# ESTIMATION AND INFERENCE IN COINTEGRATED PANELS

By

Yi Li

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Economics—Doctor of Philosophy

2017

**ABSTRACT**

ESTIMATION AND INFERENCE IN COINTEGRATED PANELS

By

Yi Li

This dissertation investigates parameter estimation and inference in cointegrated panel data model. In Chapter 1, for homogeneous cointegrated panels, a simple, new estimation method is proposed based on Vogelsang and Wagner [2014]. The estimator is labeled panel integrated modified ordinary least squares (panel IM-OLS). Similar to panel fully modified ordinary least squares (panel FM-OLS) and panel dynamic ordinary least squares (panel DOLS), the panel IM-OLS estimator has a zero mean Gaussian mixture limiting distribution. However, panel IM-OLS does not require estimation of long run variance matrices and avoids the need to choose tuning parameters such as kernel functions, bandwidths, leads and lags. Inference based on panel IM-OLS estimates does require an estimator of a scalar long run variance, and critical values for test statistics are obtained from traditional and fixed-b methods. The properties of panel IM-OLS are analyzed using asymptotic theory and finite sample simulations. Panel IM-OLS performs well relative to other estimators. Chapter 2 compares asymptotic and bootstrap hypothesis tests in cointegrated panels with cross-sectional uncorrelated units and endogenous regressors. All the tests are based on the panel IM-OLS estimator from Chapter 1. The aim of using the bootstrap tests is to deal with the size distortion problems in the finite samples of fixed-b tests. Finite sample simulations show that the bootstrap method outperforms the asymptotic method in terms of having lower size distortions. In general, the stationary bootstrap is better than the conditional-on-regressors bootstrap, although in some cases, the conditional-on-regressors bootstrap has less size distortions. The improvement in size comes with only minor power losses, which can be ignored

when the sample size is large. Chapter 3 is concerned with parameter estimation and inference in a more general case than Chapter 1 with endogenous regressors and heterogeneous long run variances in the cross section. In addition, the model allows a limited degree of cross-sectional dependence due to a common time effect. The panel IM-OLS estimator is provided for this less restricted model. Similar as in Chapter 1, this panel IM-OLS estimator has a zero mean Gaussian mixture limiting distribution. However, standard asymptotic inference is infeasible due to the existence of nuisance parameters. Inference based on panel IM-OLS relies on the stationary bootstrap. The properties of panel IM-OLS are analyzed using the stationary bootstrap in finite sample simulations.

This dissertation is dedicated to my parents and my wife.
Thank you for nursing me with love and always believing in me.

# ACKNOWLEDGMENTS

The past 5 years have been a wonderful journey for me and I would never have been able to get this far without the support from my professors, my friends and my family. It is my pleasure to acknowledge people who have given me guidance, help and encouragement.

First and foremost I would like to gratefully and sincerely thank my PhD advisor, Professor Tim Vogelsang, for his guidance, understanding and patience. His Time Series course is definitely my favorite class, which led my way to research field. I really appreciate all the support and encouragement he has given me in my research. In particular, I will never forget the effort he spent on revising this dissertation. For everything you have done for me, Professor Vogelsang, I thank you.

I would also like to express my deepest gratitude to my committee members Professor Peter Schmidt and Professor Jeffrey Wooldridge. Thank both of you for your guidance and help during my graduate study, I cannot obtain my PhD without your support. It has been such an honor for me being your student. Special thanks go to Professor Martin Wagner for his important suggestions and remarks on my dissertation as well as the effort he spent as my committee member. I must thank all of the professors from Economics Department who I learned from during the last five years for showing me what is a researcher meant to be, I will keep it in mind in my future career.

I would like to give special thanks to my cohort in Economics Department. I really enjoyed the time we studied together. We discussed, debated, exchanged ideas and learned from each other. It is definitely one of the best time periods in my life. My thanks also go to my friends in Math Department. In particular, I would like to thank Xin Yang for his expertise and patience when I had tough math problems.

Lastly, I would like to thank my family for all their love and the faith they put in me. For my parents who raised me with endless love and supported me in all my pursuits. For my grandfather who gave me all his trust and encouraged me all the time. Most importantly, I would like to express my deepest appreciation to my wife, Yanbo Shen. Without her love, understanding and encouragement, I would not have been able to finish this work.

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# KEY TO ABBREVIATIONS

OLS Ordinary Least Square

IM-OLS Integrated Modified Ordinary Least Square

FM-OLS Fully Modified Ordinary Least Square

DOLS Dynamic Ordinary Least Square

HAC Heteroskedastic and Autocorrelation Consistent RMSE Root Mean Squared Error

Fb Fixed-b

QS Quadratic Spectral

# Chapter 1

# Integrated modified OLS estimation and fixed-b inference for homogeneous cointegrated panels

This paper is concerned with parameter estimation and inference in homogeneous cointegrated panels. We propose a simple, new estimation method originated from Vogelsang and Wagner [2014]. The estimator is labeled panel integrated modified ordinary least squares (panel IM-OLS). Similar to panel fully modified ordinary least squares (panel FM-OLS) and panel dynamic ordinary least squares (panel DOLS), the panel IM-OLS estimator has a zero mean Gaussian mixture limiting distribution. However, panel IM-OLS does not require estimation of long run variance matrices and avoids the need to choose tuning parameters such as kernel functions, bandwidths, leads and lags. Inference based on panel IM-OLS estimates does require an estimator of a scalar long run variance, and we propose both traditional and fixed-b methods for obtaining critical values for test statistics. The properties of panel IM-OLS are analyzed using asymptotic theory and finite sample simulations. Panel IM-OLS performs well relative to other estimators.

## 1.1 Introduction

This paper considers the extension of the pure time series integrated modified ordinary least squares (IM-OLS) method of Vogelsang and Wagner [2014] for estimating and testing hypotheses about a cointegrating vector to a balanced panel of $N$ individuals observed over $T$ time periods. We call the estimator panel IM-OLS. We derive its limiting distribution and provide a finite sample simulation of panel IM-OLS compared with pooled OLS, panel fully modified OLS (panel FM-OLS) and panel dynamic OLS (panel DOLS).

It is well-known that in panel cointegration regression, when the regressors are endogenous, the limiting distribution of the pooled OLS estimator is contaminated by second order bias terms. Inference is difficult in this situation because the nuisance parameters cannot be removed by simple scaling methods. Consequently, panel FM-OLS and panel DOLS were proposed, which both deal with the endogeneity problem and lead to zero mean Gaussian mixture limiting distributions and in turn make standard asymptotic inference available.

The panel IM-OLS estimator is based on pooled OLS estimation of a partial sum transformation of the cointegrating panel regression. Similar to the panel FM-OLS and panel DOLS estimators, the panel IM-OLS estimator also has a zero mean Gaussian mixture limiting distribution, but it has advantage compared with its two counterparts. Panel IM-OLS estimator avoids kernel function and bandwidth choices for long run variance estimation, which is required by panel FM-OLS, and leads and lags choices to expand the regression, which is required by panel DOLS. However, for inference, panel IM-OLS does need to estimate a scalar long run variance parameter.

The limit theory considered here is obtained for a fixed number of cross-sectional units $N$, letting $T \to \infty$. This limit theory is widely used in empirical macroeconomics, empirical energy economics and empirical finance problems. In this case, even though the panel IM-OLS estimator converges to a zero mean Gaussian mixture distribution, inference based on this estimator still requires the estimation of a long run variance parameter. As in Vogelsang and Wagner [2014], there are two solutions for this problem. First, standard asymptotic inference based on a consistent estimator of the long run variance and second, fixed-b inference. The latter solution has its own benefit over standard asymptotic theory because fixed-b inference captures the impact of kernel and bandwidth choices on test statistics based upon them, whereas standard asymptotic theory does not. As will be discussed in detail later, the pooled OLS residuals of the panel IM-OLS regression need to be further adjusted to obtain pivotal fixed-b test statistics.

All estimators and tests in this paper are derived for a cross-sectionally uncorrelated homogeneous panel. For many applications this unrealistic assumption is still commonly employed when developing panel cointegration methods, especially for estimation procedures. Only a few and partial results concerning both cointegration estimation and inference are available for cross-sectionally dependent panels to date. One branch of the literature considers panel data with spatial interaction among cross-sectional units (e.g., Kapoor et al. [2007]; Yu et al. [2008], [2010], [2012]). An alternative to the spatial approach is the factor structure approach, which can capture common stochastic shocks and trends (e.g., Bai and Ng [2004]). Bai and Kao [2006] derive an extension of FM-OLS estimation to panels with short-run cross-sectional correlation. Pesaran [2006] proposes the Common Correlated Effects (CCE) approach to estimation of panel data models with multi-factor error structure, which is further developed by Kapetanios et al. [2011] allowing for nonstationary common factors. The estimation and inference is challenging for cross-sectionally dependent heterogeneous panel with endogenous regressors, but our ongoing work shows that the methods developed in this paper, with some modifications, will be able to estimate the parameter and make valid inference in that scenario.

After the theoretical analysis, we provide a finite sample simulation study to assess the performance of the estimators and tests. Benchmarks are given by pooled OLS, panel FM-OLS and panel DOLS. In the simulations, panel IM-OLS performs relatively well with smaller bias and only slightly larger RMSE than other estimators. The simulations of size and power of the tests show that fixed-b test statistics based on the panel IM-OLS estimator lead to the smallest size distortions at the price of only minor losses in size-corrected power.

The remainder of the paper is organized as follows. In the next section we present a standard panel cointegrating regression and review several key results of the benchmark estimators. Section 1.3 describes the panel IM-OLS estimator and its asymptotic distribution. Inference using the panel IM-OLS parameter estimator is discussed. Section 1.4 reports the finite sample bias and root mean squared error of the various estimators. Section 1.5 assesses the finite sample performance of the test statistics described in Section 1.3. Section 1.6 concludes the paper. Appendix contains the proofs of this paper.

## 1.2 Homogeneous cointegrated panels for benchmark estimators

Consider the following data generating process

$$y_{it} = \mu + x_{it}'\beta + u_{it} \tag{1.1}$$

$$x_{it} = x_{it-1} + v_{it} \tag{1.2}$$

where $y_{it}$ and $u_{it}$ are scalars, $x_{it}$ and $v_{it}$ are $k \times 1$ vectors with sub-index $i = 1, 2, \cdots, N$ for the $i^{th}$ cross sectional unit, sub-index $t = 1, 2, \cdots T$ for the time period; $\beta$ is $k \times 1$ vector of the slope parameters. For notational brevity here we only include the intercept $\mu$ as the deterministic component (later when we discuss the panel IM-OLS estimator, we will extend it into more general deterministic time trends such as $\mu_0 + \mu_1 t + \cdots + \mu_{p-1} t^{p-1}$). Define the error vector as $\eta_{it} = \begin{bmatrix} u_{it}, & v_{it}' \end{bmatrix}'$. It is assumed that $\eta_{it}$ is a vector of $I(0)$ processes, in which case $x_{it}$ is a non-cointegrating vector of $I(1)$ processes and there exists a cointegrating relationship among $\begin{bmatrix} y_{it}, & x_{it}' \end{bmatrix}'$ with cointegrating vector $\begin{bmatrix} 1, & -\beta' \end{bmatrix}'$.

**Assumption 1.** *Assume that $\{\eta_{it}\}_{i=1}^{N}$ are cross-sectionally uncorrelated and theirs 2nd order moment is constant.*

Note that the Assumption 1 only requires the panels are homogeneous in the 2nd order moment, it's possible that the higher order moment structure are heterogeneous across $i$.

