SFB 823

Panel cointegrating polynomial regression analysis and the environmental Kuznets curve

Robert M. de Jong, Martin Wagner

Discussion Paper

Nr. 22/2018

Panel Cointegrating Polynomial Regression Analysis and the Environmental Kuznets Curve *

Robert M. de Jong† Martin Wagner ‡

October 7, 2018

Abstract

This paper develops a modified and a fully modified OLS estimator for a panel of cointegrating polynomial regressions, i.e. regressions that include an integrated process and its powers as explanatory variables. The stationary errors are allowed to be serially correlated and the regressors are allowed to be endogenous and we allow for individual and time fixed effects. Inspired by Phillips and Moon (1999) we consider a cross-sectional i.i.d. random linear process framework. The modified OLS estimator utilizes the large cross-sectional dimension that allows to consistently estimate and subtract an additive bias term without the need to also transform the dependent variable as required in fully modified OLS estimation. Both developed estimators have zero mean Gaussian limiting distributions and thus allow for standard asymptotic inference. Our illustrative application indicates that the developed methods are a potentially useful addition to not least the environmental Kuznets curve literature's toolkit.

JEL Classification: C13, C23, Q20

Keywords: Cointegration, Environmental Kuznets Curve, Panel Data, Polynomial Transformation, Unit Roots

---

*Both authors gratefully acknowledge financial support from the Jubiläumsfonds of the Oesterreichische Nationalbank (Grant Nr. 15334). The second author furthermore gratefully acknowledges financial support from Deutsche Forschungsgemeinschaft via the Collaborative Research Center 823, *Statistical Modelling of Nonlinear Dynamic Processes* (Projects A3 and A4). The outstanding research assistance of Karsten Reichold is also gratefully acknowledged. At the time of the completion of this manuscript the second author serves as Chief Economist of the Bank of Slovenia. The views expressed in this paper are, however, solely those of the authors and not necessarily of the Bank of Slovenia or the European System of Central Banks. On top of this the usual disclaimer applies.

†Department of Economics, Ohio State University, 444 Arps Hall, Columbus, OH 43210, USA, Email: de-jong.8@osu.edu.

‡Faculty of Statistics, Technical University Dortmund, Vogelpothsweg 87, 44227 Dortmund, Germany & Institute for Advanced Studies, Vienna & Bank of Slovenia, Ljubljana, Email: mwagner@statistik.tu-dortmund.de.

1 Introduction

This paper is motivated by the large and growing literature investigating the environmental Kuznets curve (EKC) hypothesis that postulates an inverse U-shaped relationship between measures of economic development, typically GDP, and measures of pollution, often proxied by emissions. The term EKC refers by analogy to the inverted U-shaped relationship between the level of economic development and the degree of income inequality postulated by Simon Kuznets (1955) in his 1954 presidential address to the American Economic Association.[1]

In the empirical EKC literature a large variety of specification and estimation strategies are pursued, in both a time series and panel context. A large part of the literature resorts to unit root and cointegration techniques, estimating – in the panel case – equations like

$$y_{it} = \alpha_i + x_{it}\beta_1 + x_{it}^2\beta_2 + u_{it}, \tag{1}$$

with y_{it} denoting (in our application) log CO_2 emissions per capita and x_{it} log GDP per capita. If log GDP per capita is an integrated process, then the above equation involves an integrated process and its square – often also the third power is included – as regressors. It is immediate to see and well-known, see, e.g., Wagner (2012, 2015), that powers of integrated processes are not themselves integrated processes. Consequently, theory has to be developed for regressions involving integrated processes and their powers as regressors. In a time series setting this has been done, e.g., in Wagner and Hong (2016), who refer to such regressions as *cointegrating polynomial regressions* (CPRs) when the errors are stationary. In particular, that paper extends the fully modified OLS (FM-OLS) estimator of Phillips and Hansen (1990) from the cointegrating linear to the cointegrating polynomial regression case.[2]

[1] The empirical EKC literature started in the first half of the 1990s, with early important contributions including Grossman and Krueger (1993) or Holtz-Eakin and Selden (1995). Early survey papers like Stern (2004) or Yandle *et al.* (2004) already find more than 100 refereed publications, with the number growing substantially since then. For more discussion on the empirical literature and the theoretical underpinnings of the EKC see, e.g., Wagner (2015). Inverted U-shaped relationships also feature prominently in modelling the relationship between energy or material intensity and GDP per capita, see, e.g., Malenbaum (1978). Additionally, relationships involving powers of integrated processes as regressors are used in the exchange rate target zone literature, e.g., Darvas (2008) or Svensson (1992).

[2] The original motivation for Wagner and Hong (2016) to develop a fully modified type estimator for CPRs was the widespread practice in the EKC literature to consider, e.g., a quadratic CPR *incorrectly* as a cointegrating linear regression with two integrated regressors. Wagner (2015) compares results obtained with the fully modified estimator for CPRs and when using the standard estimator and finds that the estimation results are relatively similar, but that inference concerning the presence of a CPR relationship differs substantially. These findings are understood by the results of Stypka *et al.* (2018) who show that the two estimation approaches lead to the same asymptotic distribution, but that cointegration testing is asymptotically affected. The "tailor-made" estimator for CPRs, however, exhibits superior finite sample performance.

Here, we perform a similar extension of the FM-OLS estimator to CPRs in a large N and large T panel setting allowing for individual and time fixed effects. Since in the parametric EKC literature typically only quadratic and cubic formulations are considered we also focus here on the cubic formulation and abstain from considering a general polynomial degree. Our results, however, extend to higher order polynomials under corresponding moment conditions. Also, of course, the analysis can be generalized straightforwardly to consider multiple integrated processes and their powers as regressors. In terms of assumptions we follow Phillips and Moon (1999) by using a cross-sectional i.i.d. random linear process framework. Pedroni (2000), a seminal contribution to panel FM-OLS estimation, considers a non-random structure and only assumes that cross-sectional limits of (non-random) quantities like long run variances exist. This framework has been generalized and substantiated by Phillips and Moon (1999), who introduced random linear processes to the panel cointegration literature. Clearly, also our results hold, with obvious notational changes, in case of linear processes rather than random linear processes. Another aspect we do not consider here is joint asymptotics, also studied in detail in Phillips and Moon (1999). We only consider sequential asymptotics with T to infinity followed by N to infinity. Generalizations to joint asymptotics are beyond the scope of this paper.[3]

The narrow focus with respect to polynomial degree and the confinement to sequential asymptotics gives room to zoom in on another aspect that has gone unnoticed in the (linear) panel cointegration literature. The cross-sectional dimension allows to consider another estimator that we call *modified* OLS. This estimator is based on subtracting a consistent estimator of a second-order bias term but without the need to transforming the dependent variable like in FM-OLS or the addition of leads and lags of the first difference of the integrated regressor as in Dynamic OLS (Saikkonen, 1991, Stock and Watson, 1993). This is possible, since the large cross-sectional dimension transforms the individual specific random bias term into an expected value that can be consistently estimated and hence subtracted.

We compare the modified and fully modified OLS estimators, as well as tests based upon them, by means of a small simulation study and also provide a brief illustration of the developed methodology by estimating the EKC for carbon dioxide (CO_2) emissions, using two data sets, a long data set with a small cross-sectional and large time dimension ($N = 19$ and $T = 135$) taken from Wagner *et al.* (2018) and a wide data set, with $N = 89$ and $T = 54$.[4]

The paper is organized as follows. The following section introduces the model and the assumptions.

[3]Let us note for completeness that we are more general than, e.g., Phillips and Moon (1999) in one aspect by allowing for time effects in addition to individual effects.

[4]Wagner *et al.* (2018) develop fully modified OLS type estimators – feasible by construction only for small N – for systems of seemingly unrelated cointegrating polynomial regressions. Our results here complement the results of that paper for data sets with a more sizeable cross-sectional dimension.

Section 3 discusses estimation in the model with only individual specific fixed effects. The results of this section also provide important input for estimation in the model with individual and time effects discussed in Section 4. The finite sample performance is considered in Section 5 and Section 6 contains the illustration with the environmental Kuznets curve for CO_2 emissions. Section 7 briefly summarizes and concludes. Two appendices and one supplementary appendix complement the main text: The proofs of the theorems are contained in Appendix A, whilst Appendix B contains several useful lemmata. Supplementary Appendix C contains calculations in relation to some functionals of Brownian motions.

2 Model and Assumptions

The model considered in this paper is a fixed effects model with both individual and time effects and a regressor with a unit root and its square and cube, i.e.,

$$y_{it} = \alpha_i + \gamma_t + x_{it}\beta_1 + x_{it}^2\beta_2 + x_{it}^3\beta_3 + u_{it}, \tag{2}$$

$$x_{it} = x_{i,t-1} + v_{it}, \tag{3}$$

where for brevity we assume $x_{i0} = 0$. Alternatively, without any consequences for the asymptotic analysis, we could assume that the random variables x_{i0} are a sequence of cross-sectionally independent random variables with finite second moments. The quadratic case is, evidently, included with $\beta_3 = 0$ and appropriate (2×2) matrices replacing (3×3) matrices.

The cross-sectionally independently and identically distributed error processes $\eta_{it} = (u_{it}, v_{it})'$ are assumed to be random linear processes fulfilling a functional central limit theorem similar to Phillips and Moon (1999, Lemma 3), i.e.,

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{[rT]} \eta_{it} \overset{T \to \infty}{\Rightarrow} B_i(r) = \Omega_i^{1/2} W_i(r), \tag{4}$$

where $W_i(r) = (W_{ui}(r), W_{vi}(r))'$, with $B_i(r)$ partitioned analogously, is a bivariate standard Wiener process. The random long run variance matrices are partitioned as

$$\Omega_i = \begin{pmatrix} \Omega_{uui} & \Omega_{uvi} \\ \Omega_{vui} & \Omega_{vvi} \end{pmatrix}. \tag{5}$$

For later usage we also define what is usually referred to as half long run variance matrix and partition it analogously, i.e.

$$\Delta_i = \begin{pmatrix} \Delta_{uui} & \Delta_{uvi} \\ \Delta_{vui} & \Delta_{vvi} \end{pmatrix}, \tag{6}$$

with $\Omega_i = \Delta_i + \Delta_i' - \Sigma_i$, with Σ_i the random contemporaneous variance matrix.

We denote the time-demeaned variables and the averages over time by, e.g., $\tilde{y}_{it}$ and $\bar{y}_{i.}$, i.e.,

$$\tilde{y}_{it} = y_{it} - \bar{y}_{i.} = y_{it} - \frac{1}{T}\sum_{t=1}^{T} y_{it}, \tag{7}$$

with analogous quantities defined for x_{it} (and its powers), u_{it} and v_{it}. In addition, we write

$$\tilde{X}_{it} = \begin{pmatrix} x_{it} - \overline{x}_{i.} \\ x_{it}^2 - \overline{x_{i.}^2} \\ x_{it}^3 - \overline{x_{i.}^3} \end{pmatrix}. \tag{8}$$

For the two-way effects model we need the correspondingly two-way demeaned quantities, i.e. we define

$$\begin{aligned} \check{y}_{it} &= y_{it} - \frac{1}{T}\sum_{t=1}^{T} y_{it} - \frac{1}{N}\sum_{j=1}^{N} y_{jt} + \frac{1}{NT}\sum_{j=1}^{N}\sum_{t=1}^{T} y_{it} \\ &= y_{it} - \bar{y}_{i.} - \bar{y}_{.t} + \bar{y}_{..}, \end{aligned} \tag{9}$$

with again similar definitions for x_{it} (and its powers), u_{it} and v_{it}.

We are now ready to state the assumptions, denoting with $[x]$ the integer part of $x \in \mathbb{R}$.

Assumption 1. *The random variables η_{it} are i.i.d. across i for all $t \in \mathbb{N}$. Furthermore, $(\Delta_i, \Omega_i)_{i\in\mathbb{N}}$ are i.i.d. and independent of $(W_i)_{i\in\mathbb{N}}$. The matrices $(\Omega_i)_{i\in\mathbb{N}}$ are positive definite almost surely. Denoting with $G_T = diag(T^{-1}, T^{-3/2}, T^{-2})$ it holds for all $i \in \mathbb{N}$ that:*

$$\begin{aligned} T^{1/2} G_T \tilde{X}_{i,[rT]} \overset{T\to\infty}{\Rightarrow} \begin{pmatrix} B_{vi}(r) - \int_0^1 B_{vi}(s)ds \\ B_{vi}^2(r) - \int_0^1 B_{vi}^2(s)ds \\ B_{vi}^3(r) - \int_0^1 B_{vi}^3(s)ds \end{pmatrix} &= \tilde{\mathbf{B}}_{vi}(r) = \mathbf{B}_{vi}(r) - \int_0^1 \mathbf{B}_{vi}(r)dr \\ &= D_i \tilde{\mathbf{W}}_{vi}(r) = D_i \left(\mathbf{W}_{vi}(r) - \int_0^1 \mathbf{W}_{vi}(r)dr\right), \end{aligned} \tag{10}$$

with $\mathbf{W}_{vi}(r) = (W_{vi}(r), W_{vi}(r)^2, W_{vi}(r)^3)'$ and $D_i = diag(\Omega_{vvi}^{1/2}, \Omega_{vvi}, \Omega_{vvi}^{3/2})$. Furthermore,

$$\sum_{t=1}^{T} G_T \tilde{X}_{it}\tilde{u}_{it} \overset{d,T\to\infty}{\longrightarrow} \int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r) + \Delta_{vui}\left(\ 1, 2\int_0^1 B_{vi}(r)dr, 3\int_0^1 B_{vi}(r)^2dr\ \right)'. \tag{11}$$

and for $i \neq j$,

$$\sum_{t=1}^{T} G_T \tilde{X}_{it}\tilde{u}_{jt} \overset{d,T\to\infty}{\longrightarrow} \int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{uj}(r). \tag{12}$$

Weak convergence in equations (10), (11) and (12) is joint.

Conditional upon Ω_{uui} the process $B_{ui}(r)$ is Gaussian. By construction, $\Omega_{uui}^{-1/2}B_{ui}(r)$ is conditionally $N(0,r)$-distributed. This implies, for $p>0$, that:

$$\mathbb{E}|B_{ui}(r)|^p < \infty \quad \text{if} \quad \mathbb{E}|\Omega_{ui}|^{p/2} < \infty \tag{13}$$

and similarly,

$$\mathbb{E}|B_{vi}(r)|^p < \infty \quad \text{if} \quad \mathbb{E}|\Omega_{vi}|^{p/2} < \infty. \tag{14}$$

We furthermore define for later usage $B_{u\cdot v,i}(r) = B_{ui}(r) - \Omega_{uvi}\Omega_{vvi}^{-1}B_{vi}(r)$. By construction $B_{u\cdot v,i}(r)$ is, conditional upon the long run variances, independent of $B_{ui}(r)$. The conditional (long run) variance of $B_{u\cdot v,i}(r)$ is given by $\Omega_{u\cdot v,i} = \Omega_{uui} - \Omega_{vvi}^{-1}\Omega_{vui}^2$. In the sequel we use the short hand notation $\rho_i = \Omega_{uvi}\Omega_{vvi}^{-1}$.

The processes $B_{u\cdot v,i}(r)$ play a key role in the development of FM-OLS type estimators that have asymptotic distributions free of second order bias terms. Given that the long run variances are in general unknown, estimators need to be constructed invoking consistent estimators of the long run variances. The existence of consistent long run variance estimators is ensured by Assumption 2 below, whilst Assumption 3 contains the key functional central limit result underlying fully modified type estimation:

Assumption 2. *The estimators $\hat{\Delta}_i$ and $\hat{\Omega}_i$ satisfy $\hat{\Delta}_i \overset{p,T\to\infty}{\longrightarrow} \Delta_i$ and $\hat{\Omega}_i \overset{p,T\to\infty}{\longrightarrow} \Omega_i$.*

Assumption 3. *Jointly with the weak convergence results of equations (10), (11) and (12) in Assumption 1 it holds that*

$$\sum_{t=1}^{T} G_T\tilde{X}_{it}v_{it} \overset{d,T\to\infty}{\longrightarrow} \int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{vi}(r) + \Delta_{vvi}(1, 2\int_0^1 B_{vi}(r)dr, 3\int_0^1 B_{vi}^2(r)dr)' \tag{15}$$

and for $i \neq j$,

$$\sum_{t=1}^{T} G_T\tilde{X}_{it}v_{jt} \overset{d,T\to\infty}{\longrightarrow} \int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{vj}(r). \tag{16}$$

Our setting is inspired by Phillips and Moon (1999), but we restrict ourselves to a special case with respect to asymptotic theory in that we only consider sequential asymptotics with $T\to\infty$ followed by $N\to\infty$. This framework already leads to substantial complications, as will be seen below. Furthermore, with appropriate rate restrictions it seems likely that the results derived here continue to hold in joint asymptotic with T going to infinity fast enough in relation to N.

Also, at this point we formulate only high level assumptions and abstain from adding primitive assumptions that generate our Assumptions 1 to 3. The literature provides several pathways to derive these results from primitive assumptions that are well understood (see, e.g., de Jong, 2002, Ibragimov and Phillips, 2008 or Park and Phillips, 2001).[5] It is worth mentioning that Pötscher (2004) is an important contribution covering a wide class of nonlinear functions by deriving convergence results without invoking the continuous mapping theorem.

3 Individual Effects Only

In this section we consider the case when only individual fixed effects α_i are present. Time effects are added in the following section, where it will be seen that the results of this section are of prime importance also for the two-way effects case.

3.1 The OLS Estimator

If the focus is only on the coefficient vector $\beta = (\beta_1, \beta_2, \beta_3)'$, the corresponding OLS estimator is typically referred to as Least Squares Dummy Variables (LSDV for short) estimator and is given by:

$$\tilde{\beta} = (\sum_{i=1}^{N}\sum_{t=1}^{T} \tilde{X}_{it}\tilde{X}_{it}')^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T} \tilde{X}_{it}\tilde{y}_{it}. \tag{17}$$

To next discuss the asymptotic behavior of the OLS estimator we define:

$$V_1 = \mathbb{E}\left(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)\tilde{\mathbf{B}}_{vi}(r)'dr\right) = \mathbb{E}(D_i M D_i), \tag{18}$$

where

$$M = \mathbb{E}\left(\int_0^1 \tilde{\mathbf{W}}_{vi}(r)\tilde{\mathbf{W}}_{vi}(r)'dr\right) = \begin{pmatrix} 1/6 & 0 & 3/8 \\ 0 & 5/12 & 0 \\ 3/8 & 0 & 39/20 \end{pmatrix}, \tag{19}$$

as shown in Lemma 9 of Supplementary Appendix C.

Theorem 1. *Under Assumption 1, when* $\gamma_t = 0$ *for all* t, $\mathbb{E}|\Omega_{uui}|^{1/2} < \infty$, $\mathbb{E}|\Omega_{vvi}|^3 < \infty$, $\mathbb{E}|\Omega_{vvi}^{-1/2}\Omega_{uvi}| < \infty$, $\mathbb{E}|\Delta_{vui}| < \infty$, $\mathbb{E}|\Delta_{vui}\Omega_{vvi}^{1/2}| < \infty$, $\mathbb{E}|\Delta_{vui}\Omega_{vvi}| < \infty$, *and for* $p = 1, 2, 3$, $\mathbb{E}|\Omega_{uui}\Omega_{vvi}^{p}|^{1/2} < \infty$ *and* $\mathbb{E}|\Omega_{uvi}\Omega_{vvi}^{p/2-1/2}| < \infty$, *it holds that*

$$G_T^{-1}(\tilde{\beta} - \beta) \quad \stackrel{p, N\to\infty, T\to\infty}{\longrightarrow} \quad V_1^{-1}\begin{pmatrix} -(1/2)\mathbb{E}(\Omega_{uvi}) + \mathbb{E}(\Delta_{vui}) \\ 0 \\ -\mathbb{E}(\Omega_{vvi}\Omega_{uvi}) + (3/2)\mathbb{E}(\Omega_{uvi}\Delta_{vui}) \end{pmatrix}. \tag{20}$$

[5] In case one considers a standard, i.e., non-random, coefficient setting, the convergence results posited in our assumptions follow immediately from, e.g., the results of Wagner and Hong (2016). In this case, as mentioned in the introduction, it is sufficient to assume that the cross-sectional averages of, e.g., Ω_i converge to well-defined positive definite limits to derive the results in this paper.

The theorem is formulated with appropriate moment assumptions on the random long run variance matrices, which are necessary to obtain consistency, as $N, T \to \infty$ sequentially. Due to the bias term the convergence rate does not depend upon N but only upon T and that the result is a convergence in probability result to a non-random quantity. As is well-known, the bias term depends upon the dependence structure between u_{it} and v_{it} and is zero in case that Ω_{uvi} and Δ_{uvi} are zero with probability one. In this case the OLS estimator features a faster rate of convergence given by $N^{1/2}G_T^{-1}$ and the limiting distribution is normal. The panel dimension allows to remove the bias term responsible for the result independent of N given above in (20) with a one-step correction by subtracting a consistent estimator of the bias. This is simpler than the two-part transformation required for FM-OLS, where additionally also the dependent variable is modified. Hence, the panel setting in fact allows for more possibilities of corrections than the pure time series setting, which have not yet been explored in the literature.

3.2 A Modified OLS Estimator

In order to remove the additive bias and to obtain an asymptotic normal distribution at the convergence rate $N^{1/2}G_T^{-1}$ we define

$$\tilde{\beta}^m = (\sum_{i=1}^{N}\sum_{t=1}^{T}\tilde{X}_{it}\tilde{X}_{it}')^{-1}\sum_{i=1}^{N}(\sum_{t=1}^{T}\tilde{X}_{it}\tilde{y}_{it} - \tilde{C}_i), \tag{21}$$

where

$$\tilde{C}_i = \hat{\Delta}_{vui}(T, 2\sum_{t=1}^{T}x_{it}, 3\sum_{t=1}^{T}x_{it}^2)' + (-(1/2)T\hat{\Omega}_{uvi}, 0, -T^2\hat{\Omega}_{vvi}\hat{\Omega}_{uvi})'. \tag{22}$$

Note that for the modified OLS estimator, in contrast with the fully modified OLS estimator considered below, the dependent variable is not changed. The large cross-sectional dimension allows to consistently remove the expectation of the cross-product of the regressors and the errors, the endogeneity bias arising in OLS, directly.

In the theorem below, a role is played by the matrix

$$Q = \mathbb{E}(\int_0^1 \tilde{\mathbf{W}}(r)dW(r))(\int_0^1 \tilde{\mathbf{W}}(r)dW(r))') = \begin{pmatrix} 1/3 & 0 & 9/10 \\ 0 & 59/60 & 0 \\ 9/10 & 0 & 101/20 \end{pmatrix}. \tag{23}$$

The calculations to show second equality in the above equation are provided in Lemma 13 in Supplementary Appendix C.

