a = -6, b = 0$				
0.1	8.5	13.6	48.6	100.0	9.3	15.9	52.2	100.0
1.0	8.6	12.2	48.5	98.8	9.4	13.9	53.2	99.8
5.0	14.6	4.2	70.4	31.6	15.5	4.5	74.6	38.0
10.0	31.3	1.7	91.1	8.6	32.7	1.7	94.0	8.3
15.0	45.8	0.9	97.3	3.4	47.1	0.8	98.4	2.8
20.0	55.7	0.5	99.4	2.0	56.9	0.5	99.5	1.4
30.0	66.6	0.2	100.0	0.8	67.6	0.2	99.9	0.3
50.0	74.9	0.0	100.0	0.4	75.4	0.0	100.0	0.1
				$a = -12, b = 0$				
0.1	17.9	27.2	99.3	100.0	20.0	33.4	99.9	100.0
1.0	18.3	23.3	99.3	99.8	20.2	27.5	100.0	100.0
5.0	36.7	5.4	100.0	44.9	38.6	5.6	100.0	47.8
10.0	58.6	1.5	100.0	12.5	59.9	1.5	100.0	10.4
15.0	68.2	0.6	100.0	5.5	68.9	0.5	100.0	4.0
20.0	72.7	0.2	100.0	3.4	73.2	0.2	100.0	1.8
30.0	76.8	0.1	100.0	1.6	77.2	0.0	100.0	0.5
50.0	79.7	0.0	100.0	1.0	80.2	0.0	100.0	0.2

Notes: See Table 1 for an explanation.