**Assumption 2.** *Assume that $\eta_{it}$ satisfies a functional central limit theorem (FCLT) of the form*

$$T^{-1/2} \sum_{t=1}^{[rT]} \eta_{it} \Rightarrow B_i(r) = \Omega^{1/2} W_i(r), \quad r \in [0, 1].$$

In Assumption 2, $[rT]$ represents the integer part of $rT$, and $W_i(r)$ is a $(k+1) \times 1$ vector of independent standard Brownian motions. $\Omega^{1/2}$ is a $(k+1) \times (k+1)$ matrix that satisfies: $\Omega = \Omega^{1/2}\left(\Omega^{1/2}\right)'$, and

$$\Omega = \sum_{j=-\infty}^{\infty} \mathbb{E}(\eta_{it}\eta'_{it-j}) = \begin{bmatrix} \Omega_{uu} & \Omega_{uv} \\ \Omega_{vu} & \Omega_{vv} \end{bmatrix} > 0,$$

where it is clear that $\Omega_{vu} = \Omega'_{uv}$. The assumption $\Omega_{vv} > 0$ rules out cointegration in $x_{it}$. Partition $B_i(r)$ as

$$B_i(r) = \begin{bmatrix} B_{u,i}(r) \\ B_{v,i}(r) \end{bmatrix},$$

and likewise partition $W_i(r)$ as

$$W_i(r) = \begin{bmatrix} w_{u,i}(r) \\ W_{v,i}(r) \end{bmatrix},$$

where $w_{u,i}(r)$ and $W_{v,i}(r)$ are a scalar and a k-dimensional standard Brownian motion respectively. Using the Cholesky form of $\Omega^{1/2}$,

$$\Omega^{1/2} = \begin{bmatrix} \sigma_{u \cdot v} & \lambda_{uv} \\ 0_{k \times 1} & \Omega_{vv}^{1/2} \end{bmatrix},$$

it can be shown that $\sigma^2_{u \cdot v} = \Omega_{uu} - \Omega_{uv}\Omega_{vv}^{-1}\Omega_{vu}$ and $\lambda_{uv} = \Omega_{uv}\left(\Omega_{vv}^{-1/2}\right)'$. By this Cholesky decomposition, we can write

$$B_i(r) = \begin{bmatrix} B_{u,i}(r) \\ B_{v,i}(r) \end{bmatrix} = \begin{bmatrix} \sigma_{u\cdot v} w_{u,i}(r) + \lambda_{uv} W_{v,i}(r) \\ \Omega_{vv}^{1/2} W_{v,i}(r) \end{bmatrix}.$$

Next define the one-sided long run covariance matrix. For each $i \in [1, 2, \cdots, N]$,

$$\Lambda = \sum_{j=1}^{\infty} \mathbb{E}(\eta_{i,t-j}\eta_{it}') = \begin{bmatrix} \Lambda_{uu} & \Lambda_{uv} \\ \Lambda_{vu} & \Lambda_{vv} \end{bmatrix}.$$

Also define the contemporary covariances, that is, for each $i \in [1, 2, \cdots, N]$,

$$\Sigma = \mathbb{E}(\eta_{it}\eta_{it}') = \begin{bmatrix} \Sigma_{uu} & \Sigma_{uv} \\ \Sigma_{vu} & \Sigma_{vv} \end{bmatrix}.$$

Note that $\mathit{\Delta} = \Sigma + \Lambda$ is half long run variance, and it is likewise partitioned as

$$\mathit{\Delta} = \sum_{j=0}^{\infty} \mathbb{E}(\eta_{i,t-j}\eta_{it}') = \begin{bmatrix} \mathit{\Delta}_{uu} & \mathit{\Delta}_{uv} \\ \mathit{\Delta}_{vu} & \mathit{\Delta}_{vv} \end{bmatrix}.$$

The long run variance, $\Omega$, is related to $\Lambda$ and $\Sigma$ as $\Omega = \Sigma + \Lambda + \Lambda'$.

**Remark 1.** *1. If we do have heterogeneity in the 2nd order moment structure, i.e. $\Omega_i$, $\Lambda_i$, $\Sigma_i$ and $\Delta_i$ are varied for different $i$, then even though we can estimate those moments individual by individual, however, finding pivotal fixed-b statistics is challenging, and we haven't found it yet. In this case, one possible way to make valid inference is using bootstrap to mimic those non-pivotal distributions. The stationary bootstrap is one method that we could apply in this scenario.*

2. *If $\eta_{it}$ are cross-sectionally dependent, then the inference is much more complicated. Spatial approach and factor structure approach are possible solutions according to different dependence assumptions. We haven't been able to find a way to deal with the general cross-sectional dependence case. However, if the dependence only originates from time fixed-effect dummy variables, then the methods developped in this paper will go through with some natural modifications, and the bootstrap, both over time as well as across units, is needed for valid inference.*

As mentioned before, the benchmark estimators are the pooled OLS, the panel FM-OLS and the panel DOLS estimators. To conserve space, we don't provide detail results of all those estimators. But we do want to review several key results for those estimators. For the pooled OLS estimator, when the regressors are endogenous, it has an asymptotic bias due to the nuisance parameters $\varDelta_{vu}$, which cannot be removed by simple scaling methods. The panel FM-OLS estimator as considered here is an extension of the FM-OLS estimator of Phillips and Hansen [1990], which is designed to asymptotically remove $\varDelta_{vu}$ and to deal with the correlation between $B_{u,i}(r)$ and $B_{v,i}(r)$. Conditional on $B_{v,i}(r)$ for all $i = 1, 2, \ldots, N$, the limit of the scaled panel FM-OLS estimator is a mean zero mixture of normals. Asymptotically pivotal $t$ and Wald statistics with $N(0, 1)$ and chi-square limiting distributions can be constructed by estimating $\sigma^2_{u \cdot v}$. The panel DOLS estimator considered here is almost identical to Mark and Sul [2003]. The only difference is that there is no fixed effect in the data generating process (1.1). The homogeneous panel DOLS estimator of $\beta$ has the same limiting distribution as the homogeneous panel FM-OLS estimator. Hence, they are asymptotically equivalent. This result was shown by Kao and Chiang [2000], and it also can be extended to heterogeneous panels.

## 1.3 Panel integrated modified OLS

### 1.3.1 Panel IM-OLS estimator

In this section, we present a new estimator for homogeneous cointegrated panels. This estimator is an extension of Vogelsang and Wagner [2014], who propose the IM-OLS estimator for the time series case. The transformation used by IM-OLS provides an asymptotically unbiased estimator with a zero mean Gaussian mixture limiting distribution. Compared with panel FM-OLS, the transformation does not require estimators of $\Omega$, so the choice of bandwidth and kernel is avoided for parameter estimation. We consider a slightly more general version of (1.1) given by

$$y_{it} = D_t'\delta_i + x_{it}'\beta + u_{it}, \tag{1.3}$$

where $\delta_i$ and $\beta$ are $p \times 1$ and $k \times 1$ parameter vectors respectively, $x_{it}$ continues to follow (1.2) and for the deterministic component, $D_t$, we assume that there is a $p \times p$ matrix $G_D$ and a vector of functions, $D(s)$, such that

$$\lim_{T\to\infty} \sqrt{T} G_D^{-1} D_{[sT]} = D(s) \;\; with \;\; 0 < \int_0^r D(s)D(s)'ds < \infty, \;\; 0 < r \leqslant 1. \tag{1.4}$$

The deterministic component $D_t$ could include an intercept, time trend and polynomials of the time trend.

**Remark 2.** *1. Note that, in regression (1.3), the intercept from $D_t$ and $\delta_i$ together allow fixed effect estimation of the system. In a simpler case, suppose that $\delta_i$ is the same* *constant for all $i$ so that*

$$y_{it} = D_t'\delta + x_{it}'\beta + u_{it}.$$

*In this case, the estimation and inference procedures introduced later in the paper go through with minor changes.*

2. *In regression (1.3), $\beta$ is a constant for all $i$, which means the same long-run relation between $y_{it}$ and $x_{it}$ applies for all $i$. As in Philips and Moon [1999], we could also allow this coefficient differs randomly across $i$, which leads to the heterogeneous panel cointegration model. In that case, as long as the error vectors are uncorrelated across $i$ and their 2nd order moments are constant, then the results will be similar as what we have in this paper. Otherwise, if the panel has heterogeneity in both cointegration relation and 2nd moment structure, then inference is challenging and might need to apply boostrap.*

3. *We could consider a more traditional panel data setting model like*

$$y_{it} = \mu_i + \lambda_t + x_{it}'\beta + u_{it}.$$

*In this case, after the time effect $\lambda_t$ being eliminated by cross-sectional demeaning, and if it is also a homogeneous panel, then the estimation and inference procedures will be similar as in this paper. But if there is heterogeneity in the sencond moment structure, even though the estimation of the $\beta$ will not be affected, however, the inference is much more complicated as we disscussed in Remark 1, and bootstrap method could be used for the inference.*

Computing the partial sum of both sides of (1.3) gives

$$S_{it}^{y} = S_t^{D\prime}\delta_i + S_{it}^{x\prime}\beta + S_{it}^{u}, \tag{1.5}$$

where $S_{it}^{y} = \sum_{j=1}^{t} y_{ij}$, and $S_t^D$, $S_{it}^x$ and $S_{it}^u$ are defined analogously. As in Vogelsang and Wagner [2014], we need to add $x_{it}$ as regressors in (1.5) to deal with correlation between $u_{it}$ and $v_{it}$, which leads to

$$\begin{aligned} S_{it}^{y} &= S_t^{D\prime}\delta_i + S_{it}^{x\prime}\beta + x_{it}'\gamma + \left(S_{it}^{u} - x_{it}'\gamma\right) \\ &= S_t^{D\prime}\delta_i + S_{it}^{x\prime}\beta + x_{it}'\gamma + S_{it}^{\tilde{u}}. \end{aligned} \tag{1.6}$$

We now focus on the asymptotic behavior of the pooled OLS estimators of $\delta_i$, $\beta$ and $\gamma$ from (1.6), which we label the panel IM-OLS estimators of $\delta_i$, $\beta$ and $\gamma$. Define the stacked vectors and matrices as follows:

$$S^{y} = \begin{bmatrix} S_1^{y} \\ \vdots \\ S_N^{y} \end{bmatrix}, \; S_i^{y} = \begin{bmatrix} S_{i1}^{y} \\ \vdots \\ S_{iT}^{y} \end{bmatrix}; \; \theta = \begin{bmatrix} \beta \\ \gamma \\ \delta_1 \\ \vdots \\ \delta_N \end{bmatrix}; \; S^{\tilde{u}} = \begin{bmatrix} S_1^{\tilde{u}} \\ \vdots \\ S_N^{\tilde{u}} \end{bmatrix}, \; S_i^{\tilde{u}} = \begin{bmatrix} S_{i1}^{u} - x_{i1}'\gamma \\ \vdots \\ S_{iT}^{u} - x_{iT}'\gamma \end{bmatrix};$$

$$S^{\tilde{x}} = \begin{bmatrix} S_1^{\tilde{x}} \\ \vdots \\ S_N^{\tilde{x}} \end{bmatrix}, \; S_i^{\tilde{x}} = \begin{bmatrix} S_{i1}^{x\prime} & x_{i1}' & 0_{1\times p} & \cdots & S_1^{D\prime} & \cdots & 0_{1\times p} \\ \vdots & \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ S_{iT}^{x\prime} & x_{iT}' & \underbrace{0_{1\times p}}_{1^{st}\,block} & \cdots & \underbrace{S_T^{D\prime}}_{i^{th}\,block} & \cdots & \underbrace{0_{1\times p}}_{N^{th}\,block} \end{bmatrix}.$$

With the above notation, the matrix form of (1.6) is given by

$$S^{y} = S^{\tilde{x}}\theta + S^{\tilde{u}}, \tag{1.7}$$

and the OLS estimator of (1.7) is given by

$$\tilde{\theta} = \left(S^{\tilde{x}\prime}S^{\tilde{x}}\right)^{-1}\left(S^{\tilde{x}\prime}S^{y}\right), \tag{1.8}$$

which leads to

$$\begin{aligned}\tilde{\theta} - \theta &= \left(S^{\tilde{x}\prime}S^{\tilde{x}}\right)^{-1}\left(S^{\tilde{x}\prime}S^{\tilde{u}}\right) \qquad (1.9)\\ &= \left(\sum_{i=1}^{N}\sum_{t=1}^{T} q_{it}q_{it}'\right)^{-1}\left(\sum_{i=1}^{N}\sum_{t=1}^{T} q_{it}\left(S_{it}^{u} - x_{it}'\gamma\right)\right),\end{aligned}$$

where $q_{it} = \begin{bmatrix} S_{it}^{x\prime} & x_{it}' & 0_{1\times p} & \cdots & S_{t}^{D\prime} & \cdots & 0_{1\times p}\end{bmatrix}'$ for $i = 1, 2, \ldots, N$, $t = 1, 2, \ldots, T$. The submatrix of $q_{it}$, $\begin{bmatrix} 0_{1\times p} & \cdots & S_{t}^{D\prime} & \cdots & 0_{1\times p}\end{bmatrix}'$, consists of $S_{t}^{D\prime}$ as its $i^{th}$ block and other $N-1$ zero vector blocks. Define the scaling matrix

$$A_{PIM}^{-1} = \begin{bmatrix} TI_k & & 0 \\ & I_k & \\ 0 & & I_N \otimes G_D \end{bmatrix}$$

as a $(2k + Np) \times (2k + Np)$ diagonal matrix.