Theorem 2. *Under Assumptions 1, 2 and 3, when $\gamma_t = 0$ for all t, $\mathbb{E}|\Omega_{vvi}|^3 < \infty$, $\mathbb{E}|\Omega_{uui}| < \infty$, $\mathbb{E}|\Delta_{uvi}| < \infty$, $\mathbb{E}|\Delta_{uvi}\Omega_{vvi}^{1/2}| < \infty$, $\mathbb{E}|\Delta_{uvi}\Omega_{vvi}| < \infty$, $\mathbb{E}|\Omega_{uvi}^2\Omega_{vvi}^{-1}| < \infty$, and for $p = 1, 2, 3$,*

$\mathbb{E}|\Omega_{uui}\Omega_{vvi}^p| < \infty$, $\mathbb{E}|\Omega_{uvi}\Omega_{vvi}^{p-1}| < \infty$, and $\mathbb{E}|\Omega_{uvi}^2\Omega_{vvi}^{p-1}| < \infty$, it holds that

$$N^{1/2}G_T^{-1}(\tilde{\beta}^m - \beta) \overset{d, N\to\infty, T\to\infty}{\longrightarrow} N(0, V_1^{-1}\Sigma_1 V_1^{-1}), \tag{24}$$

where

$$\begin{aligned}\Sigma_1 &= \mathbb{E}\left(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r) - \mathbb{E}(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)|\Omega_i)\right)\left(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r) - \mathbb{E}(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)|\Omega_i)\right)' \\ &= \mathbb{E}(\Omega_{u\cdot v,i}D_iMD_i) + \mathbb{E}(\Omega_{uvi}^2\Omega_{vvi}^{-1}D_iQD_i) \\ &\quad - \mathbb{E}\begin{pmatrix} (1/4)\Omega_{uvi}^2 & 0 & (1/2)\Omega_{vvi}\Omega_{uvi}^2 \\ 0 & 0 & 0 \\ (1/2)\Omega_{vvi}\Omega_{uvi}^2 & 0 & \Omega_{vvi}^2\Omega_{uvi}^2 \end{pmatrix}. \end{aligned} \tag{25}$$

Using the modified OLS estimator for inference requires, of course, consistent estimators of V_1 and Σ_1. A consistent estimator of V_1 is straightforwardly given by

$$\hat{V}_1 = \tfrac{1}{N}\textstyle\sum_{i=1}^N \sum_{t=1}^T G_T\tilde{X}_{it}\tilde{X}_{it}'G_T. \tag{26}$$

An estimator for Σ_1 is constructed using consistent estimators of its components, i.e., is based on replacing in the above limiting expression the expected values by cross-sectional averages of estimates based on $\hat{\Omega}_i$ and $\hat{\Delta}_i$. Thus, e.g., the second component composing Σ_1 is estimated by $\frac{1}{N}\sum_{i=1}^N \hat{\Omega}_{uvi}^2\hat{\Omega}_{vvi}^{-1}\hat{D}_iM\hat{D}_i$ with $\hat{D}_i = \text{diag}(\hat{\Omega}_{vvi}^{1/2}, \hat{\Omega}_{vvi}, \hat{\Omega}_{vvi}^{3/2})$ and similarly for the other components.

3.3 A Fully Modified OLS Estimator

The above modified estimator utilizes the panel dimension for bias correction. In this section we consider a "classical" fully modified OLS (FM-OLS) estimator based on the Phillips and Hansen (1991) two part "full modification" that considers $\tilde{y}_{it}^+$, rather than $\tilde{y}_{it}$, and an additive bias correction. Define

$$\hat{\Delta}_{vui}^+ = \hat{\Delta}_{vui} - \hat{\Delta}_{vvi}\hat{\Omega}_{vvi}^{-1}\hat{\Omega}_{vui}, \tag{27}$$

$$\tilde{C}_i^+ = \hat{\Delta}_{vui}^+(T, 2\sum_{t=1}^T x_{it}, 3\sum_{t=1}^T x_{it}^2)' \tag{28}$$

and

$$\tilde{y}_{it}^+ = \tilde{y}_{it} - \hat{\Omega}_{uvi}\hat{\Omega}_{vvi}^{-1}v_{it} = \tilde{X}_{it}'\beta + \tilde{u}_{it} - \hat{\Omega}_{uvi}\hat{\Omega}_{vvi}^{-1}v_{it}. \tag{29}$$

With the necessary quantities defined, the FM-OLS estimator is given by:

$$\tilde{\beta}^+ = (\sum_{i=1}^N\sum_{t=1}^T \tilde{X}_{it}\tilde{X}_{it}')^{-1}\sum_{i=1}^N(\sum_{t=1}^T \tilde{X}_{it}\tilde{y}_{it}^+ - \tilde{C}_i^+) \tag{30}$$

Assumptions 1 to 3 are tailor-made to establish consistency of the FM-OLS estimator in this context. This extends the time series results contained in Wagner and Hong (2016) to the panel case.

Theorem 3. *Under Assumptions 1, 2 and 3, when $\gamma_t = 0$ for all t, $\mathbb{E}|\Omega_{vvi}|^3 < \infty$, $\mathbb{E}|\Omega_{uui}|^{1/2} < \infty$, $\mathbb{E}|\Omega_{vvi}^{-1/2}\Omega_{uvi}| < \infty$, $\mathbb{E}|\Omega_{vvi}|^{3/2} < \infty$, $\mathbb{E}|\Delta_{vui}| < \infty$, $\mathbb{E}|\Delta_{vui}\Omega_{vvi}^{1/2}| < \infty$, $\mathbb{E}|\Delta_{vui}\Omega_{vvi}| < \infty$, and for $p = 1, 2, 3$, $\mathbb{E}|\Omega_{uui}\Omega_{vvi}^{p}|^{1/2} < \infty$, and $\mathbb{E}|\Omega_{uvi}^2\Omega_{vvi}^{p-1}|^{1/2} < \infty$, $\mathbb{E}|\Omega_{uvi}\Omega_{vvi}^{(p-1)/2}| < \infty$, it holds that*

$$N^{1/2}G_T^{-1}(\tilde{\beta}^+ - \beta) \overset{d, N\to\infty, T\to\infty}{\longrightarrow} N(0, V_1^{-1}\Sigma_1^+ V_1^{-1}). \tag{31}$$

where

$$\Sigma_1^+ = \mathbb{E}(\Omega_{u\cdot v,i}D_iMD_i).$$

Inference for the FM-OLS estimator requires an estimate of $\mathbb{E}(\Omega_{u\cdot v,i}D_iMD_i)$. This can be done in several ways, one based on replacing the expectation by cross-sectional averaging, that is, $\frac{1}{N}\sum_{i=1}^{N}\hat{\Omega}_{u\cdot v,i}\hat{D}_iM\hat{D}_i$. The other one replaces $\hat{D}_iM\hat{D}_i$ in this formula by the sample analogue, i.e., $\frac{1}{N}\sum_{i=1}^{N}\hat{\Omega}_{u\cdot v,i}\sum_{t=1}^{T}G_T\tilde{X}_{it}\tilde{X}_{it}'G_T$.

Remark 1. *As can be expected, the result given above for the FM-OLS estimator simplifies considerably if we assume that the processes are cross-sectionally i.i.d. with non-random (and thus identical) long run variance matrices, e.g., $\Omega_{uvi} = \Omega_{uv}$ for $i = 1, 2, \ldots, N$ and similar for all other quantities. In this case the limiting distribution has a "standard" rather than a "sandwich" variance covariance matrix given by*

$$\Omega_{u\cdot v}(DMD)^{-1} = \Omega_{u\cdot v}\begin{pmatrix} (1/6)\Omega_{vv} & 0 & (3/8)\Omega_{vv}^2 \\ 0 & (5/12)\Omega_{vv}^2 & 0 \\ (3/8)\Omega_{vv}^2 & 0 & (39/20)\Omega_{vv}^3 \end{pmatrix}^{-1}. \tag{32}$$

The result in (32) of course relates to existing results for linear panel cointegration. In case of a scalar regressor as discussed here, the limiting variance is given by $6\Omega_{u\cdot v}\Omega_{vv}^{-1}$, with similar results contained in Pedroni (2000), for a standardized version of the estimator, and in Phillips and Moon (1999, Section 6, p. 1090–1091).

4 Individual and Time Effects

In this section we turn to the case when both individual and time effects are present. It is convenient for the analysis of the asymptotic behavior of the three considered estimators to express quantities in differences to the corresponding quantities arising in the one-way model, i.e.,

$$\check{X}_{it} = \tilde{X}_{it} - V_{NTt}, \tag{33}$$

with $V_{NTt} = \frac{1}{N}\sum_{i=1}^N (X_{it} - \bar{X}_{i.}) = \frac{1}{N}\sum_{i=1}^N \tilde{X}_{it}$,

$$\check{u}_{it} = u_{it} - \bar{u}_{i.} - \bar{u}_{.t} + \bar{u} = \tilde{u}_{it} - w_{NTt}, \tag{34}$$

with $w_{NTt} = \frac{1}{N}\sum_{i=1}^N (u_{it} - \bar{u}_{i.}) = \frac{1}{N}\sum_{i=1}^N \tilde{u}_{it}$.

4.1 The OLS Estimator

In the two-way effects model, the OLS (LSDV) estimator is given by

$$\check{\beta} = (\sum_{i=1}^N \sum_{t=1}^T \check{X}_{it}\check{X}'_{it})^{-1} \sum_{i=1}^N \sum_{t=1}^T \check{X}_{it}\check{y}_{it}. \tag{35}$$

The asymptotic behavior of the LSDV estimator is quite similar to the one-way effects case and as before, in order to state its asymptotic behavior, we need to define the corresponding V-matrix first:

$$\begin{aligned} V_2 &= \mathbb{E}(D_i M D_i) - \mathrm{diag}(0, (1/12)(\mathbb{E}\Omega_{vvi})^2, 0)' \\ &= V_1 - \mathrm{diag}(0, (1/12)(\mathbb{E}\Omega_{vvi})^2, 0)'. \end{aligned} \tag{36}$$

It is interesting to note that V_2 and V_1 in fact only differ in the $(2,2)$-element. This difference stems from the additional averaging across individual units in the two-way model. This averaging generates an additional term in the variance covariance matrix of the limiting distribution given by the expected value of $\tilde{\mathbf{B}}_{vi}(r)$ times its transpose (related to V_{NTt} defined above). Since, the first and third element of $\mathbb{E}\tilde{\mathbf{B}}_{vi}(r)$ are equal to zero, only the (2,2)-element changes between V_1 and V_2.[6]

With the necessary quantities defined, we are now in a position to give the result for the LSDV estimator in the two-way model.

Theorem 4. *Under Assumption 1,* $\mathbb{E}|\Omega_{uui}| < \infty$, $\mathbb{E}|\Omega_{vvi}|^3 < \infty$, $\mathbb{E}|\Omega_{vvi}^{-1/2}\Omega_{uvi}| < \infty$, $\mathbb{E}|\Delta_{vui}| < \infty$, $\mathbb{E}|\Delta_{vui}\Omega_{vvi}^{1/2}| < \infty$, $\mathbb{E}|\Delta_{vui}\Omega_{vvi}| < \infty$, *and for* $p = 1, 2, 3$, $\mathbb{E}|\Omega_{uui}\Omega_{vvi}^p|^{1/2} < \infty$ *and* $\mathbb{E}|\Omega_{uvi}\Omega_{vvi}^{p/2-1/2}| < \infty$*, it holds that*

$$G_T^{-1}(\check{\beta} - \beta) \overset{p, N\to\infty, T\to\infty}{\longrightarrow} V_2^{-1} \begin{pmatrix} -(1/2)\mathbb{E}(\Omega_{uvi}) + \mathbb{E}(\Delta_{vui}) \\ 0 \\ -\mathbb{E}(\Omega_{vvi}\Omega_{uvi}) + (3/2)\mathbb{E}(\Omega_{uvi}\Delta_{vui}) \end{pmatrix}. \tag{37}$$

The second component of the bias term is unchanged compared to the one-way case discussed in Theorem 1 and the first, as discussed, differs *prior to inversion* only in its (2,2)-element.

[6]Such a pattern also occurs when considering higher order polynomials, driven by the zero odd (cross-)moments of normal distributions.

4.2 A Modified OLS Estimator

Parallelling the structure of the discussion for the one-way effects model we now define the correction terms for the modified estimator, which depends upon $\tilde{C}_i$ just as in the one-way effects model, that is

$$\check{\beta}^m = (\sum_{i=1}^N \sum_{t=1}^T \check{X}_{it} \check{X}_{it}')^{-1} \sum_{i=1}^N (\sum_{t=1}^T \check{X}_{it} \check{y}_{it} - \tilde{C}_i). \tag{38}$$

Its limiting distribution is given in the following theorem.

Theorem 5. *Let Assumptions 1, 2 and 3 hold. Assume that* $\mathbb{E}|\Omega_{uui}| < \infty$, $\mathbb{E}|\Omega_{vvi}|^3 < \infty$, $\mathbb{E}|\Delta_{vui}| < \infty$, $\mathbb{E}|\Delta_{vui}\Omega_{vvi}^{1/2}| < \infty$, $\mathbb{E}|\Delta_{vui}\Omega_{vvi}| < \infty$, $\mathbb{E}|\Omega_{vvi}^{-1/2}\Omega_{uvi}| < \infty$, $\mathbb{E}|\Omega_{uui}\Omega_{vvi}^3| < \infty$, *and for* $p = 1, 2, 3$, $\mathbb{E}|\Omega_{uvi}\Omega_{vvi}^{(p-1)/2}| < \infty$ *and* $\mathbb{E}|\Omega_{uui}\Omega_{vvi}^p|^{1/2} < \infty$. *Then it holds that*

$$N^{1/2} G_T^{-1} (\check{\beta}^m - \beta) \overset{d, N\to\infty, T\to\infty}{\longrightarrow} N(0, V_2^{-1}\Sigma_2 V_2^{-1}), \tag{39}$$

where

$$\begin{aligned} \Sigma_2 &= \mathbb{E}\left(\int_0^1 \tilde{\mathbf{B}}_{vi}(r) dB_{ui}(r) - \mathbb{E}(\int_0^1 \tilde{\mathbf{B}}_{vi}(r) dB_{ui}(r)|\Omega_i) - \int_0^1 \mathbb{E}\tilde{\mathbf{B}}_{vi}(r) dB_{ui}(r)\right) \\ &\quad \times \left(\int_0^1 \tilde{\mathbf{B}}_{vi}(r) dB_{ui}(r) - \mathbb{E}(\int_0^1 \tilde{\mathbf{B}}_{vi}(r) dB_{ui}(r)|\Omega_i) - \int_0^1 \mathbb{E}\tilde{\mathbf{B}}_{vi}(r) dB_{ui}(r)\right)' \\ &= \Sigma_1 - (1/6)\mathrm{diag}(0, \mathbb{E}(\Omega_{uui}\Omega_{vvi})\mathbb{E}(\Omega_{vvi}), 0) + (1/12)\mathrm{diag}(0, \mathbb{E}(\Omega_{uui})(\mathbb{E}(\Omega_{vvi}))^2, 0). \end{aligned} \tag{40}$$

4.3 A Fully Modified OLS Estimator

Finally, again as in the one-way effects model we define a classical FM-OLS estimator, which uses the correspondingly transformed dependent variable $\check{y}_{it}^+ = \check{y}_{it} - \hat{\Omega}_{uv}\hat{\Omega}_{vv}^{-1} v_{it}$. The FM-OLS estimator for the two-way effects model is defined as

$$\check{\beta}^+ = (\sum_{i=1}^N \sum_{t=1}^T \check{X}_{it} \check{X}_{it}')^{-1} \sum_{i=1}^N (\sum_{t=1}^T \check{X}_{it} \check{y}_{it}^+ - \tilde{C}_i^+). \tag{41}$$

Theorem 6. *Under the same assumptions as in Theorem 5, plus* $\mathbb{E}|\Delta_{vvi}| < \infty$, $\mathbb{E}|\Delta_{vvi}\Omega_{vvi}^{1/2}| < \infty$ *and* $\mathbb{E}|\Delta_{vvi}\Omega_{vvi}| < \infty$, *it holds that*

$$N^{1/2} G_T^{-1} (\check{\beta}^+ - \beta) \overset{d, N\to\infty, T\to\infty}{\longrightarrow} N\left(0, V_2^{-1}\Sigma_2^+ V_2^{-1}\right) \tag{42}$$

where

$$\Sigma_2^+ = \Sigma_1^+ - \mathrm{diag}(0, (1/6)\mathbb{E}(\Omega_{vvi})\mathbb{E}(\Omega_{u\cdot v,i}\Omega_{vvi}), 0) + \mathrm{diag}(0, (1/12)\mathbb{E}(\Omega_{u\cdot v,i})(\mathbb{E}(\Omega_{vvi}))^2, 0).$$

5 Finite Sample Performance

We now turn to a brief simulation investigation of the developed estimators and consider for brevity only the cubic case with two-way effects, i.e., our most general setting.[7] The data are given by

$$y_{it} = \alpha_i + \gamma_t + \beta_1 x_{it} + \beta_2 x_{it}^2 + \beta_3 x_{it}^3 + u_{it}, \tag{43}$$

where u_{it} and $v_{it} = \Delta x_{it}$ are generated as:

$$\begin{aligned} u_{it} &= \rho_{1i} u_{i,t-1} + \varepsilon_{it} + \rho_{2i}\eta_{it}, \quad u_{i0} = 0, \\ v_{it} &= \nu_{it} + 0.5\nu_{i,t-1}, \end{aligned}$$

with $(\varepsilon_{it}, \nu_{it})' \sim \mathcal{N}(0, I_2)$ cross-sectionally independent. The parameters ρ_{1i} control the level of serial correlation in the error terms u_{it}, and ρ_{2i} control the extent of regressor endogeneity. The parameters ρ_{i1}, ρ_{i2} are cross-sectionally i.i.d. distributed and are independent of $(\varepsilon_{it}, \nu_{it})'$. In particular we consider $\rho_{1i} = \rho_1 + \mathcal{U}_{1i}$ and $\rho_{2i} = \rho_2 + \mathcal{U}_{2i}$ with $\mathcal{U}_{1i}, \mathcal{U}_{2i}$ independently and identically distributed uniform random variables over the interval $[-0.05, 0.05]$, with $\rho_1, \rho_2 \in \{0, 0.3, 0.6, 0.8\}$.[8] The slope parameters are chosen as $\beta_1 = 5$, $\beta_2 = -3$ and $\beta_3 = 0.3$. The individual effects α_i are i.i.d. $\mathcal{N}(0,1)$ and the time effects are set to $\gamma_t = t$, i.e., we simply consider a common linear trend.

For the construction of the modified and fully modified estimators, consistent estimators of the long run variances and half long run variances are required. Based upon those, different variants of the estimators (and test statistics based upon them) are conceivable, as discussed in the theory sections. We have compared several versions in preliminary simulations and focus for brevity on the best performing variants here only. Estimator performance, in terms of bias and root mean squared error (RMSE), is improved by using cross-sectional averages, i.e., the modified OLS estimator as defined in (38) is implemented using $\tilde{C}_i$ as given in (22) with the individual specific estimates $\hat{\Delta}_{vui}$, $\hat{\Omega}_{uvi}$ and $\hat{\Omega}_{vvi}$ replaced for each cross-section member by the cross-sectional average, e.g., $\hat{\Delta}_{vu} = \frac{1}{N}\sum_{i=1}^N \hat{\Delta}_{vui}$ and similarly for the other quantities appearing. The FM-OLS estimator as defined in (41) is implemented using $\tilde{C}_i^+$ as defined in (28) using $\hat{\Delta}_{vu} = \frac{1}{N}\sum_{i=1}^N \hat{\Delta}_{vui}$ for all $i = 1, \ldots, N$ instead of $\hat{\Delta}_{vui}$. The modified dependent variable is used in the form $\check{y}_{it}^+ = \check{y}_{it} - \hat{\Omega}_{uv}\hat{\Omega}_{vv}^{-1} v_{it}$, i.e., again with the cross-sectional average of the long run variances used for all cross-section members. All long run variances are estimated using the Bartlett kernel and the Andrews (1991) bandwidth selection rule. The results are very similar for other choices, e.g., the Quadratic Spectral kernel or the Newey and West (1994) bandwidth rule. The sample sizes considered are all combinations of $T = 50, 100, 200$

[7] The results are very similar for the quadratic case.

[8] The addition of cross-sectionally i.i.d. random variables to the coefficients ρ_1 and ρ_2 is obviously a simple way of generating data in a random linear process fashion. Considering non-random ρ_{1i} and ρ_{2i} leads to very similar results in the simulations.

and $N = 10, 25, 50, 100$. For each setting the number of replications is 5,000 and all test decisions are performed at the nominal 5% level.

For hypothesis testing we consider one variant for the modified estimator and two variants for the fully modified estimator. In particular, for the modified OLS estimator, we consider $\hat{\Sigma}_{\breve{\beta}^m\breve{\beta}^m} = \frac{1}{N}G_T\hat{V}_2^{-1}\hat{\Sigma}_2\hat{V}_2^{-1}G_T$, with

$$\hat{V}_2 = \frac{1}{N}\sum_{i=1}^{N}\hat{D}_iM\hat{D}_i - \text{diag}\left(0, \frac{1}{12}\hat{\Omega}_{vv}^2, 0\right), \tag{44}$$

with $\hat{D}_i = \text{diag}(\hat{\Omega}_{vvi}^{1/2}, \hat{\Omega}_{vvi}, \hat{\Omega}_{vvi}^{3/2})$ and M as given in (19) and

$$\begin{aligned}\hat{\Sigma}_2 = &\frac{1}{N}\sum_{i=1}^{N}\hat{\Omega}_{u\cdot v,i}\hat{D}_iM\hat{D}_i + \frac{1}{N}\sum_{i=1}^{N}\hat{\Omega}_{uvi}^2\hat{\Omega}_{vvi}^{-1}\hat{D}_iQ\hat{D}_i- \\ &-\frac{1}{N}\sum_{i=1}^{N}\begin{pmatrix}\frac{1}{4}\hat{\Omega}_{uvi}^2 & 0 & \frac{1}{2}\hat{\Omega}_{vvi}^2\hat{\Omega}_{uvi}^2 \\ 0 & 0 & 0 \\ \frac{1}{2}\hat{\Omega}_{vvi}^2\hat{\Omega}_{uvi}^2 & 0 & \hat{\Omega}_{vvi}^2\hat{\Omega}_{uvi}^2\end{pmatrix} - \frac{1}{6}\text{diag}\left(0, \left(\frac{1}{N}\sum_{i=1}^{N}\hat{\Omega}_{uui}\hat{\Omega}_{vvi}\right)\hat{\Omega}_{vv}, 0\right) + \\ &+\frac{1}{12}\text{diag}\left(0, \hat{\Omega}_{uu}\hat{\Omega}_{vv}^2, 0\right),\end{aligned} \tag{45}$$

with Q as defined in (23). We consider two estimators of the covariance matrices of the FM-OLS estimator $\breve{\beta}^+$. The first is given by

$$\hat{\Sigma}_{\breve{\beta}^+\breve{\beta}^+} = \frac{1}{N}G_T\hat{V}_2^{-1}\hat{\Sigma}_2^+\hat{V}_2^{-1}G_T, \tag{46}$$

with

$$\hat{\Sigma}_2^+ = \frac{1}{N}\sum_{i=1}^{N}\hat{\Omega}_{u\cdot v,i}\hat{D}_iM\hat{D}_i - \frac{1}{6}\text{diag}\left(0, \hat{\Omega}_{vv}\frac{1}{N}\sum_{i=1}^{N}\hat{\Omega}_{u\cdot v,i}\hat{\Omega}_{vvi}, 0\right) + \frac{1}{12}\text{diag}\left(0, \hat{\Omega}_{u\cdot v}\hat{\Omega}_{vv}^2, 0\right). \tag{47}$$

The second variant is based on a "standard expression" for an FM-OLS variance covariance matrix, i.e.,

$$\hat{\Sigma}_{\breve{\beta}^+\breve{\beta}^+}^{std} = \hat{\Omega}_{u\cdot v}\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\breve{X}_{it}\breve{X}_{it}'\right)^{-1}, \tag{48}$$

compare also Remark 1.

We start with assessing estimator performance measured by bias and RMSE, where we include as a benchmark also the OLS estimator, labelled $\breve{\beta}$, in the discussion and in Tables 1 to 4. We mostly focus on β_1 and β_2, since the results for β_3 are similar to the results for either β_1 or β_2, depending upon question considered.[9] For β_1 (and qualitatively very similarly for β_3), see Table 1,

[9]Additional simulation results are available upon request.

the modified OLS estimator leads to the smallest bias, with the differences to the fully modified OLS estimator often negligible. The exception is, of course $\rho_1 = \rho_2 = 0$, where the OLS estimator, that is in this case second order bias free and by construction tuning parameter free, exhibits the best performance. For β_2 (Table 2) the best performance is, to a certain extent surprisingly, exhibited by the OLS estimator and the FM-OLS estimator, with the differences between the two estimators often very small. Here the exception is $N = 50$, where for the larger values of T the modified OLS estimator has the smallest bias.

Table 1: Bias of estimators of β_1

	$N=10$			$N=25$			$N=50$			$N=100$		
$\rho_1=\rho_2$	$\check{\beta}_1$	$\check{\beta}_1^m$	$\check{\beta}_1^+$	$\check{\beta}_1$	$\check{\beta}_1^m$	$\check{\beta}_1^+$	$\check{\beta}_1$	$\check{\beta}_1^m$	$\check{\beta}_1^+$	$\check{\beta}_1$	$\check{\beta}_1^m$	$\check{\beta}_1^+$
$T=50$												
0	-0.000	-0.000	0.000	-0.000	-0.000	-0.000	0.000	-0.000	-0.000	0.000	0.000	0.000
0.3	0.017	-0.001	0.005	0.016	0.002	0.004	0.015	0.003	0.003	0.015	0.003	0.003
0.6	0.070	0.015	0.032	0.066	0.023	0.028	0.065	0.025	0.026	0.064	0.025	0.025
0.8	0.185	0.084	0.116	0.178	0.095	0.105	0.175	0.097	0.102	0.172	0.098	0.100
$T=100$												
0	-0.000	-0.000	-0.000	0.000	-0.000	-0.000	-0.000	-0.000	-0.000	0.000	0.000	0.000
0.3	0.008	-0.002	0.002	0.008	0.000	0.001	0.008	0.001	0.001	0.008	0.001	0.001
0.6	0.036	0.003	0.013	0.035	0.008	0.011	0.034	0.009	0.010	0.033	0.010	0.010
0.8	0.103	0.030	0.053	0.100	0.039	0.047	0.098	0.041	0.045	0.096	0.042	0.043
$T=200$												
0	-0.000	-0.000	-0.000	0.000	0.000	0.000	-0.000	-0.000	-0.000	0.000	0.000	0.000
0.3	0.004	-0.001	0.001	0.004	0.000	0.001	0.004	0.000	0.000	0.004	0.000	0.000
0.6	0.018	-0.001	0.005	0.018	0.003	0.004	0.017	0.003	0.004	0.017	0.004	0.004
0.8	0.054	0.007	0.022	0.053	0.014	0.019	0.051	0.016	0.018	0.051	0.017	0.018

With respect to RMSE, Table 3 shows that for β_1 the fully modified OLS estimator generally exhibits the best performance, followed by the modified OLS estimator. The differences between these two tend to become very small as N increases. For β_2 (Table 4) the ordering is different, with the fully modified estimator exhibiting the best performance and the OLS estimator coming in second. The modified estimator leads to partly substantially larger RMSE than the other two estimators in case of β_2, especially for small N. *Grosso modo*, from a bias and RMSE perspective the fully modified OLS estimator is the best choice, especially given that the performance of the modified OLS estimator is comparably quite poor for β_2.[10]

Let us now turn to test performance, where we start by considering the empirical null rejection

[10] Note for completeness that for $N = 100$ the modified OLS estimator leads to smallest bias and RMSE for β_3.

Table 2: Bias ($\times 10^4$) of estimators of β_2

	$N=10$			$N=25$			$N=50$			$N=100$		
$\rho_1=\rho_2$	$\check{\beta}_2$	$\check{\beta}_2^m$	$\check{\beta}_2^+$	$\check{\beta}_2$	$\check{\beta}_2^m$	$\check{\beta}_2^+$	$\check{\beta}_2$	$\check{\beta}_2^m$	$\check{\beta}_2^+$	$\check{\beta}_2$	$\check{\beta}_2^m$	$\check{\beta}_2^+$
						$T=50$						
0	-0.009	-0.170	-0.018	0.091	0.130	0.103	0.142	0.145	0.144	0.002	0.003	0.002
0.3	0.050	-0.784	0.065	0.153	0.283	0.163	0.199	0.162	0.197	-0.001	0.027	0.004
0.6	0.283	-1.650	0.371	0.329	0.632	0.335	0.355	0.229	0.348	0.010	0.090	0.017
0.8	0.864	-2.689	0.980	0.703	1.167	0.697	0.657	0.397	0.662	0.025	0.178	0.039
						$T=100$						
0	-0.032	-0.097	-0.029	0.014	0.017	0.016	0.023	0.021	0.023	0.020	0.019	0.020
0.3	-0.007	-0.114	-0.028	0.027	0.021	0.020	0.023	0.042	0.027	0.028	0.030	0.027
0.6	0.091	0.031	0.018	0.037	0.019	0.009	0.026	0.091	0.040	0.049	0.053	0.046
0.8	0.344	0.467	0.195	-0.037	-0.042	-0.099	0.041	0.185	0.071	0.092	0.098	0.086
						$T=200$						
0	-0.001	0.016	0.002	-0.003	-0.002	-0.003	-0.002	-0.002	-0.002	0.002	0.002	0.002
0.3	0.011	0.079	0.004	0.000	0.026	-0.002	-0.004	-0.002	-0.003	0.002	-0.001	0.003
0.6	0.047	0.249	0.017	0.007	0.094	-0.000	-0.010	-0.001	-0.006	0.003	-0.009	0.006
0.8	0.102	0.520	0.031	0.026	0.230	0.011	-0.016	0.005	-0.007	0.006	-0.023	0.010

probabilities of t-tests for β_1 in Table 5 and β_2 in Table 6. The most pervasive finding is that the test based on the FM-OLS estimator with the estimated variance given by (48) leads to the best performance (smallest size distortions) in most cases. In some cases, when N is large and ρ_1 and ρ_2 are large, the t-test based on the modified OLS estimator leads to the smallest size distortions. For β_2 and small N the modified OLS estimator based test is outperformed by both variants of FM-OLS based tests. When comparing the findings for β_1 and β_2 (and β_3) a striking feature is the *size divergence* observed for testing hypothesis concerning β_1. For N large relative to T, the null rejection probabilities tend to one. This phenomenon has been found to be widespread in panel unit root and cointegration testing, see, e.g., Hlouskova and Wagner (2006) and Wagner and Hlouskova (2009). The faster convergence rate of the estimators of β_2 and β_3 ameliorate this problem. In this respect it is a surprising observation that size divergence is less present for β_2 than for β_3, for which the coefficient estimators converge faster than for β_2. When the simulations are performed using true rather than estimated half long run and long run variances the test performances are throughout much better, which identifies, as is well-known in the literature, long run variance estimation as a man culprit for poor performance.

We close this section with a brief look on size corrected power, where we consider all three coefficients. The rejection probabilities are calculated using the empirical critical values from the

Table 3: RMSE of estimators of β_1

	$N=10$			$N=25$			$N=50$			$N=100$		
$\rho_1=\rho_2$	$\check{\beta}_1$	$\check{\beta}_1^m$	$\check{\beta}_1^+$	$\check{\beta}_1$	$\check{\beta}_1^m$	$\check{\beta}_1^+$	$\check{\beta}_1$	$\check{\beta}_1^m$	$\check{\beta}_1^+$	$\check{\beta}_1$	$\check{\beta}_1^m$	$\check{\beta}_1^+$
$T=50$												
0	0.018	0.019	0.019	0.010	0.010	0.010	0.007	0.007	0.007	0.005	0.005	0.005
0.3	0.030	0.030	0.026	0.021	0.015	0.015	0.018	0.010	0.010	0.016	0.008	0.007
0.6	0.082	0.062	0.054	0.071	0.035	0.037	0.067	0.030	0.031	0.065	0.028	0.028
0.8	0.201	0.136	0.141	0.184	0.107	0.115	0.178	0.102	0.107	0.174	0.100	0.102
$T=100$												
0	0.009	0.009	0.009	0.005	0.005	0.005	0.003	0.003	0.003	0.002	0.002	0.002
0.3	0.015	0.017	0.013	0.011	0.008	0.007	0.009	0.005	0.005	0.008	0.004	0.004
0.6	0.043	0.037	0.026	0.037	0.017	0.017	0.035	0.013	0.013	0.034	0.012	0.011
0.8	0.114	0.079	0.069	0.104	0.048	0.053	0.099	0.045	0.048	0.097	0.044	0.045
$T=200$												
0	0.004	0.005	0.005	0.002	0.003	0.003	0.002	0.002	0.002	0.001	0.001	0.001
0.3	0.008	0.008	0.006	0.005	0.004	0.004	0.004	0.003	0.002	0.004	0.002	0.002
0.6	0.022	0.019	0.012	0.019	0.008	0.008	0.018	0.006	0.006	0.017	0.005	0.005
0.8	0.060	0.042	0.032	0.055	0.021	0.023	0.052	0.019	0.020	0.051	0.018	0.019

t-test null simulations with data generated under the alternative using an equidistant grid of 21 points for the parameter values (including also the null parameter values). Reflecting the different convergence rates we consider for β_1 the interval $[5, 5.08]$, for β_2 the interval $[-3, -2.996]$ and for β_3 the interval $[0.3, 0.3002]$. Figures 1 to 3 display the results for $T = 100$ all values of N and $\rho_1 = \rho_2 = 0.6$. Size corrected power is lowest for the t-test based on the modified estimator, the two variants based on the FM-OLS estimator perform relatively similarly, with some visible performance advantages of the "standard" version in case of small N. With increasing N, the performance disadvantages of the modified OLS estimator based test diminish, reflecting the fact that the modified OLS estimator crucially rests upon a cross-sectional limit in the bias correction step. Combining size and power behavior we conclude that the "standard" variant of the FM-OLS based tests leads to the best performance.

Table 4: RMSE ($\times 10^4$) of estimators of β_2

	$N=10$			$N=25$			$N=50$			$N=100$		
$\rho_1=\rho_2$	$\check{\beta}_2$	$\check{\beta}_2^m$	$\check{\beta}_2^+$	$\check{\beta}_2$	$\check{\beta}_2^m$	$\check{\beta}_2^+$	$\check{\beta}_2$	$\check{\beta}_2^m$	$\check{\beta}_2^+$	$\check{\beta}_2$	$\check{\beta}_2^m$	$\check{\beta}_2^+$
						$T=50$						
0	14.851	16.784	14.864	6.173	6.263	6.183	3.787	3.802	3.789	2.441	2.443	2.442
0.3	20.854	39.644	20.375	8.747	10.955	8.560	5.346	5.883	5.233	3.513	3.669	3.391
0.6	36.837	100.570	33.303	15.571	24.660	14.262	9.457	11.749	8.667	6.392	6.993	5.683
0.8	66.161	179.831	58.813	29.020	45.185	26.072	17.732	21.503	15.899	12.178	12.900	10.580
						$T=100$						
0	5.342	5.927	5.371	2.186	2.215	2.191	1.328	1.330	1.328	0.841	0.842	0.842
0.3	7.652	18.295	7.498	3.132	4.096	3.071	1.905	2.165	1.866	1.226	1.307	1.188
0.6	14.065	49.865	12.632	5.759	10.041	5.259	3.497	4.661	3.187	2.306	2.645	2.049
0.8	27.665	99.151	23.986	11.796	20.848	10.387	7.160	9.582	6.273	4.814	5.418	4.090
						$T=200$						
0	1.842	1.942	1.847	0.771	0.775	0.771	0.470	0.471	0.470	0.300	0.300	0.300
0.3	2.648	6.803	2.601	1.115	1.484	1.093	0.680	0.772	0.667	0.443	0.472	0.426
0.6	5.053	20.954	4.480	2.082	3.765	1.899	1.265	1.727	1.155	0.846	0.995	0.743
0.8	10.922	46.609	9.020	4.447	8.457	3.834	2.692	3.791	2.323	1.834	2.164	1.512

Table 5: Empirical null rejection probabilities of t-tests for β_1

	$N=10$			$N=25$			$N=50$			$N=100$		
$\rho_1=\rho_2$	$\check{\beta}_1^m$	$\check{\beta}_1^+$	$\check{\beta}_1^{++}$	$\check{\beta}_1^m$	$\check{\beta}_1^+$	$\check{\beta}_1^{++}$	$\check{\beta}_1^m$	$\check{\beta}_1^+$	$\check{\beta}_1^{++}$	$\check{\beta}_1^m$	$\check{\beta}_1^+$	$\check{\beta}_1^{++}$
						$T=50$						
0	0.194	0.205	0.094	0.131	0.143	0.084	0.110	0.121	0.084	0.093	0.104	0.076
0.3	0.242	0.237	0.128	0.155	0.171	0.118	0.140	0.162	0.124	0.147	0.165	0.137
0.6	0.329	0.384	0.263	0.316	0.443	0.366	0.462	0.591	0.537	0.706	0.803	0.774
0.8	0.512	0.688	0.567	0.731	0.877	0.831	0.936	0.985	0.973	0.997	1.000	1.000
						$T=100$						
0	0.180	0.178	0.094	0.111	0.115	0.077	0.091	0.094	0.072	0.084	0.091	0.074
0.3	0.241	0.207	0.111	0.141	0.140	0.096	0.116	0.120	0.091	0.117	0.119	0.106
0.6	0.306	0.309	0.192	0.234	0.317	0.249	0.296	0.400	0.354	0.475	0.570	0.540
0.8	0.403	0.544	0.417	0.509	0.722	0.646	0.762	0.904	0.880	0.951	0.992	0.988
						$T=200$						
0	0.163	0.167	0.081	0.101	0.105	0.069	0.075	0.077	0.062	0.072	0.074	0.065
0.3	0.217	0.183	0.097	0.121	0.126	0.084	0.096	0.095	0.074	0.093	0.096	0.086
0.6	0.274	0.248	0.143	0.176	0.233	0.172	0.203	0.268	0.228	0.297	0.381	0.352
0.8	0.320	0.414	0.286	0.338	0.542	0.463	0.513	0.721	0.674	0.804	0.924	0.910

Table 6: Empirical null rejection probabilities of t-tests for β_2

	$N=10$			$N=25$			$N=50$			$N=100$		
$\rho_1=\rho_2$	$\check{\beta}_2^m$	$\check{\beta}_2^+$	$\check{\beta}_2^{++}$	$\check{\beta}_2^m$	$\check{\beta}_2^+$	$\check{\beta}_2^{++}$	$\check{\beta}_2^m$	$\check{\beta}_2^+$	$\check{\beta}_2^{++}$	$\check{\beta}_2^m$	$\check{\beta}_2^+$	$\check{\beta}_2^{++}$
$T=50$												
0	0.308	0.335	0.088	0.163	0.198	0.079	0.113	0.147	0.074	0.085	0.113	0.072
0.3	0.355	0.351	0.109	0.190	0.216	0.098	0.127	0.161	0.092	0.090	0.129	0.093
0.6	0.398	0.380	0.126	0.205	0.252	0.123	0.122	0.181	0.111	0.079	0.156	0.113
0.8	0.420	0.443	0.161	0.211	0.306	0.153	0.128	0.242	0.142	0.087	0.212	0.144
$T=100$												
0	0.297	0.312	0.079	0.151	0.168	0.069	0.099	0.117	0.062	0.067	0.086	0.059
0.3	0.357	0.332	0.095	0.188	0.184	0.083	0.113	0.135	0.078	0.075	0.098	0.074
0.6	0.401	0.361	0.108	0.211	0.215	0.098	0.129	0.160	0.094	0.074	0.125	0.090
0.8	0.422	0.404	0.132	0.219	0.268	0.121	0.129	0.199	0.109	0.077	0.163	0.110
$T=200$												
0	0.285	0.295	0.072	0.149	0.158	0.059	0.097	0.106	0.059	0.070	0.081	0.057
0.3	0.362	0.310	0.087	0.179	0.174	0.071	0.111	0.121	0.072	0.077	0.093	0.073
0.6	0.404	0.335	0.097	0.212	0.196	0.081	0.126	0.142	0.082	0.084	0.116	0.083
0.8	0.422	0.366	0.108	0.223	0.227	0.091	0.131	0.170	0.087	0.082	0.136	0.093

Figure 1: Size corrected power of t-tests for β_1 for $T = 100$, $\rho_1, \rho_2 = 0.6$ and all values of N

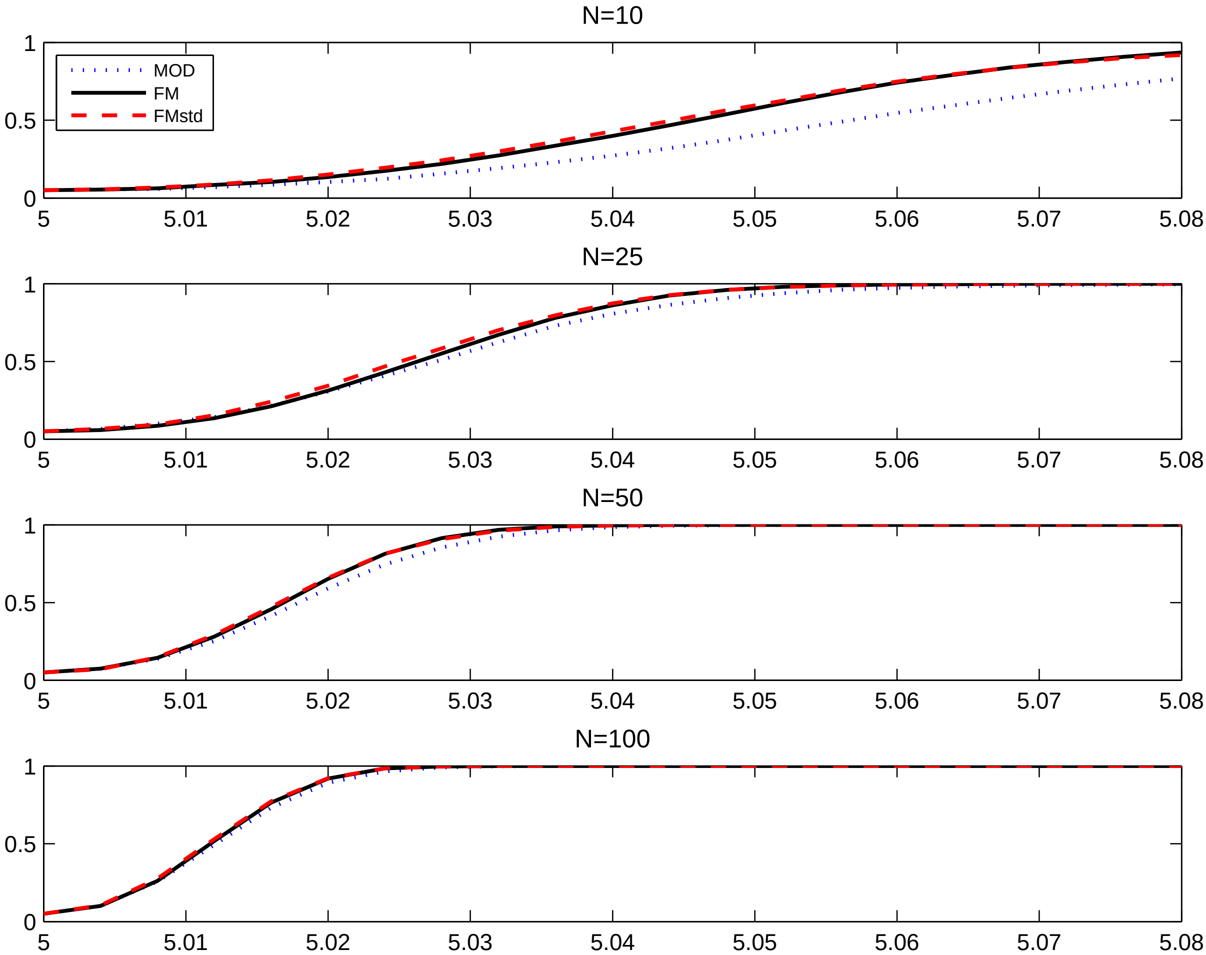


Figure 2: Size corrected power of t-tests for β_2 for $T = 100$, $\rho_1, \rho_2 = 0.6$ and all values of N

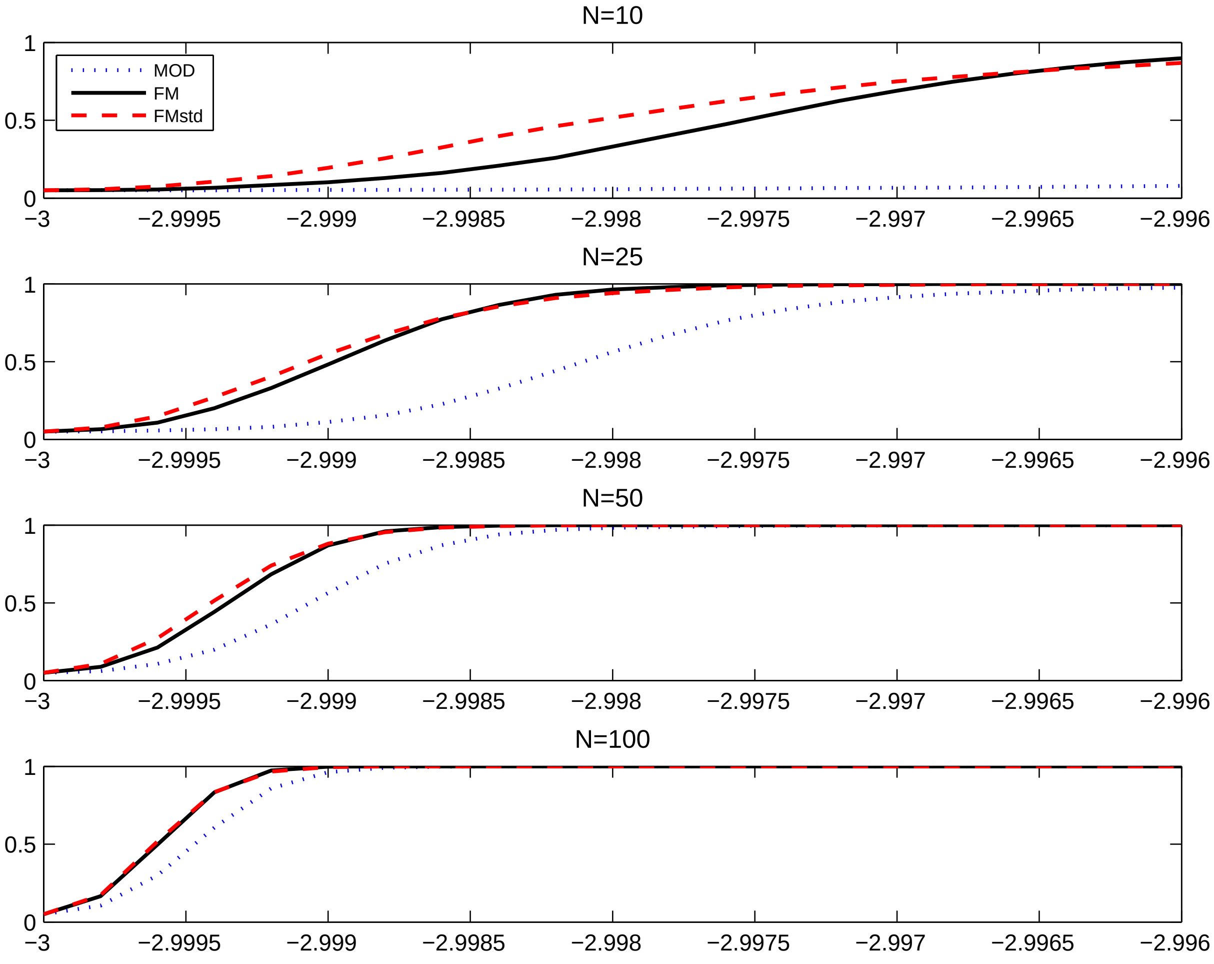
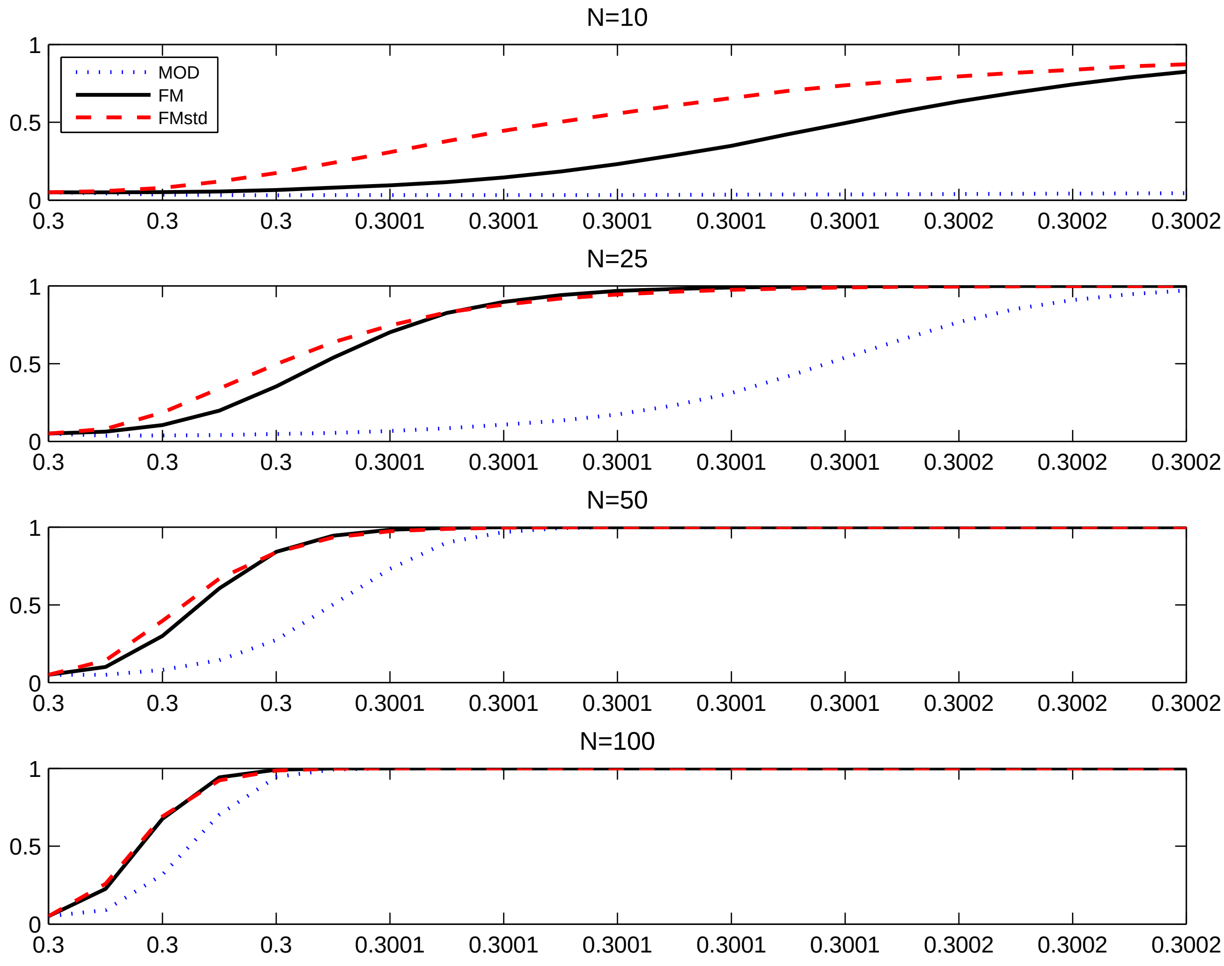


Figure 3: Size corrected power of t-tests for β_3 for $T = 100$, $\rho_1, \rho_2 = 0.6$ and all values of N



6 The Environmental Kuznets Curve

In this section we briefly illustrate the developed methods by estimating EKCs for CO_2 emissions. The dependent variable is the logarithm of CO_2 emissions per capita and the explanatory variable is the logarithm of GPD per capita and its powers. We consider both the quadratic and the cubic specification and both the individual effects only (one-way) and the individual and time effects (two-way) specification. All estimation results given in this section are based on the fully modified estimator, with "standard" inference in the terminology of the above finite sample performance section. Also as in the simulations we use the Bartlett kernel and the Andrews (1991) bandwidth selection rule.