The following theorem gives the asymptotic distribution of $\tilde{\delta}_i$, $\tilde{\beta}$, $\tilde{\gamma}$.

**Theorem 1.** *Assume that the data are generated by (1.2) and (1.3), that the deterministic components satisfy (1.4) for all* $i \in [1, 2, \cdots, N]$*, and that Assumptions 1 and 2 hold. Define*

*$\theta$ by stacking the vectors $\delta_i$, $\beta$ and $\Omega_{vv}^{-1}\Omega_{vu}$. Then for fixed $N$, as $T \to \infty$*

$$\begin{bmatrix} T\left(\tilde{\beta}-\beta\right) \\ \left(\tilde{\gamma}-\Omega_{vv}^{-1}\Omega_{vu}\right) \\ G_D\left(\tilde{\delta}_1-\delta_1\right) \\ \vdots \\ G_D\left(\tilde{\delta}_N-\delta_N\right) \end{bmatrix} = A_{PIM}^{-1}\left(\tilde{\theta}-\theta\right)$$

$$= \left(A_{PIM}S^{\tilde{x}\prime}S^{\tilde{x}\prime}A_{PIM}\right)^{-1}\left(A_{PIM}S^{\tilde{x}\prime}S^{\tilde{u}}\right) - \begin{pmatrix} 0_{k\times 1} \\ \Omega_{vv}^{-1}\Omega_{vu} \\ 0_{p\times 1} \\ \vdots \\ 0_{p\times 1} \end{pmatrix}$$

$$\Rightarrow \sigma_{u\cdot v}\left(\Pi'\right)^{-1}\left(\sum_{i=1}^{N}\int_0^1 g_{1,i}(s)g_{1,i}(s)'ds\right)^{-1}\left(\sum_{i=1}^{N}\int_0^1 g_{1,i}(s)w_{u,i}(s)ds\right)$$

$$= \sigma_{u\cdot v}\left(\Pi'\right)^{-1}\left(\sum_{i=1}^{N}\int_0^1 g_{1,i}(s)g_{1,i}(s)'ds\right)^{-1}\left(\sum_{i=1}^{N}\int_0^1 [G_{1,i}(1)-G_{1,i}(s)]dw_{u,i}(s)\right)$$

$$\equiv \Psi$$

*where* $\Pi = \begin{bmatrix} \Omega_{vv}^{1/2} & & 0 \\ & \Omega_{vv}^{1/2} & \\ 0 & & I_N \otimes I_p \end{bmatrix}$, $g_{1,i}(r) = \begin{bmatrix} \int_0^r w_{v,i}(s)ds \\ w_{v,i}(r) \\ 0_{p\times 1} \\ \vdots \\ \int_0^r D(s)ds \\ \vdots \\ 0_{p\times 1} \end{bmatrix}$, $G_{1,i}(r) = \int_0^r g_{1,i}(s)ds$.

Conditional on $g_{1,i}(r)$ for all $i \in \{1, 2, \ldots, N\}$, it holds that

$$\Psi \sim N\left(0, V_{PIM}\right), \tag{1.10}$$

where $V_{PIM}$ is given by

$$\begin{aligned} V_{PIM} &= \sigma_{u\cdot v}^2 \left(\Pi'\right)^{-1} \left(\sum_{i=1}^{N} \int_0^1 g_{1,i}(s) g_{1,i}(s)' ds\right)^{-1} \times \\ & \left(\sum_{i=1}^{N} \int_0^1 [G_{1,i}(1) - G_{1,i}(s)][G_{1,i}(1) - G_{1,i}(s)]' ds\right) \times \\ & \left(\sum_{i=1}^{N} \int_0^1 g_{1,i}(s) g_{1,i}(s)' ds\right)^{-1} \Pi^{-1} \end{aligned} \tag{1.11}$$

It is clear that this conditional asymptotic variance differs from the conditional asymptotic variance of the panel FM-OLS and panel DOLS estimator of $\delta$ and $\beta$. Denoting with $m_i(s) = \left[D(s)', \quad w_{v,i}(s)'\right]'$ and with $\Pi_{PFM} = \operatorname{diag}\left(I_p, \quad \Omega_{vv}^{1/2}\right)$ the latter is given by

$$V_{PFM} = \sigma_{u\cdot v}^2 \left(\Pi_{PFM}'\right)^{-1} \left(\int_0^1 \sum_{i=1}^{N} m_i(s) m_i(s)' ds\right)^{-1} \left(\Pi_{PFM}\right)^{-1}. \tag{1.12}$$

It is important to note that we can extend Theorem 1 further to obtain a sequential result. Since the parameters in $\theta$ require different scaling for the sequential limits, and our interests are mainly on $\beta$, therefore we only provide the result for $\left(\tilde{\beta} - \beta\right)$. That is, if all the assumptions in Theorem 1 hold, and first $T \to \infty$, then $N \to \infty$, we will have following asymptotic distribution

$$\sqrt{N} T \left(\tilde{\beta} - \beta\right) \Rightarrow \Phi \tag{1.13}$$

where $\Phi \sim N\left(0, V_{seq}^{\beta}\right)$, and the asymptotic variance $V_{seq}^{\beta}$ is given by

$$\begin{aligned} V_{seq}^{\beta} &= \sigma_{u\cdot v}^{2}\left[\left(\Omega_{vv}^{1/2}\right)'\right]^{-1}\left(\int_{0}^{1}A_{1}(r)dr\right)^{-1}\left(\int_{0}^{1}A_{2}(r)dr\right)\left(\int_{0}^{1}A_{1}(r)dr\right)^{-1}\left(\Omega_{vv}^{1/2}\right)^{-1} \\ &= 5.6\sigma_{u\cdot v}^{2}\Omega_{vv}^{-1} \end{aligned} \tag{1.14}$$

where $A_1(r) = \left(r^3/3\right) I_k$ and $A_2(r) = \left[\left(1 - r - r^4 + r^5\right)/12\right] I_k$. The importance of this result is that it leads to standard inference based on the large T and large N approximation. The details of the derivation are in the Appendix.

**Remark 3.** *Above is the sequential limit for the homogeneous panel. If our panel has heterogeneity in the 2nd moment structure, then the variance, $V_{seq}^{\beta}$, will take the same expression, but $\sigma_{u\cdot v}^{2} = \underset{N\to\infty}{lim}\frac{1}{N}\sum_{i=1}^{N}\sigma_{u\cdot v,i}^{2}$ and $\Omega_{vv} = \underset{N\to\infty}{lim}\frac{1}{N}\sum_{i=1}^{N}\Omega_{vv,i}$.*

### 1.3.2 Inference using Panel IM-OLS

This section provides a discussion of hypothesis testing using the panel IM-OLS estimator. The zero mean Gaussian mixture limiting distribution of the panel IM-OLS estimator given in Theorem 1 and the conditional asymptotic variance given in (1.11) offer the theoretical basis for this discussion. In particular we consider Wald tests for testing multiple linear hypotheses of the form

$$H_0 : R\theta = r$$

where $R \in \mathbb{R}^{q\times(2k+Np)}$ with full rank $q$ and $r \in \mathbb{R}^q$. Because the vector $\tilde{\theta}$ has elements that converge at different rates, we need restriction on $R$ to get formal Wald statistics. We assume that there exists a nonsingular $q \times q$ scaling matrix $A_R$ such that

$$\lim_{T\to\infty} A_R^{-1} R A_{PIM} = R^*$$

where $R^*$ has rank $q$.

In order to carry out statistical inference, we need to scale out the asymptotic variance of panel IM-OLS. Suppose that $\breve{\sigma}^2_{u\cdot v}$ is an estimator for $\sigma^2_{u\cdot v}$. Then an estimator for $V_{PIM}$ is given by

$$\begin{aligned}
\breve{V}_{PIM} &= \breve{\sigma}^2_{u\cdot v} \left( T^{-2} \sum_{t=1}^{T} \sum_{i=1}^{N} A_{PIM} q_{it} q_{it}' A_{PIM} \right)^{-1} \times \\
& \left( T^{-4} \sum_{t=1}^{T} \sum_{i=1}^{N} A_{PIM} \left[ S^q_{iT} - S^q_{i,t-1} \right] \left[ S^q_{iT} - S^q_{i,t-1} \right]' A_{PIM} \right) \times \\
& \left( T^{-2} \sum_{t=1}^{T} \sum_{i=1}^{N} A_{PIM} q_{it} q_{it}' A_{PIM} \right)^{-1},
\end{aligned}$$

where $S^q_{it} = \sum_{j=1}^{t} q_{ij}$, and $S^q_{i0} = 0$ for all $i = 1, 2, \cdots, N$.

There are several obvious candidates for $\breve{\sigma}^2_{u\cdot v}$. The first is based on the pooled OLS residuals. Let the pooled OLS residuals be $\hat{u}_{it} = y_{it} - x_{it}'\hat{\beta} - D_t'\hat{\delta}_i$, where $\hat{\delta}_i$ and $\hat{\beta}$ are the pooled OLS estimators. Using these residuals, we can define estimators for the error vector $\hat{\eta}_{it} = \begin{bmatrix} \hat{u}_{it}, & \Delta x_{it}' \end{bmatrix}'$. Then $\hat{\Gamma}_{ij} = T^{-1} \sum_{t=j+1}^{T} \hat{\eta}_{it} \hat{\eta}_{i,t-j}'$, and

$$\hat{\Omega}_i = \begin{bmatrix} \hat{\Omega}_{uu,i} & \hat{\Omega}_{uv,i} \\ \hat{\Omega}_{vu,i} & \hat{\Omega}_{vv,i} \end{bmatrix} = \hat{\Gamma}_{i0} + \sum_{j=1}^{T-1} k(\frac{j}{M}) \left( \hat{\Gamma}_{ij} + \hat{\Gamma}_{ij}' \right).$$

The estimator for $\sigma^2_{u\cdot v}$ is given by

$$\hat{\sigma}^2_{u\cdot v} = \frac{1}{N}\sum_{i=1}^{N}\left[\hat{\Omega}_{uu,i} - \hat{\Omega}_{uv,i}\left(\hat{\Omega}_{vv,i}\right)^{-1}\hat{\Omega}_{vu,i}\right].$$

The second estimation approach is to use $\Delta\tilde{S}^u_{it}$, the first differences of the pooled OLS residuals of the panel IM-OLS regression (1.6), to directly estimate $\sigma^2_{u\cdot v}$:

$$\tilde{\sigma}^2_{u\cdot v} = \frac{1}{N}\sum_{i=1}^{N}\left[T^{-1}\sum_{j=2}^{T}\sum_{h=2}^{T} k\left(\frac{|j-h|}{M}\right)\Delta\tilde{S}^u_{ij}\Delta\tilde{S}^u_{ih}\right].$$

Extending a result in Vogelsang and Wagner [2014], it can be shown that $\tilde{\sigma}^2_{u\cdot v}$ is not a consistent estimator under traditional assumptions on the bandwidth and kernel functions. Under traditional bandwidth assumptions, we can show that the limit of $\tilde{\sigma}^2_{u\cdot v}$ is larger than $\sigma^2_{u\cdot v}$ which leads to asymptotically conservative results when we build test statistics by $\tilde{\sigma}^2_{u\cdot v}$ and use critical values from the standard normal or a chi-square distribution.