We use two data sets, the *long data set* with small $N = 19$ and large $T = 135$ and the *wide data set* with large $N = 89$ and small $T = 54$. The long data set is as in Wagner *et al.* (2018) and comprises data on per capita CO_2 emissions and per capita GDP for the period 1878–2013 for 19 early industrialized countries.[11] We also consider a subset of six out of these 19 countries for the period 1870–2013 as analyzed in Wagner *et al.* (2018) in a seemingly unrelated regressions (SUR) setting. These six countries are Austria, Belgium, Finland, the Netherlands, Switzerland and the United Kingdom.

The estimation results for the long data set (for both $N = 19$ and $N = 6$) are given in Table 7. The sample range for the $N = 6$ country set is 1,725 to 26,102 and for the $N = 19$ country set the range is 794 to 31,933 (measured in 1990 Geary-Khamis Dollars). Let us start with the quadratic specification. Three of the four estimated EKCs (with the exception of the two-way specification and $N = 6$) have significant coefficients with a negative coefficient to squared log GDP per capita and thus correspond to an inverted U shape. For these three estimated EKCs also the turning points are (with one slight exception) inside the sample range and for the $N = 6$ case close to the pooled turning point estimated with SUR methods in Wagner *et al.* (2018). For the six country data set the estimated time effects $\hat{\gamma}_t$ "take out" a very large amount of common variation in the GDP-emissions relationship across these six countries, and therefore this specification appears to a certain extent over-parameterized. The cubic specification leads to mostly insignificant coefficients, except for $N = 6$ with two-way effects. Furthermore, the fitted curves are essentially monotonic over the sample ranges with again especially in the two-way case the inverted U shape taken out by the estimated time effects.

[11] The 19 countries are given by Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom and USA. Note that the data are in fact available from 1870 onwards, with the exception of CO_2 emissions for New Zealand. Thus, when considering all 19 countries we use 1878 as starting point to have a balanced panel for simplicity. A detailed description of the data including the sources is contained in Wagner *et al.* (2018).

Table 7: Estimation results for the long data set.

	Quadratic				Cubic			
	$N = 19$		$N = 6$		$N = 19$		$N = 6$	
	1-way	2-way	1-way	2-way	1-way	2-way	1-way	2-way
β_1	7.972	7.364	6.642	-9.372	20.954	23.638	33.432	195.860
	(9.649)	(5.432)	(3.380)	(-2.381)	(1.864)	(1.676)	(0.922)	(5.505)
β_2	-0.402	-0.362	-0.331	0.645	-1.914	-2.288	-3.352	-23.138
	(-8.536)	(-4.257)	(-3.021)	(2.737)	(-1.463)	(-1.372)	(-0.820)	(-5.653)
β_3					0.058	0.076	0.113	0.917
					(1.155)	(1.153)	(0.738)	(5.854)
TP	20,240	26,141	22,771	1,430	127,784	$--$	$--$	$--$
					28,054			

Note: The turning points (TP) are computed as $\exp\left(-\frac{\beta_1}{2\beta_2}\right)$ in the quadratic case and as $\exp\left(\pm(-\frac{\beta_1}{3\beta_3} + (\frac{\beta_2}{3\beta_3})^2)^{1/2} - \frac{\beta_2}{3\beta_3}\right)$ in the cubic case. t-statistics given in brackets.

The results for the wide data set are given in Table 8. The sample range in this data set is 340 to 100,959 Dollars (here measured in 2015 US-Dollars). Since in the cubic specification all coefficients are significant, we have to consider the quadratic specification to be misspecified and consequently focus on the cubic case. The negative third order coefficient, obviously, means that the larger turning point corresponds to an inverted U shape behavior. At the lower sample end, at around 500 Dollars, a U-shaped turning point occurs (owing to the third order specification). Plotting the estimated time effects for the cubic specification displays a trend with a break towards a smaller but still positive slope around 1970. This may contribute to the smaller estimated turning point in the two-way specification. Altogether, these are reasonable first findings based on our problem adequate estimators for panels of cointegrating polynomial regressions. A detailed empirical analysis is, however, beyond the scope of this paper.

7 Summary and Conclusions

This paper extends the fully modified OLS estimation principle for cointegrating polynomial regressions from the time series case, studied in detail in Wagner and Hong (2016), to the panel case. Following Phillips and Moon (1999) we consider a cross-sectional i.i.d. random linear process framework, however, we only consider a sequential asymptotic framework with T tending to infinity first followed by N tending to infinity. Given that in applications basically only the quadratic and

Table 8: Estimation results for the wide data set.

	Quadratic		Cubic	
	1-way	2-way	1-way	2-way
β_1	2.438	2.292	-13.410	-13.327
	(9.093)	(8.843)	(-6.962)	(-7.042)
β_2	-0.092	-0.102	1.728	1.695
	(-6.299)	(-7.117)	(7.846)	(7.828)
β_3			-0.069	-0.068
			(-8.277)	(-8.320)
TP	531,250	72,329	43,231	29,519
			443	578

Note: See note to Table 7.

cubic formulations are used we also confine ourselves, mostly for notational brevity, to the cubic specification, with either fixed individual effects or fixed individual and time effects.

Utilizing the large cross-sectional dimension we also introduce an additional estimator, labelled modified OLS. This estimator is based on the observation that the large cross-sectional dimension allows to consistently estimate an additive bias term arising in OLS estimation that can hence be removed. Contrary to FM-OLS, no modification of the dependent variable is required. This idea has, to the best of our knowledge, not yet been considered in the (linear) panel cointegration literature.

The simulations indicate that by and large the FM-OLS estimator outperforms the modified OLS estimator. However, especially in case of a large cross-sectional dimension (compared to the time series dimension) the modified estimator leads to the best performance. Based on the mostly superior estimator performance of FM-OLS, it is not surprising that FM-OLS based tests mostly outperform modified OLS based tests, with the differences vanishing with increasing cross-sectional dimension. As has been observed before in the nonstationary panel literature, hypothesis tests are prone to *size divergence* in case of a large cross-sectional dimension compared to the time series dimension. We also observe this behavior, but only when testing hypotheses for the coefficient to the level of integrated process regressor. When testing hypotheses concerning the coefficient for the second or third power of the integrated regressor, size divergence is not an issue. This may well be due to the faster time dimension convergence rate for these coefficients, $T^{3/2}$ and T^2 respectively, compared to T for the coefficient to the integrated regressor itself. A more detailed formal analysis of this finding is on the agenda for future research.

Our first small illustrative application leads to reasonable findings for two (the long and the wide) data sets. This indicates that our estimator is a potentially valuable addition to – not least – the EKC toolkit, where often panel data with a relatively large cross-sectional dimension are used and where consequently the SUR estimators of Wagner *et al.* (2018) cannot be used.

References

Andrews, D.W.K. (1991). Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation. *Econometrica* **59**, 817–858.

Darvas, Z. (2008). Estimation Bias and Inference in Overlapping Autoregressions: Implications for the Target-Zone Literature. *Oxford Bulletin of Economics and Statistics* **70**, 1–22.

de Jong, R. (2002). Nonlinear Estimators with Integrated Regressors but without Exogeneity. Mimeo.

Grossman, G.M. and A.B. Krueger (1993). Environmental Impacts of a North American Free Trade Agreement. In Garber, P. (Ed.) *The Mexico-US Free Trade Agreement*, 13–56, MIT Press, Cambridge.

Hlouskova, J. and M. Wagner (2006). The Performance of Panel Unit Root and Stationarity Tests: Results from a Large Scale Simulation Study. *Econometric Reviews* **25**, 85–116.

Holtz-Eakin, D. and T.M. Selden (1995). Stoking the Fires? CO_2 Emissions and Economic Growth. *Journal of Public Economics* **57**, 85–101.

Ibragimov, R. and P.C.B. Phillips (2008). Regression Asymptotics Using Martingale Convergence Methods. *Econometric Theory* **24**, 888–947.

Kuznets, S. (1955). Economic Growth and Income Inequality. *American Economic Review* **45**, 1–28.

Malenbaum, W. (1978). *World Demand for Raw Materials in 1985 and 2000*. New York: McGraw-Hill, EMJ Mining Information Services.

Newey, W. and K. West (1994). Automatic Lag Selection in Covariance Matrix Estimation. *Review of Economic Studies* **61**, 631–654.

Park, J.Y. and P.C.B. Phillips (2001). Nonlinear Regressions with Integrated Time Series. *Econometrica* **69**, 117–161.

Pedroni, P. (2000). Fully Modified OLS for Heterogeneous Cointegrated Panels. In Baltagi, B.H. (Ed.) *Advances in Econometrics, Vol. 15: Nonstationary Panels, Panel Cointegration, and Dynamic Panels*. Elsevier, Amsterdam, 93–130, Phillips, P.C.B. and B.E. Hansen (1990). Statistical Inference in Instrumental Variables Regression with I(1) Processes. *Review of Economic Studies* **57**, 99–125.

Phillips, P.C.B. and H.R. Moon (1999). Linear Regression Limit Theory for Nonstationary Panel Data. *Econometrica* **67**, 1057–1111.

Pötscher, B.M. (2004). Nonlinear Functions and Convergence to Brownian Motion: Beyond the Continuous Mapping Theorem. *Econometric Theory* **20**, 1–22.

Saikkonen, P. (1991). Asymptotically Efficient Estimation of Cointegrating Regressions. *Econometric Theory* **7**, 1–21.

Stern, D.I. (2004). The Rise and Fall of the Environmental Kuznets Curve. *World Development* **32**, 1419–1439.

Stock, J.H. and M.W. Watson (1993). A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems. *Econometrica* **61**, 783–820.

Stypka, O., M. Wagner, P. Grabarczyk and R. Kawka (2018). The Asymptotic Validity of "Standard" Fully Modified OLS Estimation and Inference in Cointegrating Polynomial Regressions. *Mimeo.*

Svensson, L.E.O. (1992). An Interpretation of Recent Research on Exchange Rate Target Zones. *Journal of Economic Perspectives* **6**, 119–144.

Wagner, M. (2012). The Phillips Unit Root Tests for Polynomials of Integrated Processes. *Economics Letters* **114**, 299–303.

Wagner, M. (2015). The Environmental Kuznets Curve, Cointegration and Nonlinearity. *Journal of Applied Econometrics* **30**, 948–967.

Wagner, M. and J. Hlouskova (2009). The Performance of Panel Cointegration Methods: Results From a Large Scale Simulation Study. *Econometric Reviews* **29**, 182–223.

Wagner, M. and S.H. Hong (2016). Cointegrating Polynomial Regressions: Fully Modified OLS Estimation and Inference. *Econometric Theory* **32**, 1289–1315.

Wagner, M., S.H. Hong and P. Grabarczyk (2018). Fully Modified OLS Estimation and Inference for Seemingly Unrelated Cointegrating Polynomial Regressions and the Environmental Kuznets Curve for Carbon Dioxide Emissions. *Mimeo.*

Yandle, B., M. Bjattarai, and M. Vijayaraghavan (2004). Environmental Kuznets Curves: A Review of Findings, Methods, and Policy Implications. Research Study 02.1 update, PERC.

Appendix A: Proofs of the Main Theorems

Proof of Theorem 1:

First note that $\tilde{\beta}$ satisfies

$$G_T^{-1}(\tilde{\beta}-\beta) = (N^{-1}\sum_{i=1}^{N} G_T \sum_{t=1}^{T} \tilde{X}_{it}\tilde{X}_{it}' G_T)^{-1} N^{-1}\sum_{i=1}^{N} G_T \sum_{t=1}^{T} \tilde{X}_{it}\tilde{u}_{it}.$$

Under Assumption 1,

$$G_T \sum_{t=1}^{T} \tilde{X}_{it}\tilde{X}_{it}' G_T \overset{d,T\to\infty}{\longrightarrow} \int_0^1 \tilde{\mathbf{B}}_{vi}(r)\tilde{\mathbf{B}}_{vi}(r)' dr$$

$$G_T \sum_{t=1}^{T} \tilde{X}_{it}\tilde{u}_{it} \overset{d,T\to\infty}{\longrightarrow} \int_0^1 \tilde{\mathbf{B}}_{vi}(r) dB_{ui}(r) + \Delta_{vui}(1, 2\int_0^1 B_{vi}(r)dr, 3\int_0^1 B_{vi}(r)^2 dr)'.$$

By parts 7, 8 and 10 of Lemma 1, $\mathbb{E}|\int_0^1 \tilde{\mathbf{B}}_{vi}(r)\tilde{\mathbf{B}}_{vi}(r)'dr| < \infty$, $\mathbb{E}|\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)| < \infty$ and $\mathbb{E}|\Delta_{vui}(1, 2\int_0^1 B_{vi}(r)dr, 3\int_0^1 B_{vi}(r)^2dr)| < \infty$, since the conditions for parts 7, 8 and 10 of Lemma 1 were assumed under the conditions of the theorem. Therefore,

$$N^{-1}\sum_{i=1}^{N} G_T \sum_{t=1}^{T} \tilde{X}_{it}\tilde{X}_{it}' G_T \overset{d,N\to\infty,T\to\infty}{\longrightarrow} \mathbb{E}\left(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)\tilde{\mathbf{B}}_{vi}(r)' dr\right) = V_1,$$

$$N^{-1}\sum_{i=1}^{N} G_T \sum_{t=1}^{T} \tilde{X}_{it}\tilde{u}_{it}$$

$$\overset{d,N\to\infty,T\to\infty}{\longrightarrow} \mathbb{E}(\int_0^1 \tilde{\mathbf{B}}_{vi}(r) dB_{ui}(r)) + \mathbb{E}(\Delta_{vui}\begin{pmatrix} 1, 2\int_0^1 B_{vi}(r)dr, 3\int_0^1 B_{vi}(r)^2 dr \end{pmatrix}') = M_2,$$

say. To calculate M_2, note that by Lemma 2, since all moment assumptions for this lemma are assumed,

$$M_2 = \mathbb{E}(\int_0^1 \tilde{\mathbf{B}}_{vi}(r) dB_{ui}(r)) + \mathbb{E}(\Delta_{vui}\begin{pmatrix} 1, 2\int_0^1 B_{vi}(r)dr, 3\int_0^1 B_{vi}(r)^2 dr \end{pmatrix}')$$

$$= \begin{pmatrix} -(1/2)\mathbb{E}(\Omega_{uvi}) \\ 0 \\ -\mathbb{E}(\Omega_{vvi}\Omega_{uvi}) \end{pmatrix} + \begin{pmatrix} \mathbb{E}(\Delta_{vui}) \\ 0 \\ 3\mathbb{E}(\Delta_{vu}\Omega_{vvi})\mathbb{E}(\int_0^1 W_{vi}(r)^2 dr) \end{pmatrix} = \begin{pmatrix} -(1/2)\mathbb{E}(\Omega_{uvi}) + \mathbb{E}(\Delta_{vui}) \\ 0 \\ -\mathbb{E}(\Omega_{vvi}\Omega_{uvi}) + (3/2)\mathbb{E}(\Omega_{uvi}\Delta_{vui}) \end{pmatrix}.$$

This completes the calculation of all values as stated in the theorem, thereby completing the proof. □

Proof of Theorem 2:

Using the definitions of $\tilde{\beta}^{\mathrm{m}}$ and $\tilde{C}_i$, we have

$$N^{1/2}G_T^{-1}(\tilde{\beta}^{\mathrm{m}}-\beta)=(N^{-1}\sum_{i=1}^{N}G_T\sum_{t=1}^{T}\tilde{X}_{it}\tilde{X}_{it}'G_T)^{-1}N^{-1/2}\sum_{i=1}^{N}G_T(\sum_{t=1}^{T}\tilde{X}_{it}\tilde{u}_{it}-\tilde{C}_i).$$

The proof of Theorem 1 shows that under Assumption 1, if the conditions for parts 7, 8 and 10 of Lemma 1 are assumed,

$$N^{-1}\sum_{i=1}^{N}G_T\sum_{t=1}^{T}\tilde{X}_{it}\tilde{X}_{it}'G_T \overset{p,N\to\infty,T\to\infty}{\longrightarrow} V_1.$$

By Assumptions 1, 2 and 3 and the definition of $\tilde{C}_i$,

$$G_T\tilde{C}_i=\hat{\Delta}_{vui}(1,2T^{-3/2}\sum_{t=1}^{T}x_{it},3T^{-2}\sum_{t=1}^{T}x_{it}^2)'+(-(1/2)\hat{\Omega}_{uvi},0,-\hat{\Omega}_{vvi}\hat{\Omega}_{uvi})'$$

$$\overset{d,T\to\infty}{\longrightarrow}\Delta_{vui}\Big(\ 1,2\textstyle\int_0^1 B_{vi}(r)dr,3\int_0^1 B_{vi}(r)^2dr\ \Big)'+(-(1/2)\Omega_{uvi},0,-\Omega_{vvi}\Omega_{uvi}). \tag{49}$$

Now as $T\to\infty$, by Assumption 1, the result of Equation (49), and Lemma 2,

$$G_T(\sum_{t=1}^{T}\tilde{X}_{it}\tilde{u}_{it}-\tilde{C}_i)\overset{d,T\to\infty}{\longrightarrow}\int_0^1\tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)+\Delta_{vui}(1,2\int_0^1 B_{vi}(r)dr,3\int_0^1 B_{vi}(r)^2dr)'$$
$$-(\Delta_{vui}(\ 1,2\textstyle\int_0^1 B_{vi}(r)dr,3\int_0^1 B_{vi}(r)^2dr\ )'+(-(1/2)\Omega_{uvi},0,-\Omega_{vvi}\Omega_{uvi})')$$

$$=\int_0^1\tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)-\mathbb{E}(\int_0^1\tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)|\Omega_i). \tag{50}$$

By part 9 of Lemma 1, the moment conditions of which were assumed, all elements of this vector have a finite second moment. Therefore, by the CLT for i.i.d. distributed random variables with finite variance,

$$N^{-1/2}\sum_{i=1}^{N}G_T(\sum_{t=1}^{T}\tilde{X}_{it}\tilde{u}_{it}-\tilde{C}_i)\overset{d,N\to\infty,T\to\infty}{\longrightarrow}N(0,\Sigma_1)$$

where

$$\Sigma_1=\mathbb{E}(\int_0^1\tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)-\mathbb{E}(\int_0^1\tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)|\Omega_i))(\int_0^1\tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)-\mathbb{E}(\int_0^1\tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)|\Omega_i))'.$$

Therefore,

$$N^{1/2}G_T^{-1}(\tilde{\beta}^{\mathrm{m}}-\beta)\overset{d,n\to\infty,T\to\infty}{\longrightarrow}N(0,V_1^{-1}\Sigma_1V_1^{-1}),$$

and the value of Σ_1 as stated in the theorem is calculated in Lemma 3.

□

Proof of Theorem 3:

By the definition of $\tilde{\beta}^+$,

$$N^{1/2}G_T^{-1}(\tilde{\beta}^+ - \beta) = (N^{-1}\sum_{i=1}^{N} G_T \sum_{t=1}^{T} \tilde{X}_{it}\tilde{X}_{it}' G_T)^{-1} N^{-1/2}\sum_{i=1}^{N} G_T(\sum_{t=1}^{T} \tilde{X}_{it}\tilde{u}_{it} - \sum_{t=1}^{T} \tilde{X}_{it}\hat{\Omega}_{uvi}\hat{\Omega}_{vvi}^{-1} v_{it} - \tilde{C}_i^+).$$

Again, if the conditions for parts 7, 8 and 10 of Lemma 1 are assumed,

$$N^{-1}\sum_{i=1}^{N} G_T \sum_{t=1}^{T} \tilde{X}_{it}\tilde{X}_{it}' G_T \overset{p,N\to\infty,T\to\infty}{\longrightarrow} V_1.$$

By Assumptions 1 and 2, it follows that

$$G_T\tilde{C}_i^+ \overset{d,T\to\infty}{\longrightarrow} \Delta_{vui}^+(1, 2\int_0^1 B_{vi}(r)dr, 3\int_0^1 B_{vi}^2(r)dr)', \tag{51}$$

and therefore under Assumptions 1, 2, 3, recalling that $B_{u\cdot v,i}(r) = B_{ui}(r) - \Omega_{uvi}\Omega_{vvi}^{-1}B_{vi}(r)$ and $\Omega_{u\cdot v,i} = \Omega_{uui} - \Omega_{uvi}^2\Omega_{vvi}^{-1}$,

$$\begin{aligned}
&\sum_{t=1}^{T} \tilde{G}_T X_{it}\tilde{u}_{it} - \hat{\Omega}_{uvi}\hat{\Omega}_{vvi}^{-1} G_T \sum_{t=1}^{T} \tilde{X}_{it} v_{it} - G_T\tilde{C}_i^+ \\
&\overset{d,T\to\infty}{\longrightarrow} (\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r) + \Delta_{vui}(1, 2\int_0^1 B_{vi}(r)dr, 3\int_0^1 B_{vi}(r)^2dr)') \\
&-\Omega_{vui}\Omega_{vvi}^{-1}(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{vi}(r) + \Delta_{vvi}(1, 2\int_0^1 B_{vi}(r)dr, 3\int_0^1 B_{vi}^2(r)dr)') \\
&-(\Delta_{vui} - \Delta_{vvi}\Omega_{vvi}^{-1}\Omega_{vui})(1, 2\int_0^1 B_{vi}(r)dr, 3\int_0^1 B_{vi}^2(r)dr)' \\
&= \int_0^1 \tilde{\mathbf{B}}_{vi}(r)d(B_{ui}(r) - \Omega_{uvi}\Omega_{vvi}^{-1}B_{vi}(r)) = \int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{u\cdot v,i}(r),
\end{aligned}$$

implying that, by the CLT for i.i.d. summands, because $\mathbb{E}(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{u\cdot v,i}(r)) = 0$ and

$$\mathbb{E}(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{u\cdot v,i}(r)\int_0^1 \tilde{\mathbf{B}}_{vi}(r)'dB_{u\cdot v,i}(r)) = \mathbb{E}(\Omega_{u\cdot v,i}D_iMD_i)$$

as shown in the proof of Lemma 3, we have

$$N^{-1/2}\sum_{i=1}^{N} G_T(\sum_{t=1}^{T} \tilde{X}_{it}\tilde{u}_{it} - \sum_{t=1}^{T} \tilde{X}_{it}\hat{\Omega}_{uv}\hat{\Omega}_{vv}^{-1} v_{it} - \tilde{C}_i^+) \overset{d,N\to\infty,T\to\infty}{\longrightarrow} N(0, \mathbb{E}(\Omega_{u\cdot v,i}D_iMD_i)).$$

It now follows that

$$N^{1/2}G_T^{-1}(\tilde{\beta}^+ - \beta) \overset{d,N\to\infty,T\to\infty}{\longrightarrow} N(0, V_1^{-1}\mathbb{E}(\Omega_{u\cdot v,i}D_iMD_i)V_1^{-1}),$$

which is the result as stated. □

Proof of Theorem 4:

Because $\check{X}_{it} = \tilde{X}_{it} - V_{NTt}$ and $\check{u}_{it} = \tilde{u}_{it} - w_{NTt}$, $V_{NTt} = N^{-1}\sum_{i=1}^{N}\tilde{X}_{it}$, and $w_{NTt} = N^{-1}\sum_{i=1}^{N}\tilde{u}_{it}$, we have

$$N^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\tilde{X}_{it}V'_{NTt} = \sum_{t=1}^{T}V_{NTt}V'_{NTt},$$

$$N^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\tilde{X}_{it}w_{NTt} = \sum_{t=1}^{T}V_{NTt}w_{NTt},$$

and

$$N^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}V_{NTt}\tilde{u}_{it} = \sum_{t=1}^{T}V_{NTt}w_{NTt},$$

and recalling that $\check{y}_{it} = \check{X}'_{it}\beta + \check{u}_{it}$, it follows that

$$G_T^{-1}(\tilde{\beta} - \beta) = G_T^{-1}(\sum_{i=1}^{N}\sum_{t=1}^{T}\check{X}_{it}\check{X}'_{it})^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\check{X}_{it}\check{u}_{it}$$

$$= (N^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}G_T(\tilde{X}_{it} - V_{NTt})(\tilde{X}_{it} - V_{NTt})'G_T)^{-1}N^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}G_T(\tilde{X}_{it} - V_{NTt})(\tilde{u}_{it} - w_{NTt})$$

$$= (N^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}G_T(\tilde{X}_{it}\tilde{X}'_{it} - V_{NTt}V'_{NTt})G_T)^{-1}N^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}G_T(\tilde{X}_{it}\tilde{u}_{it} - V_{NTt}w_{NTt}).$$

Theorem 1 showed that, if the conditions for parts 7, 8 and 10 of Lemma 1 are assumed,

$$N^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}G_T\tilde{X}_{it}\tilde{X}'_{it}G_T \overset{p,N\to\infty,T\to\infty}{\longrightarrow} V_1$$

and under the moment conditions of Lemma 2,

$$N^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}G_T\tilde{X}_{it}\tilde{u}_{it} \overset{d,N\to\infty,T\to\infty}{\longrightarrow} \begin{pmatrix} -(1/2)\mathbb{E}(\Omega_{uvi}) + \mathbb{E}(\Delta_{vui}) \\ 0 \\ -\mathbb{E}(\Omega_{vvi}\Omega_{uvi}) + (3/2)\mathbb{E}(\Omega_{uvi}\Delta_{vui}) \end{pmatrix},$$

and therefore it now suffices to show that under Assumption 1,

$$N^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}G_TV_{NTt}V'_{NTt}G_T \overset{p,N\to\infty,T\to\infty}{\longrightarrow} \mathrm{diag}(0,(1/12)(\mathbb{E}(\Omega_{vvi}))^2,0)$$

and

$$\sum_{t=1}^{T}G_TV_{NTt}w_{NTt} \overset{p,N\to\infty,T\to\infty}{\longrightarrow} 0.$$

The first result follows from Lemma 4, and it follows from Lemma 5 that under Assumption 1,

$$|N^{1/2}\sum_{t=1}^{T} G_T V_{NTt} w_{NTt} - N^{-1/2}\sum_{j=1}^{N}\int_0^1 \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r))dB_{uj}(r)| \overset{p,N\to\infty,T\to\infty}{\longrightarrow} 0.$$

To show the second result, note that $\mathbb{E}(\int_0^1 \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r))dB_{uj}(r)) = 0$ and that

$$\mathbb{E}(\int_0^1 \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r))dB_{uj}(r))(\int_0^1 \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r))dB_{uj}(r))' = \mathbb{E}(B_{uj}(1)^2)\int_0^1 \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r))\mathbb{E}(\tilde{\mathbf{B}}_{vi}(r))'dr < \infty$$

if $\mathbb{E}|\Omega_{uuj}| < \infty$ and $\mathbb{E}|\Omega_{vvj}^3| < \infty$. Therefore, it follows that

$$|\sum_{t=1}^{T} G_T V_{NTt} w_{NTt}| \overset{p,N\to\infty,T\to\infty}{\longrightarrow} 0,$$

which completes the proof. □

Proof of Theorem 5:

It follows from the definition of $\check{\beta}^m$ that

$$\begin{aligned}
&N^{1/2}G_T^{-1}(\check{\beta}^m - \beta)\\
&= (N^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T} G_T \check{X}_{it}\check{X}_{it}' G_T)^{-1} N^{-1/2}\sum_{i=1}^{N}(\sum_{t=1}^{T} G_T \check{X}_{it}\check{u}_{it} - \tilde{C}_i)\\
&= (N^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T} G_T \check{X}_{it}\check{X}_{it}' G_T)^{-1} N^{-1/2}\sum_{i=1}^{N}(\sum_{t=1}^{T} G_T \tilde{X}_{it}\tilde{u}_{it} - \sum_{t=1}^{T} G_T V_{NTt} w_{NTt} - G_T\tilde{C}_i).
\end{aligned}$$

The proof of Theorem 4 shows that, if the conditions for parts 7, 8 and 10 of Lemma 1 and Assumption 1 are assumed,

$$N^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T} G_T \check{X}_{it}\check{X}_{it}' G_T \overset{d,n\to\infty,T\to\infty}{\longrightarrow} V_2.$$

Furthermore, by Lemma 5, under the conditions of this lemma,

$$|N^{1/2}\sum_{t=1}^{T} G_T V_{NTt} w_{NTt} - N^{-1/2}\sum_{j=1}^{N}\int_0^1 \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r))dB_{uj}(r)| \overset{p,N\to\infty,T\to\infty}{\longrightarrow} 0,$$

and the result of Equation (50) showed that

$$G_T(\sum_{t=1}^{T} \tilde{X}_{it}\tilde{u}_{it} - \tilde{C}_i) \overset{d,T\to\infty}{\longrightarrow} \int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r) - \mathbb{E}(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)|\Omega_i),$$

implying that

$$|N^{-1/2}\sum_{i=1}^{N}(G_T\sum_{t=1}^{T}\tilde{X}_{it}\tilde{u}_{it}-G_T\sum_{t=1}^{T}V_{NTt}w_{NTt}-\tilde{C}_i)$$

$$-N^{-1/2}\sum_{i=1}^{N}(\int_0^1\tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)-\mathbb{E}(\int_0^1\tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)|\Omega_i)-\int_0^1\mathbb{E}(\tilde{\mathbf{B}}_{vi}(r))dB_{ui}(r))|\overset{p,N\to\infty,T\to\infty}{\longrightarrow}0.$$

Therefore, by the CLT for i.i.d. distributed random variables,

$$N^{-1/2}\sum_{i=1}^{N}(G_T\sum_{t=1}^{T}\tilde{X}_{it}\tilde{u}_{it}-G_T\sum_{t=1}^{T}V_{NTt}w_{NTt}-\tilde{C}_i)\overset{d,N\to\infty,T\to\infty}{\longrightarrow}N(0,\Sigma_2)$$

where

$$\Sigma_2=\mathbb{E}(\int_0^1\tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)-\mathbb{E}(\int_0^1\tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)|\Omega_i)-\int_0^1\mathbb{E}(\tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)))$$
$$\times(\int_0^1\tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)-\mathbb{E}(\int_0^1\tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)|\Omega_i)-\int_0^1\mathbb{E}(\tilde{\mathbf{B}}_{vi}(r))dB_{ui}(r))'.$$

In sum,

$$N^{1/2}(G_T^{-1}(\hat{\beta}-\beta)-\breve{B})\overset{d,n\to\infty,T\to\infty}{\longrightarrow}N(0,{V_2}^{-1}\Sigma_2{V_2}^{-1})$$

and the value of Σ_2 as stated in the theorem is calculated in Lemma 6. □

Proof of Theorem 6:

By the definition of $\breve{\beta}^+$,

$$N^{1/2}G_T^{-1}(\breve{\beta}^+-\beta)=(N^{-1}\sum_{i=1}^{N}G_T\sum_{t=1}^{T}\breve{X}_{it}\breve{X}_{it}'G_T)^{-1}$$
$$\times N^{-1/2}\sum_{i=1}^{N}G_T(\sum_{t=1}^{T}\breve{X}_{it}\breve{u}_{it}-\sum_{t=1}^{T}\breve{X}_{it}\hat{\Omega}_{uv}\hat{\Omega}_{vv}^{-1}v_{it}-\tilde{C}_i^+)$$
$$=(N^{-1}\sum_{i=1}^{N}G_T\sum_{t=1}^{T}\breve{X}_{it}\breve{X}_{it}'G_T)^{-1}$$
$$\times N^{-1/2}\sum_{i=1}^{N}G_T(\sum_{t=1}^{T}\tilde{X}_{it}\tilde{u}_{it}-\sum_{t=1}^{T}V_{NTt}w_{NTt}-\sum_{t=1}^{T}\tilde{X}_{it}\hat{\Omega}_{uv}\hat{\Omega}_{vv}^{-1}v_{it}+\sum_{t=1}^{T}V_{NTt}\hat{\Omega}_{uv}\hat{\Omega}_{vv}^{-1}v_{it}-\tilde{C}_i^+)$$
$$=(N^{-1}\sum_{i=1}^{N}G_T\sum_{t=1}^{T}\breve{X}_{it}\breve{X}_{it}'G_T)^{-1}$$
$$\times N^{-1/2}\sum_{i=1}^{N}(G_T\sum_{t=1}^{T}\tilde{X}_{it}\tilde{u}_{it}-G_T\sum_{t=1}^{T}\tilde{X}_{it}\hat{\Omega}_{uv}\hat{\Omega}_{vv}^{-1}v_{it}-G_T\tilde{C}_i^++G_T\sum_{t=1}^{T}V_{NTt}\hat{\Omega}_{uv}\hat{\Omega}_{vv}^{-1}v_{it}-G_T\sum_{t=1}^{T}V_{NTt}w_{NTt}).$$

It was shown in the proof of Theorem 4, if the conditions for parts 7, 8 and 10 of Lemma 1 and Assumption 1 are assumed,

$$N^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T} G_T \check{X}_{it}\check{X}_{it}' G_T \overset{p,N\to\infty,T\to\infty}{\longrightarrow} V_2.$$

Furthermore, it was shown in the proof of Theorem 3,

$$\sum_{t=1}^{T} G_T \tilde{X}_{it}\tilde{u}_{it} - \hat{\Omega}_{uvi}\hat{\Omega}_{vvi}^{-1}\sum_{t=1}^{T} G_T \tilde{X}_{it} v_{it} - G_T \tilde{C}_i^{+} \overset{d,T\to\infty}{\longrightarrow} \int_0^1 \tilde{\mathbf{B}}_{vi}(r) dB_{u\cdot vi}(r),$$

and it follows from Lemma 5 that under Assumption 1,

$$|N^{-1/2}\sum_{t=1}^{T} G_T V_{NTt} w_{NTt} - N^{-1/2}\sum_{j=1}^{N}\int_0^1 \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r)) dB_{uj}(r)| \overset{p,N\to\infty,T\to\infty}{\longrightarrow} 0,$$

while Lemma 7 shows that

$$|N^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T} G_T V_{NTt} v_{it} - N^{-1/2}\sum_{j=1}^{N}\int_0^1 \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r)) dB_{vj}(r)| \overset{p,N\to\infty,T\to\infty}{\longrightarrow} 0,$$

implying that, since $B_{u\cdot v,i}(r) = B_{ui}(r) - \Omega_{uvi}\Omega_{vvi}^{-1}B_{vi}(r)$, we have

$$|\sum_{t=1}^{T} G_T V_{NTt}\hat{\Omega}_{uv}\hat{\Omega}_{vv}^{-1} v_{it} - \sum_{t=1}^{T} G_T V_{NTt} w_{NTt} + N^{-1/2}\sum_{j=1}^{N}\int_0^1 \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r)) dB_{u\cdot v,i}(r)| \overset{p,N\to\infty,T\to\infty}{\longrightarrow} 0.$$

Together this implies that

$$|N^{-1/2}\sum_{i=1}^{N}(G_T\sum_{t=1}^{T}\tilde{X}_{it}\tilde{u}_{it} - G_T\sum_{t=1}^{T}\tilde{X}_{it}\hat{\Omega}_{uv}\hat{\Omega}_{vv}^{-1}v_{it} - G_T\tilde{D}_i + G_T\sum_{t=1}^{T}V_{NTt}\hat{\Omega}_{uv}\hat{\Omega}_{vv}^{-1}v_{it} - G_T\sum_{t=1}^{T}V_{NTt}w_{NTt})$$

$$-N^{-1/2}\sum_{i=1}^{N}(\int_0^1 (\tilde{\mathbf{B}}_{vi}(r) - \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r))) dB_{u\cdot v,i}(r))| \overset{p,N\to\infty,T\to\infty}{\longrightarrow} 0.$$

Therefore, by the CLT for i.i.d. summands, because $\mathbb{E}(\int_0^1 (\tilde{\mathbf{B}}_{vi}(r) - \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r))) dB_{u\cdot vi}(r)) = 0$ and

$$\mathbb{E}(\int_0^1 (\tilde{\mathbf{B}}_{vi}(r) - \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r))) dB_{u\cdot vi}(r)\int_0^1 (\tilde{\mathbf{B}}_{vi}(r) - \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r)))' dB_{u\cdot vi}(r)|\Omega_i)$$

$$= \mathbb{E}(B_{u\cdot vi}(1)^2|\Omega_i)\mathbb{E}(\int_0^1 (\tilde{\mathbf{B}}_{vi}(r) - \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r)))(\tilde{\mathbf{B}}_{vi}(r) - \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r)))' dr|\Omega_i).$$

Now,

$$\mathbb{E}(\int_0^1 (\tilde{\mathbf{B}}_{vi}(r) - \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r)))(\tilde{\mathbf{B}}_{vi}(r) - \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r)))' dr|\Omega_i)$$

$$= \mathbb{E}(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)\tilde{\mathbf{B}}_{vi}(r)' dr|\Omega_i) - \int_0^1 \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r))\mathbb{E}(\tilde{\mathbf{B}}_{vi}(r)|\Omega_i)' dr$$

$$-\int_0^1 \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r))\mathbb{E}(\tilde{\mathbf{B}}_{vi}(r)|\Omega_i)'dr + \mathbb{E}(D_i)\text{diag}(0, 1/12, 0)\mathbb{E}(D_i)$$

$$= D_i M D_i - \mathbb{E}(D_i)\int_0^1 \mathbb{E}(\tilde{\mathbf{W}}_{vi}(r))\mathbb{E}(\tilde{\mathbf{W}}_{vi}(r))'dr D_i$$

$$-D_i \int_0^1 \mathbb{E}(\tilde{\mathbf{W}}_{vi}(r))\mathbb{E}(\tilde{\mathbf{W}}_{vi}(r))'dr\mathbb{E}(D_i) + \mathbb{E}(D_i)\text{diag}(0, 1/12, 0)\mathbb{E}(D_i)$$

$$= D_i M D_i - \mathbb{E}(D_i)\text{diag}(0, 1/12, 0)D_i$$

$$-D_i\text{diag}(0, 1/12, 0)\mathbb{E}(D_i) + \mathbb{E}(D_i)\text{diag}(0, 1/12, 0)\mathbb{E}(D_i)$$

so after taking expectations we get

$$\mathbb{E}(\Omega_{u\cdot v,i}D_i M D_i) - \text{diag}(0, (1/6)\mathbb{E}(\Omega_{vvi})\mathbb{E}(\Omega_{u\cdot v,i}\Omega_{vvi}), 0) + \text{diag}(0, (1/12)\mathbb{E}(\Omega_{u\cdot v,i})(\mathbb{E}(\Omega_{vvi}))^2, 0),$$

and therefore the result of the theorem now follows. □

Appendix B: Proofs of Lemmata

Below, $|\cdot|$ is defined as the Frobenius norm, viz. $|A| = (\text{tr}(A'A))^{1/2}$.

Lemma 1. *Let Assumption 1 hold, then*

1. *For* $p \geq 1$, $\mathbb{E}\int_0^1 |B_{vi}(r)|^p dr < \infty$ *if* $\mathbb{E}|\Omega_{vvi}|^{p/2} < \infty$.

2. $\mathbb{E}|\int_0^1 \mathbf{B}_{vi}(r)dB_{u\cdot v,i}(r)| < \infty$ *if for* $p = 1, 2, 3$, $\mathbb{E}|\Omega_{uui}\Omega_{vvi}^p|^{1/2} < \infty$ *and* $\mathbb{E}|\Omega_{uvi}^2\Omega_{vvi}^{p-1}|^{1/2} < \infty$.

3. $\mathbb{E}|\int_0^1 \mathbf{B}_{vi}(r)dB_{ui}(r)| < \infty$ *if for* $p = 1, 2, 3$, $\mathbb{E}|\Omega_{uui}\Omega_{vvi}^p|^{1/2} < \infty$ *and* $\mathbb{E}|\Omega_{uvi}^2\Omega_{vvi}^{p-1}|^{1/2} < \infty$;

4. $\mathbb{E}|\int_0^1 \mathbf{B}_{vi}(r)dB_{ui}(r)|^2 < \infty$ *if for* $p = 1, 2, 3$, $\mathbb{E}|\Omega_{uui}\Omega_{vvi}^p| < \infty$, $\mathbb{E}|\Omega_{uvi}\Omega_{vvi}^{p-1}| < \infty$ *and* $\mathbb{E}|\Omega_{uvi}^2\Omega_{vvi}^{p-1}| < \infty$;

5. $\mathbb{E}|B_{ui}(1)\int_0^1 \mathbf{B}_{vi}(r)dr| < \infty$ *if* $\mathbb{E}|\Omega_{uui}|^{1/2} < \infty$, $\mathbb{E}|\Omega_{vvi}^{-1/2}\Omega_{uvi}| < \infty$, $\mathbb{E}|\Omega_{vvi}|^{3/2} < \infty$, *and for* $p = 1, 2, 3$, $\mathbb{E}|\Omega_{uvi}\Omega_{vvi}^{(p-1)/2}| < \infty$;

6. $\mathbb{E}|B_{ui}(1)\int_0^1 \mathbf{B}_{vi}(r)dr|^2 < \infty$ *if* $\mathbb{E}|\Omega_{uui}| < \infty$, $\mathbb{E}|\Omega_{vvi}^{-1}\Omega_{uvi}^2| < \infty$, $\mathbb{E}|\Omega_{vvi}|^3 < \infty$, *and for* $p = 1, 2, 3$, $\mathbb{E}|\Omega_{uvi}^2\Omega_{vvi}^{p-1}| < \infty$;

7. $\mathbb{E}|\int_0^1 \tilde{\mathbf{B}}_{vi}(r)\tilde{\mathbf{B}}_{vi}(r)'dr| < \infty$ *if* $\mathbb{E}|\Omega_{vvi}|^3 < \infty$;

8. $\mathbb{E}|\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)| < \infty$ *if the conditions from 3 and 5 hold;*

9. $\mathbb{E}|\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)|^2 < \infty$ *if the conditions from 4 and 6 hold;*

10. $\mathbb{E}|\Delta_{vui}(1, 2\int_0^1 B_{vi}(r)dr, 3\int_0^1 B_{vi}(r)^2dr)'| < \infty$ *if* $\mathbb{E}|\Delta_{vui}| < \infty$, $\mathbb{E}|\Delta_{vui}\Omega_{vvi}^{1/2}| < \infty$ *and* $\mathbb{E}|\Delta_{vui}\Omega_{vvi}| < \infty$.

Proof of Lemma 1:

Note that by assumption, Ω_{vvi} is independent of $W_{vi}(r)$. Therefore the first result follows because for all $p \geq 1$

$$\mathbb{E}|\int_0^1 B_{vi}(r)^p dr| = \mathbb{E}|\Omega_{vvi}|^{p/2}\mathbb{E}\int_0^1 |W_{vi}(r)|^p dr,$$

and because $\mathbb{E}\int_0^1 |W_{vi}(r)|^p dr < \infty$, for $p \geq 1$, $\mathbb{E}\int_0^1 |B_{vi}(r)|^p dr < \infty$ if $\mathbb{E}|\Omega_{vvi}|^{p/2} < \infty$, as asserted.

To show the second result, note that since $B_{u\cdot v,i}(r) = B_{ui}(r) - \rho_i B_{vi}(r)$ for $\rho_i = \Omega_{uvi}\Omega_{vvi}^{-1}$, since $B_{vi}(r)$ and $B_{u\cdot v,i}(r)$ are independent by construction,

$$\mathbb{E}|\int_0^1 B_{vi}(r)^p dB_{u\cdot v,i}(r)|$$

$$\leq \mathbb{E}(\mathbb{E}(|\int_0^1 B_{vi}(r)^p dB_{u\cdot v,i}(r)|^2|\Omega_i))^{1/2}$$

$$= \mathbb{E}\left(\mathbb{E}(B_{u\cdot v,i}(1)^2|\Omega_i)\mathbb{E}(\int_0^1 B_{vi}(r)^{2p}dr|\Omega_i)\right)^{1/2}$$

$$= \mathbb{E}(\Omega_{u\cdot v,i}\Omega_{vvi}^p)^{1/2}\mathbb{E}(\int_0^1 W(r)^{2p}dr)^{1/2},$$

implying that $\mathbb{E}|\int_0^1 B_{vi}(r)^p dB_{u\cdot v,i}(r)| < \infty$ if $\mathbb{E}|\Omega_{u\cdot v,i}\Omega_{vvi}^p|^{1/2} < \infty$, and therefore $\mathbb{E}|\int_0^1 \mathbf{B}_{vi}(r)dB_{u\cdot v,i}(r)| < \infty$ if for $p = 1, 2, 3$, $\mathbb{E}|\Omega_{u\cdot v,i}\Omega_{vvi}^p|^{1/2} < \infty$. Since $\Omega_{u\cdot v,i} = \Omega_{uui} - \Omega_{vvi}^{-1}\Omega_{uvi}^2$, this is implied by $\mathbb{E}|\Omega_{uui}\Omega_{vvi}^p|^{1/2} < \infty$ and $\mathbb{E}|\Omega_{uvi}^2\Omega_{vvi}^{p-1}|^{1/2} < \infty$.