The third estimation approach is based on OLS residuals from a further augmented regression. As discussed in Vogelsang and Wagner [2014], an estimator of $\sigma^2_{u\cdot v}$ based on these residuals defined below has a fixed-b limit that is proportional to $\sigma^2_{u\cdot v}$, independent of $\tilde{\theta}$, and does not depend upon additional nuisance parameters, whereas the estimators $\hat{\sigma}^2_{u\cdot v}$ and $\tilde{\sigma}^2_{u\cdot v}$ both fail those requirements. Define this estimator as:

$$\tilde{\sigma}^{2*}_{u\cdot v} = \frac{1}{N}\sum_{i=1}^{N}\left[T^{-1}\sum_{j=2}^{T}\sum_{h=2}^{T} k\left(\frac{|j-h|}{M}\right)\Delta\tilde{S}^{u*}_{ij}\Delta\tilde{S}^{u*}_{ih}\right]$$

where $\Delta\tilde{S}^{u*}_{it} = \tilde{S}^{u*}_{it} - \tilde{S}^{u*}_{i,t-1}$ is the difference of the residuals $\tilde{S}^{u*}_{it}$, obtained by running the further augmented IM-OLS regression *individual by individual*, $\tilde{S}^{u*}_{it} = S^y_{it} - S^{D\prime}_t\tilde{\delta}_i - S^{x\prime}_{it}\tilde{\beta}_i -$

$x_{it}'\tilde{\gamma}_i - z_{it}'\tilde{\lambda}_i$. The augmented regressors, $z_{it}$, are given by $z_{it} = t \sum_{j=1}^{T} q_{ij}^x - \sum_{j=1}^{t-1} \sum_{s=1}^{j} q_{is}^x$, where $q_{it}^x = \left[ S_t^{D\prime}, \;\; S_{it}^{x\prime}, \;\; x_{it}' \right]'$ for all $i = 1, 2, \ldots, N$, $t = 1, 2, \ldots, T$.

Using an estimator of $\sigma_{u \cdot v}^2$, we can define the $t$ and Wald statistic as

$$\breve{t} = \frac{\left( R\tilde{\theta} - r \right)}{\sqrt{RA_{PIM}\breve{V}_{PIM}A_{PIM}R'}}$$

$$\breve{W} = \left[ \left( R\tilde{\theta} - r \right) \right]' \left[ RA_{PIM}\breve{V}_{PIM}A_{PIM}R' \right]^{-1} \left[ \left( R\tilde{\theta} - r \right) \right]$$

where $\breve{V}_{PIM}$ could be $\hat{V}_{PIM}$ using $\hat{\sigma}_{u \cdot v}^2$, which defines $\hat{t}$ and $\hat{W}$, or $\tilde{V}_{PIM}$ using $\tilde{\sigma}_{u \cdot v}^2$, which defines $\tilde{t}$ and $\tilde{W}$, or $\tilde{V}_{PIM}^*$ using $\tilde{\sigma}_{u \cdot v}^{2*}$, which defines $\tilde{t}^*$ and $\tilde{W}^*$. The asymptotic null distribution of these test statistics are given in Theorem 2. Standard asymptotic results based on traditional bandwidth and kernel assumptions are given for $\hat{t}$, $\hat{W}$, $\tilde{t}$ and $\tilde{W}$, whereas a fixed-b result is given for $\tilde{t}^*$ and $\tilde{W}^*$.

**Theorem 2.** *Assume that the data are generated by (1.2) and (1.3), that the deterministic components satisfy (1.4), and that Assumptions 1 and 2 hold. Suppose that the bandwidth, $M$, and kernel function, $k(\cdot)$, satisfy conditions such that $\hat{\sigma}_{u \cdot v}^2$ is consistent. Then for fixed $N$, as $T \to \infty$*

- $\hat{W} \Rightarrow \chi_q^2$, *where $\chi_q^2$ is a chi-square random variable with $q$ degrees of freedom. When $q = 1$, $\hat{t} \Rightarrow Z$, where $Z$ is a standard normal distribution.*

- *Under the above assumptions, $\tilde{\sigma}_{u \cdot v}^2 \Rightarrow \sigma_{u \cdot v}^2 \left( 1 + d_\gamma' d_\gamma \right)$, with $d_\gamma$ denoting the second $k \times 1$ block of $\left( \sum_{i=1}^{N} \int_0^1 g_{1,i}(s) g_{1,i}(s)' ds \right)^{-1} \left( \sum_{i=1}^{N} \int_0^1 [G_{1,i}(1) - G_{1,i}(s)] dw_{u,i}(s) \right)$. Therefore, it follows that $\tilde{W} \Rightarrow \frac{\chi_q^2}{1 + d_\gamma' d_\gamma}$, where $\chi_q^2$ is a chi-square random variable with $q$ degrees of* *freedom that is correlated with* $d_\gamma$*. When* $q=1$, $\tilde{t} \Rightarrow \frac{Z}{\sqrt{1+d_\gamma' d_\gamma}}$*, where* $Z$ *is distributed standard normal and is correlated with* $d_\gamma$*.*

- *If* $M = bT$*, where* $b \in (0,1]$ *is held fixed as* $T \to \infty$*, then*

$$\tilde{W}^* \Rightarrow \frac{\chi_q^2}{\frac{1}{N}\sum_{i=1}^{N} Q_i^*(b)}$$

*where* $Q_i^*(b)$ *is exactly same form as* $Q_b\left(\tilde{P}^*, \tilde{P}^*\right)$ *in Vogelsang and Wagner [2014],*[1] *and* $\chi_q^2$ *is a chi-square random variable with* $q$ *degrees freedom independent of* $\frac{1}{N}\sum_{i=1}^{N} Q_i^*(b)$*. When* $q=1$*,*

$$\tilde{t}^* \Rightarrow \frac{Z}{\sqrt{\frac{1}{N}\sum_{i=1}^{N} Q_i^*(b)}}$$

*where* $Z$ *is standard normal distribution independent of* $\frac{1}{N}\sum_{i=1}^{N} Q_i^*(b)$*.*

- *Due to the independence between the numerator and denominator of the limits of* $\tilde{W}^*$ *and* $\tilde{t}^*$*, we can further obtain a sequential limit result, where* $T$ *grows large, followed sequentially by the limit as* $N$ *grows large. Define*

$$\tilde{W}^*_{\mu_Q} = \mu_Q \cdot \tilde{W}^*,$$

[1] See Vogelsang and Wagner (2014) page 744 for $Q_b(P_1, P_2)$, and formula (30) for $\tilde{P}^*(r)$.

*where* $\mu_Q = E\left[Q_i^*(b)\right]$. *Then as* $T \to \infty$,

$$\tilde{W}^*_{\mu_Q} \Rightarrow \frac{\chi^2_q}{\left[\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\mu_Q}Q_i^*(b)\right]} \xrightarrow{P} \chi^2_q \quad as \quad N \to \infty.$$

*When* $q = 1$,

$$\begin{aligned} \tilde{t}^*_{\mu_Q} &= \sqrt{\mu_Q}\cdot\tilde{t}^* \\ &\Rightarrow \frac{Z}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\mu_Q}Q_i^*(b)}} \\ &\xrightarrow{P} Z \quad as \quad N \to \infty. \end{aligned}$$

When appealing to consistency of $\hat{\sigma}^2_{u\cdot v}$, inference using $\hat{W}$ is standard. In contrast $\tilde{\sigma}^2_{u\cdot v}$ is inconsistent under traditional bandwidth assumptions. Since $d_\gamma^{\prime} d_\gamma > 0$, the critical values of $\tilde{W}$ are smaller than those of the $\chi^2_q$ distribution. Thus, using $\chi^2_q$ critical values for $\tilde{W}$ leads to a conservative test under the traditional bandwidth assumptions. The fixed-b limiting distribution of $\tilde{W}^*$ is complicated due to the presence of $\frac{1}{N}\sum_{i=1}^{N} Q_i^*(b)$, which depends on $w_{v,i}(r)$ for $i = 1, 2, \cdots, N$. Therefore, critical values should be simulated taking into account the cross-sectional sample size, the specifications of deterministic components, the number of integrated regressors, the kernel function and bandwidth choice. For sake of brevity, in Table 1.13 and Table 1.14, we only tabulate critical values for the t-statistic for the parameter associated with $x_{it}$ in models with an intercept and 2 integrated regressors for the Bartlett and QS kernels and a grid of bandwidths indexed by b for $N = 25$. However, when both $T$

and $N$ large, inference using $\tilde{W}^*_{\mu_Q}$, with $\hat{\mu}_Q$ as an estimator of $\mu_Q$, requires merely $\chi^2_q$ critical values rather than simulated critical values, which is quite convenient. The $t$ statistics have similar results.

## 1.4 Finite sample bias and root mean squared error

In this section, we compare the performance of the pooled OLS, panel FM-OLS, panel DOLS, and panel IM-OLS estimators as measured by bias and root mean squared error (RMSE) within a small simulation study. We provide results that the individual dummy variable is not included. The data generating process is given by

$$\begin{aligned} y_{it} &= \mu + x1_{it}\beta_1 + x2_{it}\beta_2 + u_{it} \\ x1_{it} &= x1_{i,t-1} + v1_{it} \\ x2_{it} &= x2_{i,t-1} + v2_{it} \end{aligned}$$

where, for $\forall\, i \in [1, 2, \cdots, N]$, $x1_{i0} = 0$, $x2_{i0} = 0$, and $u_{it} = \rho_1 u_{i,t-1} + \epsilon_{it} + \rho_2 (e1_{it} + e2_{it})$, $u_{i0} = 0$, and $v1_{it} = e1_{it} + 0.5e1_{i,t-1}$, $v2_{it} = e2_{it} + 0.5e2_{i,t-1}$, where $\epsilon_{it}$, $e1_{it}$ and $e2_{it}$ are *i.i.d.* standard normal random variables independent of each other. The parameter values chosen are $\mu = 3$, $\beta_1 = \beta_2 = 1$. Note that the estimators of $\beta_1$ and $\beta_2$ are exactly invariant to the value of $\mu$, so the value of $\mu$ has no effect on our results. In addition, we use $\rho_1$ and $\rho_2$ from the set $\{0, 0.3, 0.6, 0.9\}$. The parameter $\rho_1$ controls serial correlation in the regression error, and $\rho_2$ determines whether the regressors are endogenous or not. The kernel function chosen for panel FM-OLS are the Bartlett and the Quadratic Spectral kernels, and the bandwidths are given by $M = bT$ with $b \in \{0.06, 0.1, 0.3, 0.5, 0.7, 0.9\}$. We also use the data dependent bandwidth from Andrews [1991]. The panel DOLS estimator is implemented using the information criterion based lead and lag length choice as developed in Kejriwal and Perron [2008], where we use the more flexible version discussed in Choi and Kurozumi [2012] in which the numbers of leads and lags included are not exactly same. The sample sizes are $N \in \{5, 10, 25\}$, $T \in \{50, 100\}$ and the number of replications is 5000.

In Tables 1.1-1.6 we display the results for $N = 5, 10, 25$, $T = 50, 100$ using the Bartlett kernel only. General patterns are similar for the QS kernel. In each of those tables, Panel A reports bias and Panel B reports RMSE.

### 1.4.1 Sample size $N = 5$

Table 1.1 shows the results for $N = 5$ with $T = 50$ case. When $\rho_2 = 0$ (no endogeneity), none of the estimators show much bias for any value of $\rho_1$. When the bandwidth is relatively small, panel FM-OLS has a little bit larger RMSE than pooled OLS. But, as the bandwidth increases, the RMSE of panel FM-OLS tends to first increase and then decreases, indicating a hump-shape in the RMSE, and the turning point is around $b = 0.5$, i.e. $M = 0.5T$. Pooled OLS and panel FM-OLS have smaller RMSE than panel IM-OLS and this holds regardless of bandwidth for panel FM-OLS. Panel IM-OLS has the largest RMSE in any cases. The RMSE of panel DOLS is a little bit smaller than that of panel IM-OLS, but it is still larger than that of pooled OLS and panel FM-OLS.