To show the third result, note that since $B_{u\cdot v,i}(r) = B_{ui}(r) - \rho_i B_{vi}(r)$ for $\rho_i = \Omega_{uvi}\Omega_{vvi}^{-1}$,

$$\mathbb{E}|\int_0^1 B_{vi}(r)^p dB_{ui}(r)| \leq \mathbb{E}|\int_0^1 B_{vi}(r)^p dB_{u\cdot v,i}(r)| + \mathbb{E}|\rho_i\int_0^1 B_{vi}(r)^p dB_{vi}(r)|,$$

and by the second part of this lemma, the first expression is finite under the stated assumptions. To show that the second expression is finite also, note that

$$\mathbb{E}|\rho_i\int_0^1 B_{vi}(r)^p dB_{vi}(r)| = \mathbb{E}|\rho_i\Omega_{vvi}^{(p+1)/2}|\mathbb{E}|\int_0^1 W_{vi}(r)^p dW_{vi}(r)|,$$

implying that $\mathbb{E}|\rho_i\int_0^1 B_{vi}(r)^p dB_{vi}(r)| < \infty$ if $\mathbb{E}|\Omega_{uvi}\Omega_{vvi}^{(p-1)/2}| = \mathbb{E}|\Omega_{uvi}^2\Omega_{vvi}^{p-1}|^{1/2} < \infty$. This condition was also assumed for $p = 1, 2, 3$, thereby ensuring that $\mathbb{E}|\int_0^1 \mathbf{B}_{vi}(r)dB_{ui}(r)| < \infty$.

To show the fourth result, we can reason similarly and conclude that $\mathbb{E}|\int_0^1 \mathbf{B}_{vi}(r)dB_{ui}(r)|^2 < \infty$ if for $p = 1, 2, 3$, $\mathbb{E}|\int_0^1 B_{vi}(r)^p dB_{u\cdot v,i}(r)|^2 < \infty$ and $\mathbb{E}|\rho_i\int_0^1 B_{vi}(r)^p dB_{vi}(r)|^2 < \infty$. Now

$$\mathbb{E}|\int_0^1 B_{vi}(r)^p dB_{u\cdot v,i}(r)|^2$$

$$\leq \mathbb{E}\mathbb{E}(|\int_0^1 B_{vi}(r)^p dB_{u\cdot v,i}(r)|^2|\Omega_i)$$

$$= \mathbb{E}\mathbb{E}(B_{u\cdot v,i}(1)^2|\Omega_i)\mathbb{E}(\int_0^1 B_{vi}(r)^{2p}dr|\Omega_i)$$

$$= \mathbb{E}(\Omega_{u\cdot v,i}\Omega_{vvi}^p)\mathbb{E}(\int_0^1 W(r)^{2p}dr),$$

implying that $\mathbb{E}|\int_0^1 B_{vi}(r)^p dB_{u\cdot v,i}(r)|^2 < \infty$ if $\mathbb{E}|\Omega_{u\cdot v,i}\Omega_{vvi}^p| < \infty$, and since $\Omega_{u\cdot v,i} = \Omega_{uui} - \Omega_{vvi}^{-1}\Omega_{vui}^2$, this is implied by $\mathbb{E}|\Omega_{uui}\Omega_{vvi}^p| < \infty$ and $\mathbb{E}|\Omega_{uvi}\Omega_{vvi}^{p-1}| < \infty$. These conditions were assumed for $p = 1, 2, 3$. Finally,

$$\mathbb{E}|\rho_i \int_0^1 B_{vi}(r)^p dB_{vi}(r)|^2 = \mathbb{E}|\rho_i \Omega_{vvi}^{(p+1)/2}|^2 \mathbb{E}|\int_0^1 W_{vi}(r)^p dW_{vi}(r)|^2,$$

implying that $\mathbb{E}|\rho_i \int_0^1 B_{vi}(r)^p dB_{vi}(r)|^2 < \infty$ if $\mathbb{E}|\Omega_{uvi}^2\Omega_{vvi}^{p-1}| < \infty$. This condition was also assumed for $p = 1, 2, 3$.

To show the fifth result, note that since $B_{u\cdot v,i}(r) = B_{ui}(r) - \rho_i B_{vi}(r)$,

$$\mathbb{E}|B_{ui}(1)\int_0^1 \mathbf{B}_{vi}(r)dr| \leq \mathbb{E}|B_{u\cdot v,i}(1)\int_0^1 \mathbf{B}_{vi}(r)dr| + \mathbb{E}|\rho_i B_{vi}(1)\int_0^1 \mathbf{B}_{vi}(r)dr|,$$

and since $B_{vi}(r)$ and $B_{u\cdot v,i}(r)$ are independent by construction,

$$\begin{aligned}\mathbb{E}|B_{u\cdot v,i}(1)\int_0^1 B_{vi}(r)^p dr| &= \mathbb{E}|B_{u\cdot v,i}(1)|\mathbb{E}|\int_0^1 B_{vi}(r)^p dr| \\ &= \mathbb{E}|\Omega_{u\cdot v,i}|^{1/2}\mathbb{E}|\Omega_{vvi}|^{p/2}\mathbb{E}|\int_0^1 W_{vi}(r)^p dr|,\end{aligned}$$

implying that $\mathbb{E}|B_{u\cdot v,i}(1)\int_0^1 \mathbf{B}_{vi}(r)dr| < \infty$ if $\mathbb{E}|\Omega_{u\cdot v,i}|^{1/2} < \infty$ and $\mathbb{E}|\Omega_{vvi}|^{3/2} < \infty$. Since $\Omega_{u\cdot v,i} = \Omega_{uui} - \Omega_{vvi}^{-1}\Omega_{uvi}^2$, the condition $\mathbb{E}|\Omega_{u\cdot v,i}|^{1/2} < \infty$ is implied by $\mathbb{E}|\Omega_{uui}|^{1/2} < \infty$ and $\mathbb{E}|\Omega_{vvi}^{-1/2}\Omega_{uvi}| < \infty$, which was assumed. Also,

$$\begin{aligned}\mathbb{E}|\rho_i B_{vi}(1)\int_0^1 B_{vi}(r)^p dr| &= \mathbb{E}|\rho_i \Omega_{vvi}^{(p+1)/2}\int_0^1 W_{vi}(r)^p dr| \\ &\leq \mathbb{E}|\Omega_{uvi}\Omega_{vvi}^{(p-1)/2}|\ \mathbb{E}(\int_0^1 |W_{vi}(r)|^p dr),\end{aligned}$$

and therefore $\mathbb{E}|\rho_i B_{vi}(1)\int_0^1 \mathbf{B}_{vi}(r)dr| < \infty$ if for $p = 1, 2, 3$, $\mathbb{E}|\Omega_{uvi}\Omega_{vvi}^{(p-1)/2}| < \infty$, which was also assumed.

To show the sixth result, note that the reasoning of the fourth result ensures that $\mathbb{E}|B_{ui}(1)\int_0^1 \mathbf{B}_{vi}(r)dr|^2 < \infty$ if $\mathbb{E}|B_{u\cdot v,i}(1)\int_0^1 \mathbf{B}_{vi}(r)dr|^2 < \infty$ and $\mathbb{E}|\rho_i B_{vi}(1)\int_0^1 \mathbf{B}_{vi}(r)dr|^2 < \infty$. Similarly to the earlier reasoning,

$$\mathbb{E}|B_{u\cdot v,i}(1)\int_0^1 B_{vi}(r)^p dr|^2 = \mathbb{E}|B_{u\cdot v,i}(1)|^2\mathbb{E}|\int_0^1 B_{vi}(r)^p dr|^2 = \mathbb{E}|\Omega_{u\cdot v,i}|\mathbb{E}|\Omega_{vvi}|^p\mathbb{E}\int_0^1 |W_{vi}(r)|^p dr,$$

implying that $\mathbb{E}|B_{u\cdot v,i}(1)\int_0^1 \mathbf{B}_{vi}(r)dr|^2 < \infty$ if $\mathbb{E}|\Omega_{u\cdot v,i}| < \infty$ and $\mathbb{E}|\Omega_{vvi}|^3 < \infty$. Since $\Omega_{u\cdot v,i} = \Omega_{uui} - \Omega_{vvi}^{-1}\Omega_{uvi}^2$, the condition $\mathbb{E}|\Omega_{u\cdot v,i}| < \infty$ is implied by $\mathbb{E}|\Omega_{uui}| < \infty$ and $\mathbb{E}|\Omega_{vvi}^{-1}\Omega_{uvi}^2| < \infty$, which was assumed. Also,

$$\mathbb{E}|\rho_i B_{vi}(1)\int_0^1 B_{vi}(r)^p dr|^2 = \mathbb{E}|\rho_i \Omega_{vvi}^{(p+1)/2}\int_0^1 W_{vi}(r)^p dr|^2 \leq \mathbb{E}|\Omega_{uvi}\Omega_{vvi}^{(p+1)/2-1}|^2\mathbb{E}(\int_0^1 |W_{vi}(r)|^p dr)^2,$$

and therefore $\mathbb{E}|\rho_i B_{vi}(1) \int_0^1 \mathbf{B}_{vi}(r)dr| < \infty$ if $\mathbb{E}|\Omega_{uvi}\Omega_{vvi}^{(p-1)/2}|^2 < \infty$ for $p = 1, 2, 3$.

To show the seventh result, using properties of the Frobenius norm, note that

$$
\begin{aligned}
\mathbb{E}|\int_0^1 \tilde{\mathbf{B}}_{vi}(r)\tilde{\mathbf{B}}_{vi}(r)'dr| &\leq \mathbb{E}\int_0^1 |\tilde{\mathbf{B}}_{vi}(r)|^2 dr \\
&= \mathbb{E}\int_0^1 |\mathbf{B}_{vi}(r) - \int_0^1 \mathbf{B}_{vi}(s)ds|^2 dr \\
&\leq 4\int_0^1 \mathbb{E}|\mathbf{B}_{vi}(r)|^2 dr,
\end{aligned}
$$

and by part 1 of this lemma using $p = 6$, the last expression is finite if $\mathbb{E}|\Omega_{vvi}|^3 < \infty$.

For the eight result, note that by definition

$$
\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r) = \int_0^1 \mathbf{B}_{vi}(r)dB_{ui}(r) - B_{ui}(1)\int_0^1 \mathbf{B}_{vi}(r)dr,
$$

so the combined conditions of the third and fifth part of this lemma suffice, and those were the assumed conditions here.

Similarly, for the ninth part of this lemma, it suffices to show that $\mathbb{E}|\int_0^1 \mathbf{B}_{vi}(r)dB_{ui}(r)|^2 < \infty$ and that $\mathbb{E}|B_{ui}(1)\int_0^1 \mathbf{B}_{vi}(r)dr|^2 < \infty$. So the combined conditions of the fourth and sixth part of this lemma suffice, and those were the assumed conditions here.

Finally, the tenth result follows from noting that

$$
\mathbb{E}|\Delta_{vui}(1, 2\int_0^1 B_{vi}(r)dr, 3\int_0^1 B_{vi}(r)^2 dr)'| = \mathbb{E}|\Delta_{vui}(1, 2\Omega_{vvi}^{1/2}\int_0^1 W_{vi}(r)dr, 3\Omega_{vvi}\int_0^1 W_{vi}(r)^2 dr)'|
$$

and therefore $\mathbb{E}|\Delta_{vui}| < \infty$, $\mathbb{E}|\Delta_{vui}\Omega_{vvi}^{1/2}| < \infty$ and $\mathbb{E}|\Delta_{vui}\Omega_{vvi}| < \infty$ suffices. □

Lemma 2. *Let Assumption 1 hold and assume that* $\mathbb{E}|\Omega_{uui}|^{1/2} < \infty$, $\mathbb{E}|\Omega_{vvi}^{-1/2}\Omega_{uvi}| < \infty$, $\mathbb{E}|\Omega_{vvi}|^{3/2} < \infty$, *and for* $p = 1, 2, 3$, $\mathbb{E}|\Omega_{uui}\Omega_{vvi}^{p}|^{1/2} < \infty$ *and* $\mathbb{E}|\Omega_{uvi}\Omega_{vvi}^{p/2-1/2}| < \infty$. *Then*

$$\mathbb{E}(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)|\Omega_i) = \begin{pmatrix} -(1/2)\Omega_{uvi} \\ 0 \\ -\Omega_{vvi}\Omega_{uvi} \end{pmatrix}.$$

Proof of Lemma 2:

It follows from part 8 of Lemma 1 that $\mathbb{E}|\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)| < \infty$ under the conditions of this lemma. Furthermore, using $\mathbb{E}(\int_0^1 B_{vi}(r)^p dB_{ui}(r)|\Omega_i) = 0$ for $p = 1, 2, 3$,

$$\begin{aligned}\mathbb{E}(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)|\Omega_i) &= \mathbb{E}(\int_0^1 \mathbf{B}_{vi}(r)dB_{ui}(r)|\Omega_i) - \mathbb{E}(B_{ui}(1)\int_0^1 \mathbf{B}_{vi}(s)ds|\Omega_i) \\ &= -\mathbb{E}(\int_0^1 \mathbf{B}_{vi}(s)B_{ui}(1)ds|\Omega_i) \\ &= -\mathbb{E}(\int_0^1 \mathbf{B}_{vi}(s)B_{ui}(s)ds|\Omega_i) - \mathbb{E}(\int_0^1 \mathbf{B}_{vi}(s)(B_{ui}(1) - B_{ui}(s))ds|\Omega_i),\end{aligned}$$

and both objects are well-defined as follows from

$$\mathbb{E}|\int_0^1 B_{vi}(s)^p B_{ui}(s)ds| = \mathbb{E}|\Omega_{uui}^{1/2}\Omega_{vvi}^{p/2}|\mathbb{E}|\int_0^1 W_{vi}(s)^p W_{ui}(s)ds| < \infty,$$

if for $p = 1, 2, 3$, $\mathbb{E}|\Omega_{uui}^{1/2}\Omega_{vvi}^{p/2}| < \infty$, which was assumed. Also, using the independent increments property of Brownian motion, for $p = 1, 2, 3$,

$$\mathbb{E}(\int_0^1 B_{vi}(s)^p(B_{ui}(1) - B_{ui}(s))ds|\Omega_i) = \Omega_{uui}^{1/2}\Omega_{vvi}^{p/2}\mathbb{E}(\int_0^1 W_{vi}(s)^p(W_{ui}(1) - W_{ui}(s))ds) = 0.$$

Therefore,

$$\mathbb{E}(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)|\Omega_i) = -\mathbb{E}(\int_0^1 \mathbf{B}_{vi}(s)B_{ui}(s)ds|\Omega_i).$$

Since $(s^{1/2}B_{ui}(1), s^{1/2}B_{vi}(1)) \overset{d}{=} (B_{ui}(s), B_{vi}(s))$, using the substitution $B_{ui}(r) = B_{u\cdot v,i}(r)+\rho_i B_{vi}(r)$, the conditional independence of the processes $B_{u\cdot v,i}(r)$ and $B_{vi}(r)$, and remembering that $\rho_i = \Omega_{uvi}\Omega_{vvi}^{-1}$, it follows that the previous expression equals

$$-\mathbb{E}(\begin{pmatrix} B_{vi}(1)B_{ui}(1)\int_0^1 sds \\ B_{vi}(1)^2 B_{ui}(1)\int_0^1 s^{3/2}ds \\ B_{vi}(1)^3 B_{ui}(1)\int_0^1 s^2 ds \end{pmatrix}|\Omega_i) = -\mathbb{E}(\begin{pmatrix} (1/2)B_{vi}(1)(B_{u\cdot v,i}(1) + \rho_i B_{vi}(1)) \\ (2/5)B_{vi}(1)^2(B_{u\cdot v,i}(1) + \rho_i B_{vi}(1)) \\ (1/3)B_{vi}(1)^3(B_{u\cdot v,i}(1) + \rho_i B_{vi}(1)) \end{pmatrix}|\Omega_i)$$

$$= \begin{pmatrix} -(1/2)\rho_i \mathbb{E}(B_{vi}(1))^2|\Omega_i) \\ 0 \\ -(1/3)\rho_i \mathbb{E}(B_{vi}(1))^4|\Omega_i) \end{pmatrix} = \begin{pmatrix} -(1/2)\rho_i \Omega_{vvi}) \\ 0 \\ -\rho_i \Omega_{vvi}^2 \end{pmatrix} = \begin{pmatrix} -(1/2)\Omega_{uvi} \\ 0 \\ -\Omega_{vvi}\Omega_{uvi} \end{pmatrix}.$$

□

Lemma 3. *Let Assumption 1 hold, then*

$$\Sigma_1 = \mathbb{E}\left(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r) - \mathbb{E}(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)|\Omega_i)\right)\left(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r) - \mathbb{E}(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)|\Omega_i)\right)'$$

$$= \mathbb{E}(\Omega_{u\cdot v,i} D_i M D_i) + \mathbb{E}(\Omega_{uvi}^2 \Omega_{vvi}^{-1} D_i Q D_i)$$

$$-\mathbb{E}\begin{pmatrix} (1/4)\Omega_{uvi}^2 & 0 & (1/2)\Omega_{vvi}\Omega_{uvi}^2 \\ 0 & 0 & 0 \\ (1/2)\Omega_{vvi}\Omega_{uvi}^2 & 0 & \Omega_{vvi}^2\Omega_{uvi}^2 \end{pmatrix},$$

with M as given in (19) *and Q as given in* (23).

Proof of Lemma 3:

Using

$$\mathbb{E}\left((Z_i - \mathbb{E}(Z_i|\Omega_i))(Z_i - \mathbb{E}(Z_i|\Omega_i))'\right)$$

$$= \mathbb{E}(Z_i Z_i') - \mathbb{E}\left(\mathbb{E}(Z_i|\Omega_i)\mathbb{E}(Z_i|\Omega_i)'\right) \tag{52}$$

for $Z_i = \int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)$ leads to

$$\Sigma_1 = \mathbb{E}\left((\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r))(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r))'\right)$$

$$-\mathbb{E}\left(\mathbb{E}(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)|\Omega_i)\mathbb{E}(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)|\Omega_i)'\right) \tag{53}$$

with

$$\mathbb{E}(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)|\Omega_i) = \begin{pmatrix} -(1/2)\Omega_{uvi} \\ 0 \\ -\Omega_{vvi}\Omega_{uvi} \end{pmatrix} \tag{54}$$

as calculated in Lemma 2. It thus remains to consider the first term above. Using $B_{ui}(r) = B_{u\cdot v,i}(r) + \rho_i B_{vi}(r)$, we find

$$\mathbb{E}((\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r))(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r))'|\Omega_i)$$

$$
\begin{aligned}
&= \mathbb{E}\left(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{u\cdot v,i}(r)\int_0^1 \tilde{\mathbf{B}}_{vi}(r)'dB_{u\cdot v,i}(r)|\Omega_i\right)\\
&+\rho_i\mathbb{E}\left(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{u\cdot v,i}(r)\int_0^1 \tilde{\mathbf{B}}_{vi}(r)'dB_{vi}(r)|\Omega_i\right)\\
&+\rho_i\mathbb{E}\left(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{vi}(r)\int_0^1 \tilde{\mathbf{B}}_{vi}(r)'dB_{u\cdot v,i}(r)|\Omega_i\right)\\
&+\rho_i^2\mathbb{E}\left(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{vi}(r)\int_0^1 \tilde{\mathbf{B}}_{vi}(r)'dB_{vi}(r)|\Omega_i\right).) \qquad (55)
\end{aligned}
$$

Note that the second and third term are well-defined by the Cauchy-Schwartz inequality if the first and fourth term are finite. Since the first and fourth term are positive semi-definite, they are well-defined if the diagonal elements are well-defined, which is shown below. For the first term, we have

$$
\mathbb{E}\left((\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{u\cdot v,i}(r))(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{u\cdot v,i}(r))'|\Omega_i\right) = \mathbb{E}(B_{u\cdot v,i}(1)^2|\Omega_i)D_iMD_i = \Omega_{u\cdot v,i}D_iMD_i
$$

and for the fourth term, we have

$$
\begin{aligned}
&\rho_i^2\mathbb{E}\left((\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{vi}(r))(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{vi}(r))'|\Omega_i\right)\\
&= \rho_i^2\Omega_{vvi}D_i\mathbb{E}\left((\int_0^1 \tilde{\mathbf{W}}_{vi}(r)dW_{vi}(r))(\int_0^1 \tilde{\mathbf{W}}_{vi}(r)dW_{vi}(r))'|\Omega_i\right)D_i.\\
&= \Omega_{uvi}^2\Omega_{vvi}^{-1}D_iQD_i
\end{aligned}
$$

with Q as in Equation (23). The second and third terms vanish because of the conditional independence of the processes $B_{u\cdot v,i}(r)$ and $B_{vi}(r)$.

Combining these findings for the expression considered in Equation (55) with the results of Equations (53) and (54) leads to

$$
\Sigma_1 = \mathbb{E}(\Omega_{u\cdot v,i}D_iMD_i) + \mathbb{E}(\Omega_{uvi}^2\Omega_{vvi}^{-1}D_iQD_i) + \mathbb{E}(\begin{pmatrix} -(1/2)\Omega_{uvi} \\ 0 \\ -\Omega_{vvi}\Omega_{uvi} \end{pmatrix}\begin{pmatrix} -(1/2)\Omega_{uvi} \\ 0 \\ -\Omega_{vvi}\Omega_{uvi} \end{pmatrix}').
$$

□

Lemma 4. *Let Assumption 1 hold and assume additionally that* $\mathbb{E}|\Omega_{vvi}|^3 < \infty$. *Then*

$$
G_TN^{-1}\sum_{t=1}^T V_{NTt}V_{NTt}'G_T \overset{p,N\to\infty,T\to\infty}{\longrightarrow} \text{diag}(0,(1/12)(\mathbb{E}(\Omega_{vvi}))^2,0).
$$

Proof. Note that because $V_{NTt} = N^{-1}\sum_{i=1}^{N}\tilde{X}_{it}$, we have, by the continuous mapping theorem and Assumption 1,

$$\sum_{t=1}^{T} G_T V_{NTt} V_{NTt}' G_T = \sum_{t=1}^{T} N^{-2} \sum_{i=1}^{N} G_T \tilde{X}_{Tti} \sum_{j=1}^{N} \tilde{X}_{Ttj}' G_T$$

$$= N^{-2} \sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T} G_T \tilde{X}_{Tti}\tilde{X}_{Ttj}' G_T \overset{d,T\to\infty}{\longrightarrow} N^{-2}\sum_{i=1}^{N}\sum_{j=1}^{N}\int_0^1 \tilde{\mathbf{B}}_{vi}(r)\tilde{\mathbf{B}}_{vj}(r)'dr$$

$$= \int_0^1 (N^{-1}\sum_{i=1}^{N}\tilde{\mathbf{B}}_{vi}(r))(N^{-1}\sum_{j=1}^{N}\tilde{\mathbf{B}}_{vj}(r))'dr.$$

It now follows that under the assumptions of the lemma $\int_0^1 (N^{-1}\sum_{i=1}^{N}\tilde{\mathbf{B}}_{vi}(r))(N^{-1}\sum_{j=1}^{N}\tilde{\mathbf{B}}_{vj}(r))'$ converges to $\int_0^1 \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r))\mathbb{E}(\tilde{\mathbf{B}}_{vj}(r))'dr$, as can be verified by showing that the expectation and the variance of the difference converges to zero. Thus,

$$|G_T \sum_{t=1}^{T} V_{NTt}V_{NTt}' G_T - \int_0^1 \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r))\mathbb{E}(\tilde{\mathbf{B}}_{vj}(r))'dr| \overset{p,N\to\infty,T\to\infty}{\longrightarrow} 0.$$

Next, note that

$$\int_0^1 \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r))\mathbb{E}(\tilde{\mathbf{B}}_{vj}(r))'dr = \int_0^1 \mathbb{E}(D_i)\mathbb{E}(\tilde{\mathbf{W}}_{vi}(r))\mathbb{E}(\tilde{\mathbf{W}}_{vi}(r))'\mathbb{E}(D_i)dr$$

$$= \mathbb{E}(D_i)\int_0^1 \begin{pmatrix} 0 & 0 & 0 \\ 0 & (r-1/2)^2 & 0 \\ 0 & 0 & 0 \end{pmatrix} dr\mathbb{E}(D_i)$$

$$= \mathbb{E}(D_i)\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1/12 & 0 \\ 0 & 0 & 0 \end{pmatrix}\mathbb{E}(D_i) = \text{diag}(0, (1/12)(\mathbb{E}(\Omega_{vvi}))^2, 0)$$

because

$$\mathbb{E}(\tilde{\mathbf{W}}_{vi}(r)) = \mathbb{E}\begin{pmatrix} W_{vi}(r) - \int_0^1 W_{vi}(s)ds \\ W_{vi}^2(r) - \int_0^1 W_{vi}^2(s)ds \\ W_{vi}^3(r) - \int_0^1 W_{vi}^3(s)ds \end{pmatrix}$$

$$= \begin{pmatrix} 0 \\ r - \int_0^1 sds \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ r - 1/2 \\ 0 \end{pmatrix}$$

and

$$\mathbb{E}(D_i) = \text{diag}(\mathbb{E}(\Omega_{vvi}^{1/2}), \mathbb{E}(\Omega_{vvi}), \mathbb{E}(\Omega_{vvi}^{3/2})).$$

Therefore,

$$|\sum_{t=1}^{T} G_T V_{NTt} V_{NTt}' G_T - \text{diag}(0, (1/12)(\mathbb{E}(\Omega_{vvi}))^2, 0)| \overset{p,N\to\infty,T\to\infty}{\longrightarrow} 0,$$

which is the asserted result. □

Lemma 5. *Let Assumption 1 hold. In addition, assume* $\mathbb{E}|\Omega_{vvi}|^3 < \infty$, $\mathbb{E}|\Omega_{uui}| < \infty$, $\mathbb{E}|\Omega_{uui}\Omega_{vvi}^3| < \infty$, $\mathbb{E}|\Delta_{vui}| < \infty$, $\mathbb{E}|\Delta_{vui}\Omega_{vvi}^{1/2}| < \infty$ *and* $\mathbb{E}|\Delta_{vui}\Omega_{vvi}| < \infty$*. Then*

$$|N^{1/2}\sum_{t=1}^{T} G_T V_{NTt} w_{NTt} - N^{-1/2}\sum_{j=1}^{N}\int_0^1 \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r))dB_{uj}(r)| \overset{p,N\to\infty,T\to\infty}{\longrightarrow} 0.$$

Proof. By Assumption 1 and 3, and because $V_{NTt} = N^{-1}\sum_{i=1}^{N}\tilde{X}_{it}$ and $w_{NTt} = N^{-1}\sum_{i=1}^{N}\tilde{u}_{it}$,

$$N^{1/2}\sum_{t=1}^{T} G_T V_{NTt} w_{NTt} = N^{-3/2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T} G_T \tilde{X}_{Tti}\tilde{u}_{tj}$$

$$= N^{-3/2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T} G_T \tilde{X}_{Tti} u_{tj}$$

$$= N^{-3/2}\sum_{i=1}^{N}\sum_{j=1,j\neq i}^{N}\sum_{t=1}^{T} G_T \tilde{X}_{Tti} u_{tj} + N^{-3/2}\sum_{i=1}^{N}\sum_{t=1}^{T} G_T \tilde{X}_{Tti} u_{ti}$$

$$\overset{d,T\to\infty}{\longrightarrow} N^{-3/2}\sum_{i=1}^{N}\sum_{j=1,j\neq i}^{N}\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{uj}(r) + N^{-3/2}\sum_{i=1}^{N}\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)$$

$$+N^{-3/2}\sum_{i=1}^{N}\Delta_{vui}(1, 2\int_0^1 B_{vi}(r)dr, 3\int_0^1 B_{vi}(r)^2 dr)'.$$

It now follows that the last term is $O_p(N^{-1/2})$ as $N \to \infty$ because

$$\mathbb{E}|\Delta_{vui}(1, 2\int_0^1 B_{vi}(r)dr, 3\int_0^1 B_{vi}(r)^2 dr)| < \infty$$

if $\mathbb{E}|\Delta_{vui}| < \infty$, $\mathbb{E}|\Delta_{vui}\Omega_{vvi}^{1/2}| < \infty$ and $\mathbb{E}|\Delta_{vui}\Omega_{vvi}| < \infty$, which was assumed. Putting the two remaining terms together again then gives

$$N^{-3/2}\int_0^1 (\sum_{i=1}^{N}\tilde{\mathbf{B}}_{vi}(r))d(\sum_{j=1}^{N} B_{uj}(r))$$

$$= N^{-3/2}\int_0^1 (\sum_{i=1}^{N}(\mathbf{B}_{vi}(r) - \int_0^1 \mathbf{B}_{vi}(s)ds))d(\sum_{j=1}^{N} B_{uj}(r))$$

$$= N^{-3/2}\int_0^1 (\sum_{i=1}^N \mathbf{B}_{vi}(r))d(\sum_{j=1}^N B_{uj}(r)) - N^{-3/2}\sum_{i=1}^N(\int_0^1 \mathbf{B}_{vi}(s)ds)\sum_{j=1}^N B_{uj}(1)$$

$$= N^{-3/2}\int_0^1 (\sum_{i=1}^N \mathbb{E}(\mathbf{B}_{vi}(r)))d(\sum_{j=1}^N B_{uj}(r)) - N^{-3/2}\sum_{i=1}^N(\int_0^1 \mathbb{E}(\mathbf{B}_{vi}(s))ds)\sum_{j=1}^N B_{uj}(1)$$

$$+N^{-3/2}\int_0^1 (\sum_{i=1}^N(\mathbf{B}_{vi}(r)-\mathbb{E}(\mathbf{B}_{vi}(r)))d(\sum_{j=1}^N B_{uj}(r))-N^{-3/2}\sum_{i=1}^N(\int_0^1(\mathbf{B}_{vi}(s)-\mathbb{E}(\mathbf{B}_{vi}(s)))ds)\sum_{j=1}^N B_{uj}(1) \tag{56}$$

The fourth term is negligible asymptotically because

$$N^{-3/2}\sum_{i=1}^N(\int_0^1 (\mathbf{B}_{vi}(s) - \mathbb{E}(\mathbf{B}_{vi}(s)))ds)\sum_{j=1}^N B_{uj}(1) = O_p(N^{-1/2})$$

as $N \to \infty$ because $\sum_{j=1}^N B_{uj}(1) = O_p(N^{1/2})$, since $\mathbb{E}(B_{uj}(1)^2) < \infty$ if $\mathbb{E}|\Omega_{uuj}| < \infty$, and

$$\sum_{i=1}^N(\int_0^1 (\mathbf{B}_{vi}(s) - \mathbb{E}(\mathbf{B}_{vi}(s)))ds) = O_p(N^{1/2})$$

because for $p = 1, 2, 3$,

$$\mathbb{E}(\int_0^1 (\mathbf{B}^p_{vi}(s) - \mathbb{E}(\mathbf{B}^p_{vi}(s)))ds)^2 < \infty$$

if $\sup_{r\in[0,1]} \mathbb{E}(B_{vi}(r)^6) < \infty$. This last condition holds by the result of Equation (14) and the assumption $\mathbb{E}|\Omega_{vvi}|^3 < \infty$.

The third term is also negligible asymptotically. This is because by the martingale property of $\mathbf{B}_{vi}(r) - \mathbb{E}(\mathbf{B}_{vi}(r))$ and of $\sum_{j=1}^N B_{uj}(r)$, setting

$$Z_n(r) = \sum_{i=1}^N(\mathbf{B}_{vi}(r) - \mathbb{E}(\mathbf{B}_{vi}(r))),$$

we find

$$\mathbb{E}(N^{-3/2}\int_0^1 Z_n(r)d(\sum_{j=1}^N B_{uj}(r)))(N^{-3/2}\int_0^1 Z_n(r)d(\sum_{j=1}^N B_{uj}(r)))'|\Omega_1,\ldots,\Omega_N)$$

$$= N^{-3}\mathbb{E}((\sum_{j=1}^N B_{uj}(1))^2|\Omega_1,\ldots,\Omega_N)\int_0^1 \mathbb{E}(Z_n(r)Z_n(r)'|\Omega_1,\ldots,\Omega_N)dr$$

$$= N^{-3}\sum_{j=1}^N \mathbb{E}(B_{uj}(1)^2|\Omega_1,\ldots,\Omega_N)\int_0^1 \sum_{i=1}^N \mathbb{E}((\mathbf{B}_{vi}(r)-\mathbb{E}(\mathbf{B}_{vi}(r)))(\mathbf{B}_{vi}(r)-\mathbb{E}(\mathbf{B}_{vi}(r)))'|\Omega_1,\ldots,\Omega_N)dr$$

$$= N^{-1}(N^{-1}\sum_{j=1}^{N}\Omega_{uuj})N^{-1}\sum_{i=1}^{N}D_i\int_0^1 \mathbb{E}((\mathbf{W}_{vi}(r)-\mathbb{E}(\mathbf{W}_{vi}(r)))(\mathbf{W}_{vi}(r)-\mathbb{E}(\mathbf{W}_{vi}(r)))')drD_i.$$

Therefore, the third term is also $O_p(N^{-1/2})$ because

$$\mathbb{E}|\Omega_{uuj}D_i\int_0^1 \mathbb{E}((\mathbf{W}_{vi}(r)-\mathbb{E}(\mathbf{W}_{vi}(r)))(\mathbf{W}_{vi}(r)-\mathbb{E}(\mathbf{W}_{vi}(r)))')drD_i| < \infty$$

for all i and j if $\mathbb{E}|\Omega_{uui}\Omega_{vvi}^3| < \infty$, $\mathbb{E}|\Omega_{uui}| < \infty$, and $\mathbb{E}|\Omega_{vvi}|^3 < \infty$, which was assumed. The first two terms in Equation (56) equal

$$N^{-3/2}\int_0^1 (\sum_{i=1}^{N}\mathbb{E}(\tilde{\mathbf{B}}_{vi}(r)))d(\sum_{j=1}^{N}B_{uj}(r)) = N^{-1/2}\sum_{j=1}^{N}\int_0^1 \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r))dB_{uj}(r),$$

and therefore the lemma is now proven. □

Lemma 6.

$$\Sigma_2 = \mathbb{E}(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r) - \mathbb{E}(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)|\Omega_i) - \int_0^1 \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r))dB_{ui}(r))$$
$$\times(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r) - \mathbb{E}(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)|\Omega_i) - \int_0^1 \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r))dB_{ui}(r))'.$$

$$= \Sigma_1 - (1/6)\mathrm{diag}(0, \mathbb{E}(\Omega_{uui}\Omega_{vvi})\mathbb{E}(\Omega_{vvi}), 0) + (1/12)\mathrm{diag}(0, \mathbb{E}(\Omega_{uui})(\mathbb{E}(\Omega_{vvi}))^2, 0). \quad (57)$$

Proof. We have, by the definition of Σ_1,

$$\Sigma_2 = \Sigma_1 - \mathbb{E}((\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r) - \mathbb{E}(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)|\Omega_i))(\int_0^1 \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r))'dB_{ui}(r)))$$
$$-\mathbb{E}((\int_0^1 \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r))dB_{ui}(r))(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)'dB_{ui}(r) - \mathbb{E}(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)'dB_{ui}(r)|\Omega_i)))$$
$$+\mathbb{E}((\int_0^1 \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r))dB_{ui}(r))(\int_0^1 \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r))dB_{ui}(r))'). \quad (58)$$

The third term of this expression is the transpose of the second term. For the second term of the last equation, we have

$$\mathbb{E}((\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r) - \mathbb{E}(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)|\Omega_i))(\int_0^1 \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r)')dB_{ui}(r)))$$
$$= \mathbb{E}(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)\int_0^1 \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r)')dB_{ui}(r))$$

$$-\mathbb{E}(\mathbb{E}(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)|\Omega_i)\int_0^1 \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r)')dB_{ui}(r)) \tag{59}$$

and

$$\mathbb{E}(\mathbb{E}(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)|\Omega_i)\int_0^1 \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r)')dB_{ui}(r))$$
$$= \mathbb{E}(\mathbb{E}(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)|\Omega_i)\mathbb{E}(\int_0^1 \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r)')dB_{ui}(r)|\Omega_i)) = 0$$

since

$$\mathbb{E}(\int_0^1 \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r))dB_{ui}(r)|\Omega_i) = 0.$$

Now for the first term of Equation (59), we have

$$\mathbb{E}(\int_0^1 \tilde{\mathbf{B}}_{vi}(r)dB_{ui}(r)\int_0^1 \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r)')dB_{ui}(r))$$
$$= \mathbb{E}(\Omega_{uui}D_i\int_0^1 \tilde{\mathbf{W}}_{vi}(r)dW_{ui}(r)\int_0^1 \mathbb{E}(\tilde{\mathbf{W}}_{vi}(r)')dW_{ui}(r))\mathbb{E}(D_i)$$
$$= \mathbb{E}(\Omega_{uui}D_i)\int_0^1 \mathbb{E}(\tilde{\mathbf{W}}_{vi}(r))\mathbb{E}(\tilde{\mathbf{W}}_{vi}(r)')dr\mathbb{E}(D_i)$$
$$= (1/12)\text{diag}(0, \mathbb{E}(\Omega_{uui}\Omega_{vvi})\mathbb{E}(\Omega_{vvi}), 0),$$

and therefore the second and third term in Equation (58) add up to the second term in Equation (57). Finally, for the fourth term of Equation (58), we find

$$\mathbb{E}(\int_0^1 \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r))dB_{ui}(r))(\int_0^1 \mathbb{E}(\tilde{\mathbf{B}}_{vi}(r)')dB_{ui}(r))$$
$$= \mathbb{E}(\Omega_{uui})\mathbb{E}(D_i)\mathbb{E}(\int_0^1 \mathbb{E}(\tilde{\mathbf{W}}_{vi}(r))dW_{ui}(r)(\int_0^1 \mathbb{E}(\tilde{\mathbf{W}}_{vi}(r)')dW_{ui}(r)))\mathbb{E}(D_i)$$
$$= \mathbb{E}(\Omega_{uui})\mathbb{E}(D_i)\mathbb{E}(W_{ui}(1)^2)\mathbb{E}((\int_0^1 \mathbb{E}(\tilde{\mathbf{W}}_{vi}(r))\mathbb{E}(\tilde{\mathbf{W}}_{vi}(r)')dr))\mathbb{E}(D_i)$$
$$= \mathbb{E}(\Omega_{uui})\mathbb{E}(D_i)\text{diag}(0, 1/12, 0)\mathbb{E}(D_i)$$
$$= (1/12)\text{diag}(0, \mathbb{E}(\Omega_{uui})(\mathbb{E}(\Omega_{vvi}))^2, 0),$$

which forms the last term of the expression of Equation (57), thereby completing the proof. □

Lemma 7. *Let Assumption 1 hold and assume additionally that* $\mathbb{E}|\Omega_{vvi}|^3 < \infty$, $\mathbb{E}|\Omega_{uui}| < \infty$, $\mathbb{E}|\Delta_{vvi}| < \infty$, $\mathbb{E}|\Delta_{vvi}\Omega_{vvi}^{1/2}| < \infty$ *and* $\mathbb{E}|\Delta_{vvi}\Omega_{vvi}| < \infty$. *Then*

$$|N^{-1/2}\sum_{i=1}^{N} G_T \sum_{t=1}^{T} V_{NTt}v_{it} - N^{-1/2}\sum_{j=1}^{N}\int_0^1 \mathbb{E}\tilde{\mathbf{B}}_{vi}(r)dB_{vj}(r)| \overset{p,N\to\infty,T\to\infty}{\longrightarrow} 0.$$

Proof. Because $V_{NTt} = N^{-1} \sum_{i=1}^{N} \tilde{X}_{it}$ and $\tilde{v}_{jt} = v_{jt} - \bar{v}_{j.}$,

$$N^{-1/2} \sum_{i=1}^{N} G_T \sum_{t=1}^{T} V_{NTt} v_{it} = N^{-3/2} \sum_{i=1}^{N} \sum_{j=1}^{N} G_T \sum_{t=1}^{T} \tilde{X}_{Tti} v_{jt} = N^{-3/2} \sum_{i=1}^{N} \sum_{j=1}^{N} G_T \sum_{t=1}^{T} \tilde{X}_{Tti} \tilde{v}_{jt},$$

and therefore the proof of Lemma 5 applies here, mutatis mutandis. □

Supplementary Appendix C: Functionals of Brownian Motions

First, in order to find the values of M and Q, we need to calculate the $\phi(.,.)$ and $\psi(.,.)$ expressions below:

Lemma 8. *Define*

$$\phi(m,n) = \int_0^1 \int_0^1 \mathbf{1}(r<s)\mathbb{E}(W(s)^m W(r)^n)drds$$
$$\psi(m,n) = \int_0^1 \int_0^1 \mathbf{1}(r<s)(1-s)\mathbb{E}(W(s)^m W(r)^n)drds.$$

Then

$$\phi(m,n) = \sum_{j=0}^{m} \mathbf{1}(\text{j + n even})\mathbf{1}(\text{m} - \text{j even})2^{-(m+n)/2}\Gamma(j+n+1)\Gamma(m+1)/(\Gamma(j+1)\Gamma(n/2+m/2+3)),$$

and

$$\psi(m,n) = (n/2+m/2+3)^{-1}\phi(m,n).$$

Proof of Lemma 8:

Let Z denote a $N(0,1)$ distributed random variable, and note that for all integers $j \geq 0$, $\mathbb{E}(Z^j) = \mathbf{1}(j \text{ even})\pi^{-1/2}2^{j/2}\Gamma(j/2+1/2)$. Then, noting that

$$\int_0^1 \int_0^1 \mathbf{1}(r<s)r^{(j+n)/2}(s-r)^{(m-j)/2}drds$$
$$= \int_0^1 r^{(j+n)/2}(\int_{s=r}^{s=1}(s-r)^{(m-j)/2}ds)dr$$
$$= \int_0^1 r^{(j+n)/2}\Big[((m-j)/2+1)^{-1}(s-r)^{(m-j)/2+1}\Big]_{s=r}^{s=1}dr$$
$$= ((m-j)/2+1)^{-1}\int_0^1 r^{(j+n)/2}(1-r)^{(m-j)/2+1}dr$$
$$= ((m-j)/2+1)^{-1}\text{Beta}((j+n)/2+1,(m-j)/2+2)$$
$$= ((m-j)/2+1)^{-1}\Gamma((j+n)/2+1)\Gamma((m-j)/2+2)/\Gamma(n/2+m/2+3),$$

it follows that

$$\phi(m,n) = \int_0^1 \int_0^1 \mathbf{1}(r<s)\mathbb{E}(W(s)^m W(r)^n)drds$$
$$= \int_0^1 \int_0^1 \mathbf{1}(r<s)\mathbb{E}((W(s)-W(r)+W(r))^m W(r)^n)drds$$

$$= \int_0^1 \int_0^1 \mathbf{1}(r < s) \sum_{j=0}^{m} \mathbb{E}(W(r)^{j+n}) \mathbb{E}(W(s) - W(r))^{m-j} (\Gamma(m+1)/(\Gamma(j+1)\Gamma(m-j+1)) drds$$

$$= \int_0^1 \int_0^1 \mathbf{1}(r < s) \sum_{j=0}^{m} \mathbb{E}(Z^{j+n}) r^{(j+n)/2} (s-r)^{(m-j)/2} \mathbb{E}(Z^{m-j}) (\Gamma(m+1)/(\Gamma(j+1)\Gamma(m-j+1)) drds$$

$$= \sum_{j=0}^{m} \mathbb{E}(Z^{j+n}) \mathbb{E}(Z^{m-j}) (\Gamma(m+1)/(\Gamma(j+1)\Gamma(m-j+1)) \int_0^1 \int_0^1 \mathbf{1}(r < s) r^{(j+n)/2} (s-r)^{(m-j)/2} drds$$

$$= \sum_{j=0}^{m} \mathbf{1}(\mathrm{j+n\ even}) \mathbf{1}(\mathrm{m-j\ even}) \pi^{-1} 2^{(j+n)/2} 2^{(m-j)/2} \Gamma((j+n+1)/2) \Gamma((m-j+1)/2)$$

$$\times (\Gamma(m+1)/(\Gamma(j+1)\Gamma(m-j+1)))((m-j)/2+1)^{-1} \Gamma((j+n)/2+1) \Gamma((m-j)/2+2)/\Gamma(n/2+m/2+3)$$

$$= \sum_{j=0}^{m} \mathbf{1}(\mathrm{j+n\ even}) \mathbf{1}(\mathrm{m-j\ even}) \pi^{-1} 2^{(m+n)/2} \Gamma((j+n+1)/2) \Gamma((j+n+1)/2+1/2)$$

$$\times \Gamma((m-j+1)/2) \Gamma((m-j+1)/2+1/2) \ (\Gamma(m+1)/(\Gamma(j+1)\Gamma(m-j+1)))/\Gamma(n/2+m/2+3).$$

Now by the duplication formula $\Gamma(z)\Gamma(z+1/2) = 2^{1-2z} \sqrt{\pi}\ \Gamma(2z)$, the last expression can be written as

$$\sum_{j=0}^{m} \mathbf{1}(\mathrm{j+n\ even}) \mathbf{1}(\mathrm{m-j\ even}) \pi^{-1} 2^{(m+n)/2} (2^{1-2(j+n+1)/2} \Gamma(j+n+1) \sqrt{\pi})$$

$$\times (2^{1-2(m-j+1)/2} \Gamma(m-j+1) \sqrt{\pi}) (\Gamma(m+1)/(\Gamma(j+1)\Gamma(m-j+1)))/\Gamma(n/2+m/2+3)$$

$$= \sum_{j=0}^{m} \mathbf{1}(\mathrm{j+n\ even}) \mathbf{1}(\mathrm{m-j\ even}) 2^{-(m+n)/2} \Gamma(j+n+1) \Gamma(m+1)/(\Gamma(j+1)\Gamma(n/2+m/2+3)),$$

which completes the result for $\phi(m,n)$.

Now consider

$$\psi(m,n) = \int_0^1 \int_0^1 \mathbf{1}(r < s)(1-s) \mathbb{E}(W(s)^m W(r)^n) drds$$

and note that

$$\int_{s=r}^{s=1} (1-s)(s-r)^{(m-j)/2} ds = ((m-j)/2+1)^{-1} \int_{s=r}^{s=1} (1-s) d((s-r)^{(m-j)/2+1})$$

$$= ((m-j)/2+1)^{-1} \Big[(1-s)(s-r)^{(m-j)/2+1}\Big]_{s=r}^{s=1} + ((m-j)/2+1)^{-1} \int_{s=r}^{s=1} (s-r)^{(m-j)/2+1} ds$$

$$= ((m-j)/2+1)^{-1} \int_{s=r}^{s=1} (s-r)^{(m-j)/2+1} ds.$$

Therefore, following the same reasoning as before, it follows that

$$\psi(m,n) = \sum_{j=0}^{m} \mathbb{E}(Z^{j+n}) \mathbb{E}(Z^{m-j}) (\Gamma(m+1)/(\Gamma(j+1)\Gamma(m-j+1))$$

$$\times \int_0^1 r^{(j+n)/2}((m-j)/2+1)^{-1} \int_{s=r}^{s=1} (s-r)^{(m-j)/2+1} ds dr$$

$$= \sum_{j=0}^{m} \mathbb{E}(Z^{j+n})\mathbb{E}(Z^{m-j})(\Gamma(m+1)/(\Gamma(j+1)\Gamma(m-j+1))((m-j)/2+1)^{-1}((m-j)/2+2)^{-1}$$

$$\times \int_0^1 r^{(j+n)/2} \Big[(s-r)^{(m-j)/2+2}\Big]_{s=r}^{s=1} dr$$

$$= \sum_{j=0}^{m} \mathbb{E}(Z^{j+n})\mathbb{E}(Z^{m-j})(\Gamma(m+1)/(\Gamma(j+1)\Gamma(m-j+1))((m-j)/2+1)^{-1}((m-j)/2+2)^{-1}$$

$$\times \int_0^1 r^{(j+n)/2}(1-r)^{(m-j)/2+2} dr$$

$$= \sum_{j=0}^{m} \mathbb{E}(Z^{j+n})\mathbb{E}(Z^{m-j})(\Gamma(m+1)/(\Gamma(j+1)\Gamma(m-j+1))((m-j)/2+1)^{-1}((m-j)/2+2)^{-1}$$

$$\times \mathrm{Beta}((j+n)/2+1, (m-j)/2+3)$$

$$= \sum_{j=0}^{m} \mathbb{E}(Z^{j+n})\mathbb{E}(Z^{m-j})(\Gamma(m+1)/(\Gamma(j+1)\Gamma(m-j+1))((m-j)/2+1)^{-1}((m-j)/2+2)^{-1}$$

$$\times \Gamma((j+n)/2+1)\Gamma((m-j)/2+3)/\Gamma(m/2+n/2+4)$$

$$= \sum_{j=0}^{m} \mathbf{1}(\mathrm{j+n\ even})\mathbf{1}(\mathrm{m-j\ even})\pi^{-1}2^{(m+n)/2}\Gamma((j+n+1)/2)\Gamma((m-j+1)/2)$$

$$\times(\Gamma(m+1)/(\Gamma(j+1)\Gamma(m-j+1))((m-j)/2+1)^{-1}((m-j)/2+2)^{-1}$$

$$\times \Gamma((j+n)/2+1)\Gamma((m-j)/2+3)/\Gamma(m/2+n/2+4)$$

$$= \sum_{j=0}^{m} \mathbf{1}(\mathrm{j+n\ even})\mathbf{1}(\mathrm{m-j\ even})\pi^{-1}2^{(m+n)/2}\Gamma((j+n+1)/2)\Gamma((j+n+1)/2+1/2)$$

$$\times(\Gamma(m+1)/(\Gamma(j+1)\Gamma(m-j+1))\Gamma((m-j+1)/2)\Gamma((m-j+1)/2+1/2)$$

$$\times 1/\Gamma(m/2+n/2+4)$$

$$= \sum_{j=0}^{m} \mathbf{1}(\mathrm{j+n\ even})\mathbf{1}(\mathrm{m-j\ even})\pi^{-1}2^{(m+n)/2}(2^{1-2(j+n+1)/2}\Gamma(j+n+1)\sqrt{\pi})$$

$$\times(\Gamma(m+1)/(\Gamma(j+1)\Gamma(m-j+1))(2^{1-2(m-j+1)/2}\Gamma(m-j+1)\sqrt{\pi})$$

$$\times 1/\Gamma(m/2+n/2+4)$$

$$= \sum_{j=0}^{m} \mathbf{1}(\mathrm{j+n\ even})\mathbf{1}(\mathrm{m-j\ even})2^{-(m+n)/2}\Gamma(j+n+1)\Gamma(m+1)/(\Gamma(j+1)\Gamma(m/2+n/2+4))$$

$$= (n/2+m/2+3)^{-1} \sum_{j=0}^{m} \mathbf{1}(\mathrm{j+n\ even})\mathbf{1}(\mathrm{m-j\ even})2^{-(m+n)/2}\Gamma(j+n+1)\Gamma(m+1)$$

$$\times 1/(\Gamma(j+1)\Gamma(n/2+m/2+3)),$$

and therefore,

$$\psi(m,n) = (n/2+m/2+3)^{-1}\phi(m,n).$$

This concludes the proof. □

The values for $\phi(m,n)$ and $\psi(m,n)$ for $m,n = 0,1,2,\ldots,5$ are as given in Tables 9 and 10.