When $\rho_2 \neq 0$ (there is endogeneity), the estimators show different patterns. For a given value of $\rho_1$, as $\rho_2$ increases, the bias of pooled OLS increases. Panel FM-OLS shows the same pattern, but its bias is smaller than that of pooled OLS, which is expected from the theory. In addition, the bias of panel FM-OLS also depends on the bandwidth and value of $\rho_1$, as $\rho_1$ is relatively small, the bias of panel FM-OLS increases as bandwidth increases, however, as $\rho_1$ is far away from zero, the bias of panel FM-OLS is seen to initially fall as the bandwidth increases and then tends to increase as the bandwidth become large. But no matter how large the bandwidth is, the bias of panel FM-OLS does not exceed that of pooled OLS. On the contrary, the biases of panel IM-OLS and panel DOLS are much less sensitive to $\rho_2$, especially when $\rho_1$ is relatively small, and are always smaller than those of pooled OLS and panel FM-OLS. The bias of panel DOLS is similar to the bias of panel IM-OLS when $\rho_1$ is small whereas for larger values of $\rho_1$, the bias of panel DOLS tends to be larger than that of panel IM-OLS. The overall picture in this case is that panel IM-OLS has smaller bias than panel DOLS which in turn has lower bias than both pooled OLS and panel FM-OLS. The magnitude of the bias of both panel IM-OLS and panel DOLS are less sensitive to the values of $\rho_2$ than for pooled OLS and panel FM-OLS.

Considering the RMSE when there is endogeneity, we see that for given value of $\rho_1$, as $\rho_2$ increases, the RMSE of pooled OLS increases. Panel FM-OLS shows the same pattern, but its RMSE is smaller than that of pooled OLS, especially when $\rho_1$ is relatively small. Focusing on the bandwidth we see that the RMSE of panel FM-OLS has the same pattern as its bias, if $\rho_1$ is small, the RMSE of panel FM-OLS increases as bandwidth increases, whereas if $\rho_1$ is relatively large, the RMSE of panel FM-OLS is seen to initially fall as bandwidth increases and then tends to increase as the bandwidth becomes large. For RMSE of panel IM-OLS, it is larger than that of pooled OLS when both $\rho_1$ and $\rho_2$ are small, otherwise, it is smaller than that of pooled OLS. The RMSE of panel DOLS also has similar pattern but it is smaller than that of panel IM-OLS. The RMSE of panel IM-OLS does not vary with $\rho_2$ when $\rho_1$ is small. The comparison of RMSE for panel IM-OLS and panel FM-OLS depend on value of $\rho_1$, $\rho_2$ and $b$. When both $\rho_1$ and $\rho_2$ are small, the RMSE of panel IM-OLS is larger than that of panel FM-OLS, no matter what bandwidth used. However, when both $\rho_1$ and $\rho_2$ are large, the RMSE of panel IM-OLS could be smaller than that of panel FM-OLS with very large bandwidth.

Also, in Table 1.1, we can see that when there is endogeneity but no serial correlation, then panel FM-OLS using the data dependent bandwidth performs better than all other estimators with very small bias and smallest RMSE. And this is true for all different combinations of $N$ and $T$, which we can see from Table 1.2 to Table 1.6.

When we increase $T$ to 100, all the estimators tend to have smaller bias than the $T = 50$ case, which is expected because the estimators are consistent. Almost all of the results are similar to the $T = 50$ case. One slight difference is that when there is endogeneity and both $\rho_1$ and $\rho_2$ are large, the bias of panel IM-OLS is a little bit larger than that panel DOLS, even though they are still less biased than pooled OLS. The other difference when we increase $T$ from 50 to 100, is that when both $\rho_1$ and $\rho_2$ are very large, the RMSEs of panel IM-OLS and panel DOLS are very similar and both are smaller than that of panel FM-OLS, no matter what bandwidth used, and in turn smaller than RMSE of pooled OLS.

### 1.4.2 Sample size $N = 10$

The results of bias in $N = 10$ case are similar to the $N = 5$ case. From Panel B of Table 1.3, most of the results for RMSEs are similar as $N = 5$ case except that when both $\rho_1$ and $\rho_2$ are large, the RMSEs of panel IM-OLS and panel DOLS are very similar and both are smaller than that of panel FM-OLS for both $T = 50$ and $T = 100$. Also, when $\rho_1$ and $\rho_2$ are very large, the RMSEs of panel IM-OLS is slightly smaller than that of panel DOLS for $T = 50$ and this relation is reversed when $T = 100$.

### 1.4.3 Sample size $N = 25$

When we increase the cross section sample size to $N = 25$, the bias results are similar, but a different pattern emerges in the RMSEs. From Panel B of Tables 1.5 and 1.6, when there is endogeneity, in both cases $T = 50$ and $T = 100$, the RMSE of pooled OLS is the largest in any cases. This implies that when there is endogeneity, pooled OLS will have the largest bias and largest RMSE. Also, we can see that when both $\rho_1$ and $\rho_2$ are large, the RMSEs of panel IM-OLS and panel DOLS are very similar and both are smaller than that of panel FM-OLS with any bandwidth for both $T = 50$ and $T = 100$. Similar as the $N = 10$ case, when both $\rho_1$ and $\rho_2$ are very large, the RMSEs of panel IM-OLS is slightly smaller than that of panel DOLS when $T = 50$, and the RMSEs of panel IM-OLS is slightly larger than that of panel DOLS when $T = 100$.

### 1.4.4 Summary of finite sample bias and RMSE

The simulation shows that, when there is no endogeneity ($\rho_2 = 0$), pooled OLS dominates other estimators with no bias and smallest variance. When there is no serial correlation ($\rho_1 = 0$), panel FM-OLS with the data dependent bandwidth performs better than other estimators. When both serial correlation and endogeneity exist ($\rho_1 \neq 0, \rho_2 \neq 0$), the relative performance of the estimators depends on the values of $N$, $T$, $\rho_1$ and $\rho_2$. Panel IM-OLS is more effective in reducing bias than the other estimators and both bias and RMSE of panel IM-OLS are less sensitive to the parameters $\rho_1$ and $\rho_2$ than are the bias and RMSE of panel FM-OLS. For $N$ small ($N = 5, 10$) and $T$ small ($T = 50$), panel IM-OLS has the smallest bias but with larger RMSE as a cost, except that when both $\rho_1$ and $\rho_2$ are large where panel IM-OLS has the smallest RMSE. For $N$ small and $T$ relatively large ($T = 100$), panel IM-OLS and panel DOLS are similar, and dominate pooled OLS and panel FM-OLS. When $N$ is relatively large ($N = 25$), pooled OLS has the largest bias and largest RMSE in all cases, and if $T$ is small and $\rho_1$, $\rho_2$ are relatively large, then panel IM-OLS is better than the other estimators in reducing bias and has smallest RMSE. When $N$ and $T$ are large, and $\rho_1$, $\rho_2$ are large, then panel DOLS is a little bit better than panel IM-OLS, which in turn is better than pooled OLS and panel FM-OLS.

## 1.5 Finite sample performance of test statistics

In this section we provide some finite sample results about the tests' performance using the simulation design from Section 1.4. Here, we only report results for cases where $\rho_1 = \rho_2$. The results include t-statistics for testing the null hypothesis $H_0 : \beta_1 = 1$ and Wald statistics for testing the joint null hypothesis $H_0 : \beta_1 = 1, \beta_2 = 1$. The pooled OLS statistics serve as a benchmark. The panel FM-OLS statistics were implemented using $\hat{\sigma}^2_{u^+}$. The panel IM-OLS statistics were implemented in three ways. The first uses $\hat{\sigma}^2_{u\cdot v}$ and is labeled panel IM(O), the second uses $\tilde{\sigma}^2_{u\cdot v}$ and is labeled panel IM(D) and the third uses $\tilde{\sigma}^{2*}_{u\cdot v}$ and is labeled panel IM(Fb). We report results for both the Bartlett and Quadratic Spectral kernels. As for the choice of bandwidth for panel FM-OLS and panel IM-OLS statistics, we follow Vogelsang and Wagner [2014]. One bandwidth choice is the data dependent bandwidth rule of Andrews [1991]. The other choice is the fixed-b bandwidth, that is $b = M/T$, where $M = 1, 2, \cdots, T$. Rejections for the pooled OLS, panel FM-OLS, panel DOLS, panel IM(O) and panel IM(D) are carried out using $N(0,1)$ critical values for all values of $M$. From Theorem 2, the panel IM(D) test statistic is asymptotically conservative under traditional asymptotic theory. In contrast, rejections for panel IM(Fb) are carried out using fixed-b asymptotic critical values. The empirical rejection probabilities were computed using 5000 replications, and the nominal level is 0.05 in all cases.

Tables 1.7 to 1.9 and Tables 1.10 to 1.12 report empirical null rejection probabilities using data dependent bandwidth choices for Bartlett and QS kernel. Tables 1.7 to 1.9 show results for the t-tests and Tables 1.10 to 1.12 contain results for the Wald tests. In each table Panel A corresponds to $T = 50$ and Panel B to $T = 100$. We only briefly summarize some main findings in the tables. When $\rho_1 = \rho_2 = 0$ (no serial correlation and no endogeneity), we can see that pooled OLS tests have rejection probabilities close to 0.05, but there are huge over-rejections as the value of $\rho_1$ and $\rho_2$ increase. For $\rho_1 = \rho_2 = 0$, when using the QS kernel, panel IM(Fb) tests tend to have rejection probabilities less than 0.05, whereas other tests show some over-rejections. For $\rho_1 = \rho_2 = 0$, when using the Bartlett kernel, all the tests show some over-rejections, but the over-rejection problem is less severe when $T = 100$ than $T = 50$. Note that both panel IM(O) and panel IM(D) show some over-rejections, but those are less severe than panel FM-OLS, especially when there is strong serial correlation and strong endogeneity. Generally, panel IM(D) tests have rejection probabilities that are smaller than those of panel IM(O), which is what we expected because the panel IM(D) test is conservative under standard asymptotic theory. In addition, increasing the values of $\rho_1$

and $\rho_2$ leads to over-rejection problems for all the tests. The problem with panel IM-OLS is that the data-dependent bandwidth is too small to give less size distortions. In contrast to the pure time series case, there is no test that dominates the others in that scenario.

In order to see the impact of bandwidth and kernel choices on over-rejection problem, we plot in Figures 1.1-1.3 null rejection probabilities of the t-tests as a function of $b \in (0, 1]$. The first three figures give the results for $N \in \{5, 10, 25\}$, $T = 100$ using the Bartlett kernel and $\rho_1 = \rho_2 = 0.3$. In Figure 1.1, with cross-section sample size $N = 5$, we can see that with small bandwidths, all tests have some over-rejection problems. Panel IM(D) is less severe than the other tests because it is conservative. As the bandwidth increases, all rejection probabilities increase substantially except for panel IM(Fb). The panel IM(Fb) rejection probabilities are close to 10% for all values of b, which indicates that the fixed-b approximation performs relatively well for panel IM(Fb). In Figures 1.2 and 1.3, the cross-section sample size increases to 10 and 25 and the pattern of rejection probabilities are similar as Figure 1.1. However, when the bandwidth is small, like $b = 0.08$, panel FM(Fb) has the least rejection probabilities, around 8% and 7.5%, respectively. In addition, as $N$ increases, panel IM(Fb) rejection probabilities are close to 10% and 12% when large bandwidth used.

As the values of $\rho_1, \rho_2$ increase to 0.9, there exists strong serial correlation and endogeneity. We can see from Figures 1.4-1.6 that all the tests have serious over-rejection problems regardless of bandwidth. Interestingly, for small $N$ ($N = 5, 10$), panel FM(Fb) has less of an over-rejection problem than panel IM(Fb) although both tests are severely size distorted. As $N$ increases to 25, the rejection probabilities of panel IM(Fb) tend to smaller than that of panel FM(Fb). In Vogelsang and Wagner [2014], it was pointed out that the over-rejection problems of IM(Fb) becomes less problematic as $T$ increases. We find similar patterns in our simulations. In Figures 1.7 to 1.9, we show the results with $T = 100$, and it is clear that the panel IM(Fb) over-rejections are reduced although they are still large. We believe that, similar to the pure time series case, if we further increase $T$ to 500 or 1000, the rejections for panel IM(Fb) with non-small bandwidths will be substantially reduced to reasonable size whereas the other statistics will continue to have over-rejection problems.

Figures 1.10-1.12 give some results for the QS kernel. For brevity we only report results for $N = 5$, $T = 50, 100$, $\rho_1 = \rho_2 \in \{0.3, 0.9\}$. Compared with results using Bartlett kernel, panel IM(Fb) has less over-rejection problems using QS kernel. In Figures 1.10, when serial correlation and endogeneity are not that strong, panel IM(Fb) tends to have rejections close to 10%, whereas all other tests have over-rejection problems. In Figures 1.11 and 1.12, when there is strong serial correlation and endogeneity, all the tests have some over-rejection problems, however, the QS kernel leads to less size distorted results than the Bartlett kernel.

The overall picture is that the panel IM(Fb) test is the most robust statistic in terms of controlling over-rejection problems although for given sample sizes, $N$, $T$, increasing the values of $\rho_1$, $\rho_2$ causes over-rejections to emerge. Large sample sizes of both $N$ and $T$ in conjunction with large bandwidths and the QS kernel are desirable when serial correlation and endogeneity are strong.

Now, we turn to the analysis of the power properties of the tests. For the sake of brevity we only display results for the case $\rho_1 = \rho_2 = 0.6$ for the Wald test for $N = 5$, $T = 50$ and using the QS kernel. Patterns are similar for other values of $\rho_1$, $\rho_2$ for $t$ tests for other combinations of $N, T$ with the Bartlett kernel. Starting from the null values of $\beta_1$ and $\beta_2$ equal to 1, we consider under the alternative $\beta_1 = \beta_2 = \beta \in (1, 1.4]$, using (including the null value) a total of 21 values on a grid with mesh 0.02. We focus on size-corrected power because of the potential over-rejection problems under the null hypothesis. This allows us to see power differences across tests while holding null rejection probabilities constant at

0.05. This is useful for the theoretical power comparisons, but such size-corrections are not feasible in practice.

In Figure 1.13 we display size adjusted power of the panel FM(Fb) and panel IM-OLS Wald tests using the QS kernel with $b = 0.3$. For all other values of $b$, the patterns are very similar. From Figure 1.13, we can see that panel IM(Fb) has the least power across the four tests. The use of $\tilde{\sigma}^{2*}_{u \cdot v}$ to obtain asymptotically fixed-b inference and less finite sample size distortions comes at the price of a small reduction in power.

Figure 1.14 shows the effect of the bandwidth on size adjusted power of the panel IM(Fb) test by plotting power curves for eight values of $b = 0.02, 0.06, 0.1, 0.3, 0.5, 0.7, 0.9, 1.0$. We can see that panel IM(Fb) power depends on the bandwidth and tends to decrease as bandwidth increases, but power is not that sensitive to the bandwidth. In addition, when $b \geq 0.5$, the power of panel IM(Fb) is almost constant. In Figure 1.15, we display power using the Bartlett kernel, and it is clear that all tests almost have similar power, and the power is not sensitive to $b$.

Figure 1.16 gives power comparisons across the various tests: pooled OLS, panel FM-OLS, panel DOLS, panel IM-OLS. In Figure 1.16, panel IM(Fb) test is shown for $b = 0.06, 0.1, 1.0$, and using the Andrews data dependent bandwidth. The panel FM-OLS test is implemented with the Andrews data dependent bandwidth. We note that the pooled OLS and panel FM-OLS tests have the largest size-adjusted power, with the power of panel DOLS test being slightly smaller and panel IM-OLS tests have the smallest power. But the power difference between panel IM-OLS and all other tests are relatively small.

Finally, Figures 1.17-1.19 provide size adjusted power comparisons similar to Figures 1.13, 1.14 and 1.16 but with $N = 10$. The main feature is that size adjusted power increases with $N$. In addition, as $N$ increases, the power of panel IM(Fb) becomes less sensitive to the bandwidth. With a larger $N$, the power rankings are the same as before, but the difference of the power between panel IM-OLS and panel FM-OLS is smaller.

## 1.6 Summary and conclusions

This paper considers the extension of the integrated modified ordinary least squares (IM-OLS) method of Vogelsang and Wagner [2014] for estimation and inference about a cointegrating vector in homogeneous cointegrated panels. We label the estimator panel IM-OLS. It is a tuning parameter free estimator that is based on a partial sum transformed regression augmented by the original integrated regressors themselves. The advantage is that it leads to a zero mean mixed Gaussian limiting distribution without requiring the choice of tuning parameters (like bandwidth, kernel, numbers of leads and lags). For inference based on panel IM-OLS estimates, a long run variance still needs to be scaled out. Using a consistent estimator of the corresponding long run variance leads to tests having standard asymptotic distributions. Fixed-b inference is another way to obtain pivotal test statistics. Critical values of fixed-b $t$ and Wald tests need to be simulated taking into account the specification of deterministic components, the number of integrated regressors, the kernel function and the bandwidth choice.

We provide a finite sample simulation study in which the performance of the panel IM-OLS estimator and test statistics are compared with pooled OLS, panel FM-OLS and panel DOLS. Typically, panel IM-OLS shows good performance in terms of bias and RMSE especially in the following two scenarios: (i) the panel has large sample size; (ii) small sample size panel with strong serial correlation and endogeneity. The size and power analysis of the tests show that the fixed-b test statistics are more robust, in terms of having lower size distortions than all other test statistics, especially for larger sample sizes. This robustness comes at the cost of minor power losses provided serial correlation and endogeneity is not that strong. When there is strong serial correlation and endogeneity, all tests have severe over-rejection problems, and we prefer panel IM-OLS test with QS kernel and large bandwidth in this case.

Further research will study panel IM-OLS estimator for panels that have identical dependent unit in cross section, panels that have non-identical dependent unit in cross section, for higher order cointegrating regressions and for nonlinear cointegration relationships.

APPENDIX

# Tables and Figures

Table 1.1: Finite sample bias and RMSE of the various estimator of $\beta_1$, $N = 5$, $T = 50$, Bartlett kernel

| $\rho_1$ | $\rho_2$ | P-OLS | P-IM | P-DOLS | Panel FM-OLS | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | b=0.06 | 0.1 | 0.3 | 0.5 | 0.7 | 0.9 | AND |
| Panel A: Bias | | | | | | | | | | | |
| 0 | 0 | -0.0002 | -0.0001 | -0.0004 | -0.0003 | -0.0003 | -0.0003 | -0.0003 | -0.0003 | -0.0003 | -0.0003 |
| | 0.3 | 0.0057 | -0.0004 | -0.0006 | -0.0004 | 0.0003 | 0.0021 | 0.0030 | 0.0036 | 0.0040 | 0.0001 |
| | 0.6 | 0.0116 | -0.0008 | -0.0003 | -0.0006 | 0.0009 | 0.0044 | 0.0063 | 0.0075 | 0.0082 | 0.0005 |
| | 0.9 | 0.0175 | -0.0011 | -0.0002 | -0.0007 | 0.0015 | 0.0068 | 0.0096 | 0.0113 | 0.0124 | 0.0010 |
| 0.3 | 0 | -0.0003 | -0.0002 | -0.0004 | -0.0004 | -0.0004 | -0.0004 | -0.0004 | -0.0004 | -0.0004 | -0.0004 |
| | 0.3 | 0.0100 | -0.0001 | -0.0001 | 0.0013 | 0.0017 | 0.0039 | 0.0055 | 0.0064 | 0.0071 | 0.0016 |
| | 0.6 | 0.0202 | 0.0001 | 0.0001 | 0.0029 | 0.0038 | 0.0083 | 0.0113 | 0.0132 | 0.0145 | 0.0035 |
| | 0.9 | 0.0304 | 0.0002 | 0.0002 | 0.0045 | 0.0059 | 0.0126 | 0.0171 | 0.0200 | 0.0219 | 0.0055 |
| 0.6 | 0 | -0.0004 | -0.0003 | -0.0008 | -0.0005 | -0.0005 | -0.0006 | -0.0005 | -0.0005 | -0.0005 | -0.0005 |
| | 0.3 | 0.0200 | 0.0024 | 0.0034 | 0.0083 | 0.0075 | 0.0091 | 0.0116 | 0.0134 | 0.0146 | 0.0076 |
| | 0.6 | 0.0404 | 0.0051 | 0.0068 | 0.0172 | 0.0155 | 0.0188 | 0.0238 | 0.0274 | 0.0297 | 0.0158 |
| | 0.9 | 0.0608 | 0.0078 | 0.0090 | 0.0260 | 0.0235 | 0.0285 | 0.0360 | 0.0413 | 0.0448 | 0.0240 |
| 0.9 | 0 | -0.0009 | -0.0007 | -0.0007 | -0.0011 | -0.0011 | -0.0012 | -0.0011 | -0.0010 | -0.0010 | -0.0011 |
| | 0.3 | 0.0715 | 0.0404 | 0.0521 | 0.0598 | 0.0568 | 0.0511 | 0.0529 | 0.0561 | 0.0587 | 0.0575 |
| | 0.6 | 0.1438 | 0.0815 | 0.1031 | 0.1206 | 0.1147 | 0.1035 | 0.1069 | 0.1131 | 0.1183 | 0.1162 |
| | 0.9 | 0.2162 | 0.1226 | 0.1534 | 0.1815 | 0.1726 | 0.1559 | 0.1610 | 0.1702 | 0.1780 | 0.1749 |
| Panel B: RMSE | | | | | | | | | | | |
| 0 | 0 | 0.0115 | 0.0202 | 0.0147 | 0.0119 | 0.0120 | 0.0124 | 0.0124 | 0.0124 | 0.0124 | 0.0120 |
| | 0.3 | 0.0134 | 0.0202 | 0.0151 | 0.0119 | 0.0121 | 0.0128 | 0.0131 | 0.0133 | 0.0134 | 0.0121 |
| | 0.6 | 0.0180 | 0.0202 | 0.0156 | 0.0121 | 0.0124 | 0.0141 | 0.0151 | 0.0157 | 0.0160 | 0.0123 |
| | 0.9 | 0.0239 | 0.0203 | 0.0158 | 0.0123 | 0.0130 | 0.0160 | 0.0179 | 0.0190 | 0.0197 | 0.0127 |
| 0.3 | 0 | 0.0161 | 0.0286 | 0.0211 | 0.0168 | 0.0170 | 0.0175 | 0.0175 | 0.0175 | 0.0175 | 0.0169 |
| | 0.3 | 0.0198 | 0.0286 | 0.0213 | 0.0169 | 0.0172 | 0.0183 | 0.0189 | 0.0192 | 0.0194 | 0.0171 |
| | 0.6 | 0.0283 | 0.0286 | 0.0215 | 0.0173 | 0.0179 | 0.0208 | 0.0227 | 0.0238 | 0.0246 | 0.0177 |
| | 0.9 | 0.0386 | 0.0287 | 0.0217 | 0.0180 | 0.0192 | 0.0244 | 0.0279 | 0.0300 | 0.0313 | 0.0188 |
| 0.6 | 0 | 0.0270 | 0.0490 | 0.0358 | 0.0283 | 0.0287 | 0.0297 | 0.0297 | 0.0296 | 0.0295 | 0.0286 |
| | 0.3 | 0.0352 | 0.0491 | 0.0365 | 0.0298 | 0.0301 | 0.0318 | 0.0330 | 0.0337 | 0.0341 | 0.0300 |
| | 0.6 | 0.0529 | 0.0494 | 0.0378 | 0.0345 | 0.0341 | 0.0379 | 0.0416 | 0.0439 | 0.0455 | 0.0341 |
| | 0.9 | 0.0735 | 0.0499 | 0.0391 | 0.0411 | 0.0401 | 0.0464 | 0.0529 | 0.0571 | 0.0598 | 0.0402 |
| 0.9 | 0 | 0.0843 | 0.1638 | 0.1090 | 0.0887 | 0.0910 | 0.0967 | 0.0973 | 0.0967 | 0.0958 | 0.0904 |
| | 0.3 | 0.1131 | 0.1689 | 0.1231 | 0.1089 | 0.1091 | 0.1113 | 0.1131 | 0.1142 | 0.1149 | 0.1090 |
| | 0.6 | 0.1736 | 0.1839 | 0.1563 | 0.1552 | 0.1518 | 0.1475 | 0.1512 | 0.1557 | 0.1591 | 0.1526 |
| | 0.9 | 0.2432 | 0.2068 | 0.1986 | 0.2111 | 0.2041 | 0.1935 | 0.1992 | 0.2071 | 0.2134 | 0.2059 |