Table 9: The values of $\phi(m,n)$.

$\phi(m,n)$	$n=0$	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$
$m=0$	1/2	0	1/6	0	1/4	0
$m=1$	0	1/6	0	1/4	0	3/4
$m=2$	1/3	0	7/24	0	4/5	0
$m=3$	0	3/8	0	9/10	0	31/8
$m=4$	3/4	0	11/10	0	43/10	0
$m=5$	0	3/2	0	5	0	785/28

Lemma 9.

$$M = \mathbb{E}\left(\int_0^1 \tilde{\mathbf{W}}_{vi}(r)\tilde{\mathbf{W}}_{vi}(r)'dr\right) = \begin{pmatrix} 1/6 & 0 & 3/8 \\ 0 & 5/12 & 0 \\ 3/8 & 0 & 39/20 \end{pmatrix}. \tag{60}$$

Proof of Lemma 9:

For this proof, we will use the values of $\phi(.,.)$ and $\psi(.,.)$ as defined in Lemma 8 and listed in Tables 9 and Table 10. The six results needed are below:

Table 10: The values of $\psi(m,n)$ for $m,n = 1,2,\ldots,5$.

$\psi(m,n)$	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$
$n=1$	1/24	0	3/40	0	1/4
$n=2$	0	7/120	0	11/60	0
$n=3$	1/20	0	3/20	0	5/7
$n=4$	0	2/15	0	43/70	0
$n=5$	1/8	0	31/56	0	785/224

1. $M_{1,1} = \mathbb{E}(\int_0^1 W(r)^2 dr) - \mathbb{E}(\int_0^1 W(r)dr)^2 = 1/6$ because $\mathbb{E}(\int_0^1 W(r)^2 dr) = \int_0^1 \mathbb{E}(W(r)^2)dr = \int_0^1 r dr = 1/2$ and

$$\mathbb{E}(\int_0^1 W(r)dr)^2 = \int_0^1 \int_0^1 \mathbb{E}(W(r)W(s))drds = 2\int_0^1 \int_0^1 \mathbf{1}(r < s)\mathbb{E}(W(r)W(s))drds = 2\phi(1,1) = 1/3.$$

2. $M_{1,2} = 0$ because $\mathbb{E}(\int_0^1 W(r)^3 dr) = \int_0^1 \mathbb{E}(W(r)^3)dr = 0$ and

$$\mathbb{E}(\int_0^1 W(r)^2 dr \int_0^1 W(s)ds) = \int_0^1 \int_0^1 \mathbb{E}(W(r)^2 W(s))drds$$

$$= \int_0^1 \int_0^1 \mathbf{1}(r < s)\mathbb{E}(W(r)^2 W(s))drds + \int_0^1 \int_0^1 \mathbf{1}(r > s)\mathbb{E}(W(r)^2 W(s))drds$$

$$= \phi(2,1) + \phi(1,2) = 0 + 0 = 0.$$

3. $M_{1,3} = 3/8$ because

$$\mathbb{E}(\int_0^1 W(r)^4 dr) = \int_0^1 \mathbb{E}(W(r)^4)dr = \int_0^1 3r^2 dr = 1$$

and

$$\mathbb{E}(\int_0^1 W(r)dr \int_0^1 W(r)^3 dr)$$

$$= \int_0^1 \int_0^1 \mathbb{E}(W(r)^3 W(s))drds$$

$$= \int_0^1 \int_0^1 \mathbf{1}(r \leq s)\mathbb{E}(W(r)^3 W(s))drds + \int_0^1 \int_0^1 \mathbf{1}(r > s)\mathbb{E}(W(r)^3 W(s))drds$$

$$= \phi(3,1) + \phi(1,3) = 3/8 + 1/4 = 5/8.$$

4. $M_{2,2} = 5/12$ because

$$\int_0^1 (W(r)^2 - \int_0^1 W(s)^2 ds)^2 dr = \int_0^1 W(r)^4 dr - (\int_0^1 W(r)^2 dr)^2$$

and $\mathbb{E}(\int_0^1 W(r)^4)dr = \int_0^1 \mathbb{E}(W(r)^4)dr = \int_0^1 3r^2 dr = 1$ while

$$\mathbb{E}(\int_0^1 W(r)^2 dr)^2 = \int_0^1 \int_0^1 \mathbb{E}(W(r)^2 W(s)^2)drds = 2\phi(2,2) = 7/12.$$

5. $M_{2,3} = 0$ because $\mathbb{E}(\int_0^1 W(r)^5 dr) = \int_0^1 \mathbb{E}(W(r)^5)dr = 0$ and

$$\mathbb{E}((\int_0^1 W(r)^3 dr)(\int_0^1 W(s)^2 ds)) = \int_0^1 \int_0^1 \mathbb{E}(W(r)^3 W(s)^2)drds$$

$$= \phi(3,2) + \phi(2,3) = 0 + 0 = 0.$$

6. $M_{3,3} = 39/20$ because

$$\mathbb{E}(\int_0^1 (W(r)^3 - \int_0^1 W(s)^3 ds)^2 dr) = \mathbb{E}(\int_0^1 W(r)^6 dr) - \mathbb{E}(\int_0^1 W(r)^3 dr)^2$$

and

$$\mathbb{E}(\int_0^1 W(r)^6 dr) = \int_0^1 \mathbb{E}(W(r)^6) dr = \int_0^1 r^3 \mathbb{E}(r^{-1/2} W(r))^6 dr = \mathbb{E}(Z^6) \int_0^1 r^3 dr = 15/4$$

while

$$\mathbb{E}(\int_0^1 W(r)^3 dr)^2 = \int_0^1 \int_0^1 \mathbb{E}(W(r)^3 W(s)^3) dr ds = 2\phi(3,3) = 18/10,$$

and therefore,

$$M_{3,3} = 15/4 - 9/5 = 39/20.$$

Together, these results prove the assertion of the lemma. □

Finding the value for Q is a bit more work. We will argue that $Q = A - B - B' + D$, and calculate A, B, and D first, in the lemmas below:

Lemma 10.

$$A = \mathbb{E}W(1)^2 \mathbb{E}(\int_0^1 \mathbf{W}(r)\mathbf{W}(r)' dr) = \begin{pmatrix} 1/2 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 15/4 \end{pmatrix}.$$

Proof. This follows because

$$A = \int_0^1 \mathbb{E} \begin{pmatrix} W(r)^2 & W(r)^3 & W(r)^4 \\ W(r)^3 & W(r)^4 & W(r)^5 \\ W(r)^4 & W(r)^5 & W(r)^6 \end{pmatrix} dr$$

$$= \int_0^1 \begin{pmatrix} r\mathbb{E}Z^2 & 0 & r^2\mathbb{E}Z^4 \\ 0 & r^2\mathbb{E}Z^4 & 0 \\ r^2\mathbb{E}Z^4 & 0 & r^3\mathbb{E}Z^6 \end{pmatrix} dr = \begin{pmatrix} 1/2 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 15/4 \end{pmatrix}.$$

□

Lemma 11.

$$B = \mathbb{E}W(1)(\int_0^1 \mathbf{W}(r) dW(r))(\int_0^1 \mathbf{W}(s) ds)' = \begin{pmatrix} 1/2 & 0 & 7/4 \\ 0 & 7/6 & 0 \\ 7/8 & 0 & 9/2 \end{pmatrix}.$$

Proof. Note that

$$B_{ij} = \mathbb{E}W(1)(\int_0^1 W(r)^i dW(r))(\int_0^1 W(r)^j dr)$$

and because

$$f(W(1)) - f(W(0)) = \int_0^1 f'(W(r))dW(r) + (1/2)\int_0^1 f''(W(r))dr,$$

we have for $i = 1, 2, 3$

$$(1+1)^{-1}W(1)^{i+1} - (1+1)^{-1}W(0)^{i+1} = \int_0^1 W(r)^i dW(r) + (i/2)\int_0^1 W(r)^{i-1}dr,$$

so

$$\int_0^1 W(r)^i dW(r) = (i+1)^{-1}W(1)^{i+1} - (i/2)\int_0^1 W(r)^{i-1}dr.$$

Therefore

$$\begin{aligned}
B_{ij} &= \mathbb{E}W(1)(\int_0^1 W(r)^i dW(r))(\int_0^1 W(r)^j dr) \\
&= \mathbb{E}W(1)((i+1)^{-1}W(1)^{i+1} - (i/2)\int_0^1 W(r)^{i-1}dr)(\int_0^1 W(r)^j dr) \\
&= \mathbb{E}(i+1)^{-1}W(1)^{i+2}\int_0^1 W(r)^j dr - (i/2)W(1)\int_0^1 W(r)^{i-1}dr)(\int_0^1 W(s)^j ds) \\
&= B^1_{ij} - B^2_{ij}.
\end{aligned}$$

Now

$$\begin{aligned}
B^1_{ij} &= \mathbb{E}(i+1)^{-1}W(1)^{i+2}\int_0^1 W(r)^j dr \\
&= (i+1)^{-1}\chi(i+2, j)
\end{aligned}$$

where

$$\begin{aligned}
\chi(m,n) &= \mathbb{E}W(1)^m \int_0^1 W(r)^n dr \\
&= \mathbb{E}W(1)^m \int_0^1 W(r)^n dr \\
&= \int_0^1 \mathbb{E}(W(1) - W(r) + W(r))^m W(r)^n dr \\
&= \int_0^1 \sum_{j=0}^m \mathbb{E}W(r)^{j+n}\mathbb{E}(W(1) - W(r))^{m-j} dr \Gamma(m+1)/(\Gamma(j+1)\Gamma(m-j+1)) \\
&= \sum_{j=0}^m \int_0^1 r^{j/2+n/2}(1-r)^{m/2-j/2} dr \mathbb{E}Z^{j+n}\mathbb{E}Z^{m-j}\Gamma(m+1)/(\Gamma(j+1)\Gamma(m-j+1))
\end{aligned}$$

Table 11: The values of $\chi(m,n)$.

$\chi(m,n)$	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$
$m=1$	1/2	0	1	0	15/4
$m=2$	0	7/6	0	4	0
$m=3$	3/2	0	9/2	0	93/4
$m=4$	0	11/2	0	129/5	0
$m=5$	15/2	0	30	0	785/4

$$= \sum_{j=0}^{m} B(j/2+n/2+1, m/2-j/2+1)[I(\text{j}+\text{n even})2^{-j/2-n/2}(\Gamma(j+n+1)/\Gamma(j/2+n/2+1))]$$

$$\times[I(\text{m}-\text{j even})2^{-m/2+j/2}(\Gamma(m-j+1)/\Gamma(m/2-j/2+1))]\Gamma(m+1)/(\Gamma(j+1)\Gamma(m-j+1))$$

$$= \sum_{j=0}^{m}(\Gamma(j/2+n/2+1)\Gamma(m/2-j/2+1)/\Gamma(j/2+n/2+m/2-j/2+2))$$

$$\times[I(\text{j}+\text{n even})2^{-j/2-n/2}(\Gamma(j+n+1)/\Gamma(j/2+n/2+1))]$$

$$\times[I(\text{m}-\text{j even})2^{-m/2+j/2}(\Gamma(m-j+1)/\Gamma(m/2-j/2+1))]\Gamma(m+1)/(\Gamma(j+1)\Gamma(m-j+1))$$

$$= 2^{-m/2-n/2}\sum_{j=0}^{m} I(\text{j}+\text{n even})I(\text{m}-\text{j even})\Gamma(j+n+1)\Gamma(m+1)/(\Gamma(j+1)\Gamma(m/2+n/2+2)).$$

Noting that the last expression equals

$$\frac{\sum_{j=0}^{m} I(\text{j}+\text{n even})I(\text{m}-\text{j even})\Gamma(j+n+1)\Gamma(m+1)/\Gamma(j+1)}{\Gamma(m/2+n/2+2)2^{m/2+n/2}},$$

and noting that the numerator and denominator are integers, we can now calculate the results of Table 11.

Given the values for $\chi(.,.)$, the values of $B^1_{ij} = \mathbb{E}(i+1)^{-1}W(1)^{i+2}\int_0^1 W(r)^j dr = (i+1)^{-1}\chi(i+2,j)$ can now be calculated. Noting that

$$B^1_{11} = 2^{-1}\chi(3,1) = (1/2)(3/2) = 3/4,$$

$$B^1_{21} = 3^{-1}\chi(4,1) = 0,$$

$$B^1_{31} = 4^{-1}\chi(5,1) = (1/4)(15/2) = 15/8,$$

$$B^1_{12} = 2^{-1}\chi(3,2) = 0,$$

$$B^1_{22} = 3^{-1}\chi(4,2) = (1/3)(11/2) = 11/6,$$

$$B^1_{32} = 4^{-1}\chi(5,2) = 0,$$

$$B^1_{13} = 2^{-1}\chi(3,3) = (1/2)(9/2) = 9/4,$$

$$B^1_{23} = 3^{-1}\chi(4,3) = 0,$$

$$B^1_{33} = 4^{-1}\chi(5,3) = (1/4)(30) = 15/2,$$

we have

$$B_1 = \begin{pmatrix} 3/4 & 0 & 9/4 \\ 0 & 11/6 & 0 \\ 15/8 & 0 & 15/2 \end{pmatrix}.$$

Also, for $i \geq 1$,

$$B^2_{ij} = (i/2)W(1)\int_0^1 W(r)^{i-1}dr)(\int_0^1 W(s)^j ds) = (i/2)\omega(i-1,j)$$

where

$$\begin{aligned}\omega(m,n) &= \mathbb{E}W(1)\int_0^1 W(r)^m dr)(\int_0^1 W(s)^n ds)\\ &= \mathbb{E}\int_0^1\int_0^1 W(1)W(r)^m W(s)^n I(r<s)drds\\ &\quad +\mathbb{E}\int_0^1\int_0^1 W(1)W(r)^m W(s)^n I(r>s)drds\\ &= \mathbb{E}\int_0^1\int_0^1 (W(1)-W(s)+W(s))W(r)^m W(s)^n I(r<s)drds\\ &\quad +\mathbb{E}\int_0^1\int_0^1 (W(1)-W(r)+W(r))W(r)^m W(s)^n I(r>s)drds\\ &= \mathbb{E}\int_0^1\int_0^1 W(r)^m W(s)^{n+1} I(r<s)drds + \mathbb{E}\int_0^1\int_0^1 W(r)^{m+1} W(s)^n I(r>s)drds\\ &= \phi(n+1,m) + \phi(m+1,n),\end{aligned}$$

and therefore,

$$B^2_{ij} = (i/2)\omega(i-1,j) = (i/2)\phi(j+1,i-1) + (i/2)\phi(i,j).$$

Using the table with the ϕ-values, we can now also calculate the values for the B^2_{ij}:

$$B^2_{11} = (1/2)\phi(2,0) + (1/2)\phi(1,1) = (1/2)(1/3) + (1/2)(1/6) = 1/4,$$

$$B_{12}^2 = (1/2)\phi(3,0) + (1/2)\phi(1,2) = 0,$$

$$B_{13}^2 = (1/2)\phi(4,0) + (1/2)\phi(1,3) = (1/2)(3/4) + (1/2)(1/4) = 1/2,$$

$$B_{21}^2 = \phi(2,1) + \phi(1,2) = 0,$$

$$B_{22}^2 = \phi(3,1) + \phi(2,2) = 3/8 + 7/24 = 2/3,$$

$$B_{23}^2 = \phi(4,1) + \phi(2,3) = 0,$$

$$B_{31}^2 = (3/2)\phi(2,2) + (3/2)\phi(3,1) = (3/2)(7/24) + (3/2)(3/8) = 1,$$

$$B_{32}^2 = (3/2)\phi(3,2) + (3/2)\phi(3,2) = 0,$$

$$B_{33}^2 = (3/2)\phi(4,2) + (3/2)\phi(3,3) = (3/2)(11/10) + (3/2)(9/10) = 3.$$

which then gives

$$B_2 = \begin{pmatrix} 1/4 & 0 & 1/2 \\ 0 & 2/3 & 0 \\ 1 & 0 & 3 \end{pmatrix}$$

We now get

$$B = B_1 - B_2 = \begin{pmatrix} 3/4 & 0 & 9/4 \\ 0 & 11/6 & 0 \\ 15/8 & 0 & 15/2 \end{pmatrix} - \begin{pmatrix} 1/4 & 0 & 1/2 \\ 0 & 2/3 & 0 \\ 1 & 0 & 3 \end{pmatrix}$$

$$= \begin{pmatrix} 1/2 & 0 & 7/4 \\ 0 & 7/6 & 0 \\ 7/8 & 0 & 9/2 \end{pmatrix},$$

as asserted. □

Lemma 12.

$$D = \mathbb{E}(W(1)^2(\int_0^1 \mathbf{W}(r)dr)(\int_0^1 \mathbf{W}(s)ds)') = \begin{pmatrix} 5/6 & 0 & 101/40 \\ 0 & 139/60 & 0 \\ 101/40 & 0 & 103/10 \end{pmatrix}.$$

Proof. We will first show that

$$D_{mn} = \mathbb{E}(W(1)^2 \int_0^1 \int_0^1 W(r)^m W(s)^n drds) = \psi(m,n) + \phi(m+2,n) + \psi(n,m) + \phi(n+2,m).$$

This follows because

$$\mathbb{E}(W(1)^2 \int_0^1 W(r)^m dr \int_0^1 W(s)^n ds)$$

$$= \int_0^1 \int_0^1 \mathbf{1}(r > s)\mathbb{E}(W(r)^m W(s)^n W(1)^2) drds + \int_0^1 \int_0^1 \mathbf{1}(r < s)\mathbb{E}(W(r)^m W(s)^n W(1)^2) drds$$

$$= \int_0^1 \int_0^1 \mathbf{1}(r > s)\mathbb{E}(W(r)^m W(s)^n (W(1) - W(r) + W(r))^2) drds$$

$$+ \int_0^1 \int_0^1 \mathbf{1}(r < s)\mathbb{E}(W(r)^m W(s)^n (W(1) - W(s) + W(s))^2) drds$$

$$= \int_0^1 \int_0^1 \mathbf{1}(r > s)(1 - r)\mathbb{E}(W(r)^m W(s)^n)) drds$$

$$+ \int_0^1 \int_0^1 \mathbf{1}(r > s)\mathbb{E}(W(r)^{m+2} W(s)^n) drds$$

$$+ \int_0^1 \int_0^1 \mathbf{1}(r < s)(1 - s)\mathbb{E}(W(r)^m W(s)^n drds$$

$$+ \int_0^1 \int_0^1 \mathbf{1}(r < s)\mathbb{E}(W(r)^m W(s)^{n+2}) drds$$

$$= \psi(m, n) + \phi(m + 2, n) + \psi(n, m) + \phi(n + 2, m).$$

Using our tables for $\phi(.,.)$ and $\psi(.,.)$, we now find

$$D = \begin{pmatrix} 5/6 & 0 & 101/40 \\ 0 & 139/60 & 0 \\ 101/40 & 0 & 103/10 \end{pmatrix}.$$

□

Lemma 13.

$$Q = \mathbb{E}(\int_0^1 \tilde{\mathbf{W}}(r) dW(r))(\int_0^1 \tilde{\mathbf{W}}(r) dW(r))') = \begin{pmatrix} 1/3 & 0 & 9/10 \\ 0 & 59/60 & 0 \\ 9/10 & 0 & 101/20 \end{pmatrix}. \tag{61}$$

Proof. Note that

$$Q = \mathbb{E}(\int_0^1 \tilde{\mathbf{W}}(r) dW(r))(\int_0^1 \tilde{\mathbf{W}}(r) dW(r))'$$

$$= \mathbb{E}(\int_0^1 (\mathbf{W}(r) - \int_0^1 \mathbf{W}(s) ds) dW(r))(\int_0^1 (\mathbf{W}(r) - \int_0^1 \mathbf{W}(s) ds) dW(r))'$$

$$
\begin{aligned}
&= \mathbb{E}(\int_0^1 \mathbf{W}(r)dW(r) - W(1)\int_0^1 \mathbf{W}(s)ds)(\int_0^1 \mathbf{W}(r)dW(r) - W(1)\int_0^1 \mathbf{W}(s)ds)' \\
&= \mathbb{E}(\int_0^1 \mathbf{W}(r)dW(r)))(\int_0^1 \mathbf{W}(r)dW(r))' \\
&-\mathbb{E}W(1)(\int_0^1 \mathbf{W}(r)dW(r))(\int_0^1 \mathbf{W}(s)ds)' \\
&-\mathbb{E}W(1)(\int_0^1 \mathbf{W}(s)ds)(\int_0^1 \mathbf{W}(r)dW(r))' \\
&+\mathbb{E}W(1)^2(\int_0^1 \mathbf{W}(r)dr)(\int_0^1 \mathbf{W}(s)ds)' \\
&= A - B - B' + D,
\end{aligned}
$$

and

$$
\begin{aligned}
& A - B - B' + D \\
&= \begin{pmatrix} 1/2 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 15/4 \end{pmatrix} - \begin{pmatrix} 1/2 & 0 & 7/4 \\ 0 & 7/6 & 0 \\ 7/8 & 0 & 9/2 \end{pmatrix} \\
&- \begin{pmatrix} 1/2 & 0 & 7/8 \\ 0 & 7/6 & 0 \\ 7/4 & 0 & 9/2 \end{pmatrix} + \begin{pmatrix} 5/6 & 0 & 101/40 \\ 0 & 139/60 & 0 \\ 101/40 & 0 & 103/10 \end{pmatrix} \\
&= \begin{pmatrix} 1/3 & 0 & 9/10 \\ 0 & 59/60 & 0 \\ 9/10 & 0 & 101/20 \end{pmatrix}.
\end{aligned}
$$