Table 1.2: Finite sample bias and RMSE of the various estimator of $\beta_1$, $N = 5$, $T = 100$, Bartlett kernel

| $\rho_1$ | $\rho_2$ | P-OLS | P-IM | P-DOLS | Panel FM-OLS | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | b=0.06 | 0.1 | 0.3 | 0.5 | 0.7 | 0.9 | AND |
| | | | | | Panel A: Bias | | | | | | |
| 0 | 0 | 0.0001 | 0.0001 | 0.0000 | 0.0001 | 0.0001 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 |
| | 0.3 | 0.0030 | 0.0000 | 0.0001 | 0.0002 | 0.0005 | 0.0012 | 0.0017 | 0.0020 | 0.0021 | 0.0001 |
| | 0.6 | 0.0060 | 0.0000 | 0.0002 | 0.0003 | 0.0009 | 0.0024 | 0.0033 | 0.0039 | 0.0043 | 0.0002 |
| | 0.9 | 0.0089 | -0.0001 | 0.0001 | 0.0004 | 0.0013 | 0.0036 | 0.0050 | 0.0058 | 0.0064 | 0.0003 |
| 0.3 | 0 | 0.0001 | 0.0002 | 0.0002 | 0.0001 | 0.0001 | 0.0001 | 0.0000 | 0.0000 | 0.0000 | 0.0001 |
| | 0.3 | 0.0052 | 0.0002 | 0.0003 | 0.0008 | 0.0010 | 0.0022 | 0.0029 | 0.0034 | 0.0038 | 0.0007 |
| | 0.6 | 0.0104 | 0.0002 | 0.0004 | 0.0015 | 0.0020 | 0.0043 | 0.0059 | 0.0069 | 0.0075 | 0.0014 |
| | 0.9 | 0.0156 | 0.0003 | 0.0004 | 0.0022 | 0.0030 | 0.0064 | 0.0088 | 0.0103 | 0.0113 | 0.0021 |
| 0.6 | 0 | 0.0002 | 0.0003 | 0.0002 | 0.0002 | 0.0002 | 0.0001 | 0.0001 | 0.0000 | 0.0000 | 0.0002 |
| | 0.3 | 0.0107 | 0.0011 | 0.0011 | 0.0034 | 0.0032 | 0.0047 | 0.0061 | 0.0071 | 0.0078 | 0.0035 |
| | 0.6 | 0.0212 | 0.0018 | 0.0015 | 0.0067 | 0.0062 | 0.0093 | 0.0122 | 0.0142 | 0.0155 | 0.0069 |
| | 0.9 | 0.0317 | 0.0026 | 0.0016 | 0.0099 | 0.0093 | 0.0139 | 0.0183 | 0.0212 | 0.0232 | 0.0102 |
| 0.9 | 0 | 0.0008 | 0.0012 | 0.0002 | 0.0008 | 0.0008 | 0.0006 | 0.0005 | 0.0004 | 0.0004 | 0.0008 |
| | 0.3 | 0.0431 | 0.0174 | 0.0174 | 0.0325 | 0.0294 | 0.0261 | 0.0286 | 0.0314 | 0.0335 | 0.0331 |
| | 0.6 | 0.0854 | 0.0336 | 0.0335 | 0.0642 | 0.0580 | 0.0515 | 0.0567 | 0.0623 | 0.0666 | 0.0654 |
| | 0.9 | 0.1277 | 0.0499 | 0.0486 | 0.0959 | 0.0866 | 0.0769 | 0.0849 | 0.0933 | 0.0997 | 0.0976 |
| | | | | | Panel B: RMSE | | | | | | |
| 0 | 0 | 0.0056 | 0.0102 | 0.0082 | 0.0058 | 0.0059 | 0.0061 | 0.0062 | 0.0062 | 0.0061 | 0.0058 |
| | 0.3 | 0.0067 | 0.0102 | 0.0085 | 0.0058 | 0.0059 | 0.0064 | 0.0066 | 0.0067 | 0.0067 | 0.0058 |
| | 0.6 | 0.0091 | 0.0102 | 0.0087 | 0.0059 | 0.0061 | 0.0070 | 0.0076 | 0.0079 | 0.0081 | 0.0059 |
| | 0.9 | 0.0121 | 0.0102 | 0.0089 | 0.0061 | 0.0065 | 0.0080 | 0.0090 | 0.0096 | 0.0099 | 0.0060 |
| 0.3 | 0 | 0.0080 | 0.0145 | 0.0131 | 0.0083 | 0.0083 | 0.0087 | 0.0088 | 0.0088 | 0.0087 | 0.0082 |
| | 0.3 | 0.0101 | 0.0145 | 0.0133 | 0.0083 | 0.0085 | 0.0092 | 0.0096 | 0.0098 | 0.0099 | 0.0083 |
| | 0.6 | 0.0145 | 0.0145 | 0.0135 | 0.0086 | 0.0089 | 0.0105 | 0.0115 | 0.0121 | 0.0125 | 0.0085 |
| | 0.9 | 0.0198 | 0.0145 | 0.0137 | 0.0090 | 0.0096 | 0.0123 | 0.0142 | 0.0153 | 0.0160 | 0.0089 |
| 0.6 | 0 | 0.0136 | 0.0251 | 0.0254 | 0.0142 | 0.0144 | 0.0150 | 0.0151 | 0.0151 | 0.0150 | 0.0142 |
| | 0.3 | 0.0184 | 0.0252 | 0.0258 | 0.0149 | 0.0150 | 0.0162 | 0.0170 | 0.0174 | 0.0177 | 0.0148 |
| | 0.6 | 0.0278 | 0.0252 | 0.0262 | 0.0164 | 0.0166 | 0.0193 | 0.0214 | 0.0228 | 0.0237 | 0.0165 |
| | 0.9 | 0.0387 | 0.0253 | 0.0264 | 0.0187 | 0.0188 | 0.0235 | 0.0272 | 0.0296 | 0.0311 | 0.0188 |
| 0.9 | 0 | 0.0474 | 0.0932 | 0.0855 | 0.0501 | 0.0513 | 0.0541 | 0.0545 | 0.0541 | 0.0537 | 0.0499 |
| | 0.3 | 0.0669 | 0.0950 | 0.0881 | 0.0616 | 0.0609 | 0.0619 | 0.0637 | 0.0648 | 0.0656 | 0.0618 |
| | 0.6 | 0.1046 | 0.0997 | 0.0957 | 0.0865 | 0.0822 | 0.0800 | 0.0848 | 0.0888 | 0.0917 | 0.0874 |
| | 0.9 | 0.1471 | 0.1071 | 0.1045 | 0.1164 | 0.1086 | 0.1031 | 0.1111 | 0.1182 | 0.1234 | 0.1180 |

Table 1.3: Finite sample bias and RMSE of the various estimator of $\beta_1$, $N = 10$, $T = 50$, Bartlett kernel

| $\rho_1$ | $\rho_2$ | P-OLS | P-IM | P-DOLS | Panel FM-OLS | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | b=0.06 | 0.1 | 0.3 | 0.5 | 0.7 | 0.9 | AND |
| | | | | | Panel A: Bias | | | | | | |
| 0 | 0 | -0.0001 | 0.0000 | -0.0001 | -0.0001 | -0.0001 | -0.0001 | 0.0000 | 0.0000 | 0.0000 | -0.0001 |
| | 0.3 | 0.0055 | -0.0002 | -0.0003 | -0.0003 | 0.0003 | 0.0017 | 0.0026 | 0.0033 | 0.0037 | 0.0002 |
| | 0.6 | 0.0111 | -0.0005 | 0.0000 | -0.0005 | 0.0006 | 0.0035 | 0.0053 | 0.0065 | 0.0074 | 0.0004 |
| | 0.9 | 0.0167 | -0.0008 | 0.0001 | -0.0007 | 0.0010 | 0.0053 | 0.0080 | 0.0098 | 0.0111 | 0.0006 |
| 0.3 | 0 | -0.0001 | 0.0001 | -0.0001 | -0.0001 | -0.0001 | -0.0001 | 0.0000 | -0.0001 | 0.0000 | -0.0001 |
| | 0.3 | 0.0092 | 0.0001 | 0.0002 | 0.0012 | 0.0014 | 0.0032 | 0.0046 | 0.0056 | 0.0063 | 0.0014 |
| | 0.6 | 0.0186 | 0.0002 | 0.0004 | 0.0024 | 0.0030 | 0.0064 | 0.0092 | 0.0112 | 0.0126 | 0.0028 |
| | 0.9 | 0.0279 | 0.0003 | 0.0005 | 0.0037 | 0.0045 | 0.0097 | 0.0138 | 0.0168 | 0.0189 | 0.0043 |
| 0.6 | 0 | -0.0001 | 0.0001 | -0.0003 | -0.0002 | -0.0002 | -0.0001 | -0.0001 | -0.0001 | 0.0000 | -0.0002 |
| | 0.3 | 0.0182 | 0.0024 | 0.0036 | 0.0074 | 0.0064 | 0.0074 | 0.0096 | 0.0114 | 0.0127 | 0.0066 |
| | 0.6 | 0.0365 | 0.0047 | 0.0066 | 0.0149 | 0.0130 | 0.0150 | 0.0193 | 0.0229 | 0.0254 | 0.0134 |
| | 0.9 | 0.0548 | 0.0070 | 0.0085 | 0.0224 | 0.0196 | 0.0225 | 0.0290 | 0.0343 | 0.0382 | 0.0202 |
| 0.9 | 0 | 0.0001 | 0.0005 | -0.0001 | 0.0001 | 0.0000 | 0.0001 | 0.0003 | 0.0004 | 0.0005 | 0.0001 |
| | 0.3 | 0.0672 | 0.0386 | 0.0488 | 0.0560 | 0.0530 | 0.0469 | 0.0482 | 0.0512 | 0.0541 | 0.0538 |
| | 0.6 | 0.1343 | 0.0768 | 0.0959 | 0.1119 | 0.1060 | 0.0937 | 0.0960 | 0.1020 | 0.1077 | 0.1075 |
| | 0.9 | 0.2014 | 0.1149 | 0.1419 | 0.1679 | 0.1590 | 0.1405 | 0.1439 | 0.1528 | 0.1613 | 0.1612 |
| | | | | | Panel B: RMSE | | | | | | |
| 0 | 0 | 0.0069 | 0.0118 | 0.0088 | 0.0071 | 0.0072 | 0.0074 | 0.0075 | 0.0075 | 0.0074 | 0.0072 |
| | 0.3 | 0.0091 | 0.0118 | 0.0090 | 0.0071 | 0.0072 | 0.0077 | 0.0081 | 0.0083 | 0.0084 | 0.0072 |
| | 0.6 | 0.0138 | 0.0118 | 0.0092 | 0.0072 | 0.0073 | 0.0086 | 0.0097 | 0.0104 | 0.0110 | 0.0073 |
| | 0.9 | 0.0193 | 0.0119 | 0.0094 | 0.0073 | 0.0076 | 0.0099 | 0.0119 | 0.0133 | 0.0142 | 0.0075 |
| 0.3 | 0 | 0.0098 | 0.0168 | 0.0125 | 0.0101 | 0.0102 | 0.0105 | 0.0106 | 0.0106 | 0.0105 | 0.0101 |
| | 0.3 | 0.0139 | 0.0168 | 0.0127 | 0.0102 | 0.0103 | 0.0111 | 0.0118 | 0.0122 | 0.0125 | 0.0103 |
| | 0.6 | 0.0220 | 0.0168 | 0.0128 | 0.0105 | 0.0108 | 0.0129 | 0.0148 | 0.0161 | 0.0171 | 0.0107 |
| | 0.9 | 0.0312 | 0.0168 | 0.0130 | 0.0110 | 0.0116 | 0.0154 | 0.0187 | 0.0212 | 0.0229 | 0.0114 |
| 0.6 | 0 | 0.0166 | 0.0290 | 0.0215 | 0.0172 | 0.0174 | 0.0180 | 0.0181 | 0.0181 | 0.0179 | 0.0173 |
| | 0.3 | 0.0253 | 0.0291 | 0.0221 | 0.0189 | 0.0187 | 0.0198 | 0.0210 | 0.0218 | 0.0225 | 0.0187 |
| | 0.6 | 0.0418 | 0.0294 | 0.0233 | 0.0233 | 0.0223 | 0.0244 | 0.0278 | 0.0305 | 0.0325 | 0.0225 |
| | 0.9 | 0.0600 | 0.0299 | 0.0244 | 0.0293 | 0.0273 | 0.0307 | 0.0365 | 0.0411 | 0.0444 | 0.0277 |
| 0.9 | 0 | 0.0548 | 0.0994 | 0.0688 | 0.0568 | 0.0579 | 0.0610 | 0.0616 | 0.0612 | 0.0606 | 0.0576 |
| | 0.3 | 0.0878 | 0.1068 | 0.0854 | 0.0806 | 0.0792 | 0.0776 | 0.0790 | 0.0807 | 0.0820 | 0.0795 |
| | 0.6 | 0.1477 | 0.1261 | 0.1210 | 0.1276 | 0.1227 | 0.1138 | 0.1163 | 0.1213 | 0.1259 | 0.1240 |
| | 0.9 | 0.2129 | 0.1529 | 0.1620 | 0.1805 | 0.1724 | 0.1564 | 0.1602 | 0.1684 | 0.1760 | 0.1744 |

Table 1.4: Finite sample bias and RMSE of the various estimator of $\beta_1$, $N = 10$, $T = 100$, Bartlett kernel

| $\rho_1$ | $\rho_2$ | P-OLS | P-IM | P-DOLS | Panel FM-OLS | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | b=0.06 | 0.1 | 0.3 | 0.5 | 0.7 | 0.9 | AND |
| | | | | | Panel A: Bias | | | | | | |
| 0 | 0 | 0.0000 | -0.0001 | 0.0000 | 0.0000 | 0.0000 | -0.0001 | -0.0001 | -0.0001 | -0.0001 | 0.0000 |
| | 0.3 | 0.0027 | -0.0001 | 0.0000 | 0.0000 | 0.0002 | 0.0009 | 0.0013 | 0.0016 | 0.0018 | 0.0000 |
| | 0.6 | 0.0055 | -0.0002 | 0.0001 | 0.0001 | 0.0005 | 0.0018 | 0.0026 | 0.0033 | 0.0037 | 0.0000 |
| | 0.9 | 0.0083 | -0.0003 | 0.0000 | 0.0002 | 0.0008 | 0.0027 | 0.0040 | 0.0049 | 0.0056 | 0.0000 |
| 0.3 | 0 | -0.0001 | -0.0001 | 0.0000 | -0.0001 | -0.0001 | -0.0001 | -0.0001 | -0.0001 | -0.0001 | -0.0001 |
| | 0.3 | 0.0046 | -0.0001 | 0.0001 | 0.0005 | 0.0006 | 0.0015 | 0.0022 | 0.0027 | 0.0031 | 0.0004 |
| | 0.6 | 0.0093 | -0.0001 | 0.0002 | 0.0010 | 0.0013 | 0.0031 | 0.0045 | 0.0055 | 0.0063 | 0.0009 |
| | 0.9 | 0.0140 | 0.0000 | 0.0002 | 0.0015 | 0.0020 | 0.0048 | 0.0068 | 0.0084 | 0.0094 | 0.0014 |
| 0.6 | 0 | -0.0001 | -0.0002 | 0.0000 | -0.0001 | -0.0001 | -0.0001 | -0.0002 | -0.0002 | -0.0002 | -0.0001 |
| | 0.3 | 0.0092 | 0.0005 | 0.0009 | 0.0026 | 0.0023 | 0.0034 | 0.0046 | 0.0056 | 0.0063 | 0.0027 |
| | 0.6 | 0.0186 | 0.0011 | 0.0012 | 0.0052 | 0.0047 | 0.0069 | 0.0094 | 0.0113 | 0.0127 | 0.0055 |
| | 0.9 | 0.0280 | 0.0017 | 0.0013 | 0.0079 | 0.0071 | 0.0104 | 0.0142 | 0.0170 | 0.0191 | 0.0082 |
| 0.9 | 0 | -0.0001 | -0.0005 | -0.0005 | -0.0001 | -0.0002 | -0.0003 | -0.0003 | -0.0004 | -0.0004 | -0.0001 |
| | 0.3 | 0.0381 | 0.0141 | 0.0146 | 0.0281 | 0.0251 | 0.0213 | 0.0233 | 0.0260 | 0.0282 | 0.0286 |
| | 0.6 | 0.0763 | 0.0287 | 0.0287 | 0.0563 | 0.0504 | 0.0429 | 0.0470 | 0.0523 | 0.0568 | 0.0574 |
| | 0.9 | 0.1145 | 0.0433 | 0.0417 | 0.0846 | 0.0757 | 0.0645 | 0.0707 | 0.0787 | 0.0853 | 0.0862 |
| | | | | | Panel B: RMSE | | | | | | |
| 0 | 0 | 0.0034 | 0.0059 | 0.0047 | 0.0035 | 0.0035 | 0.0036 | 0.0037 | 0.0036 | 0.0036 | 0.0035 |
| | 0.3 | 0.0045 | 0.0059 | 0.0049 | 0.0035 | 0.0035 | 0.0038 | 0.0039 | 0.0040 | 0.0041 | 0.0035 |
| | 0.6 | 0.0068 | 0.0059 | 0.0050 | 0.0035 | 0.0036 | 0.0042 | 0.0048 | 0.0051 | 0.0054 | 0.0035 |
| | 0.9 | 0.0096 | 0.0059 | 0.0051 | 0.0036 | 0.0038 | 0.0050 | 0.0059 | 0.0066 | 0.0071 | 0.0036 |
| 0.3 | 0 | 0.0048 | 0.0084 | 0.0075 | 0.0050 | 0.0050 | 0.0052 | 0.0052 | 0.0052 | 0.0051 | 0.0049 |
| | 0.3 | 0.0069 | 0.0084 | 0.0075 | 0.0050 | 0.0051 | 0.0055 | 0.0058 | 0.0060 | 0.0061 | 0.0050 |
| | 0.6 | 0.0110 | 0.0084 | 0.0077 | 0.0051 | 0.0053 | 0.0064 | 0.0073 | 0.0080 | 0.0085 | 0.0051 |
| | 0.9 | 0.0157 | 0.0084 | 0.0078 | 0.0053 | 0.0057 | 0.0076 | 0.0093 | 0.0105 | 0.0114 | 0.0053 |
| 0.6 | 0 | 0.0083 | 0.0146 | 0.0145 | 0.0086 | 0.0087 | 0.0089 | 0.0090 | 0.0089 | 0.0089 | 0.0085 |
| | 0.3 | 0.0128 | 0.0146 | 0.0148 | 0.0090 | 0.0090 | 0.0097 | 0.0103 | 0.0108 | 0.0111 | 0.0090 |
| | 0.6 | 0.0213 | 0.0147 | 0.0151 | 0.0102 | 0.0101 | 0.0119 | 0.0137 | 0.0152 | 0.0162 | 0.0103 |
| | 0.9 | 0.0307 | 0.0148 | 0.0152 | 0.0121 | 0.0117 | 0.0148 | 0.0181 | 0.0206 | 0.0223 | 0.0123 |
| 0.9 | 0 | 0.0299 | 0.0548 | 0.0519 | 0.0311 | 0.0317 | 0.0331 | 0.0333 | 0.0330 | 0.0326 | 0.0310 |
| | 0.3 | 0.0494 | 0.0566 | 0.0540 | 0.0425 | 0.0410 | 0.0399 | 0.0414 | 0.0428 | 0.0439 | 0.0429 |
| | 0.6 | 0.0845 | 0.0620 | 0.0608 | 0.0661 | 0.0611 | 0.0559 | 0.0598 | 0.0641 | 0.0676 | 0.0671 |
| | 0.9 | 0.1223 | 0.0703 | 0.0692 | 0.0930 | 0.0848 | 0.0755 | 0.0818 | 0.0890 | 0.0949 | 0.0946 |

Table 1.5: Finite sample bias and RMSE of the various estimator of $\beta_1$, $N = 25$, $T = 50$, Bartlett kernel

| $\rho_1$ | $\rho_2$ | P-OLS | P-IM | P-DOLS | Panel FM-OLS | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | b=0.06 | 0.1 | 0.3 | 0.5 | 0.7 | 0.9 | AND |
| | | | | | Panel A: Bias | | | | | | |
| 0 | 0 | -0.0001 | 0.0001 | -0.0001 | -0.0001 | -0.0001 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | -0.0001 |
| | 0.3 | 0.0054 | -0.0002 | -0.0002 | -0.0003 | 0.0002 | 0.0015 | 0.0023 | 0.0030 | 0.0034 | 0.0001 |
| | 0.6 | 0.0108 | -0.0004 | 0.0000 | -0.0005 | 0.0005 | 0.0030 | 0.0047 | 0.0060 | 0.0069 | 0.0003 |
| | 0.9 | 0.0162 | -0.0007 | 0.0001 | -0.0007 | 0.0008 | 0.0045 | 0.0071 | 0.0090 | 0.0104 | 0.0004 |
| 0.3 | 0 | -0.0001 | 0.0001 | -0.0001 | -0.0001 | -0.0001 | -0.0001 | 0.0000 | 0.0000 | -0.0001 | -0.0001 |
| | 0.3 | 0.0088 | 0.0002 | 0.0002 | 0.0010 | 0.0012 | 0.0027 | 0.0040 | 0.0050 | 0.0057 | 0.0012 |
| | 0.6 | 0.0177 | 0.0003 | 0.0004 | 0.0021 | 0.0026 | 0.0055 | 0.0080 | 0.0100 | 0.0115 | 0.0024 |
| | 0.9 | 0.0266 | 0.0003 | 0.0005 | 0.0033 | 0.0039 | 0.0083 | 0.0121 | 0.0151 | 0.0173 | 0.0037 |
| 0.6 | 0 | -0.0001 | 0.0002 | -0.0002 | -0.0002 | -0.0002 | -0.0001 | -0.0001 | -0.0001 | -0.0001 | -0.0002 |
| | 0.3 | 0.0172 | 0.0024 | 0.0034 | 0.0067 | 0.0058 | 0.0065 | 0.0085 | 0.0102 | 0.0116 | 0.0060 |
| | 0.6 | 0.0345 | 0.0045 | 0.0062 | 0.0137 | 0.0118 | 0.0131 | 0.0170 | 0.0205 | 0.0232 | 0.0122 |
| | 0.9 | 0.0518 | 0.0066 | 0.0080 | 0.0206 | 0.0178 | 0.0197 | 0.0256 | 0.0308 | 0.0348 | 0.0184 |
| 0.9 | 0 | 0.0000 | 0.0008 | -0.0002 | -0.0001 | -0.0002 | -0.0001 | 0.0000 | 0.0001 | 0.0001 | -0.0002 |
| | 0.3 | 0.0644 | 0.0378 | 0.0467 | 0.0534 | 0.0505 | 0.0441 | 0.0451 | 0.0480 | 0.0510 | 0.0512 |
| | 0.6 | 0.1289 | 0.0747 | 0.0919 | 0.1070 | 0.1011 | 0.0884 | 0.0901 | 0.0959 | 0.1018 | 0.1026 |
| | 0.9 | 0.1933 | 0.1116 | 0.1354 | 0.1605 | 0.1518 | 0.1327 | 0.1351 | 0.1439 | 0.1527 | 0.1540 |
| | | | | | Panel B: RMSE | | | | | | |
| 0 | 0 | 0.0040 | 0.0068 | 0.0049 | 0.0041 | 0.0041 | 0.0042 | 0.0042 | 0.0042 | 0.0042 | 0.0041 |
| | 0.3 | 0.0068 | 0.0068 | 0.0051 | 0.0041 | 0.0041 | 0.0045 | 0.0049 | 0.0052 | 0.0055 | 0.0041 |
| | 0.6 | 0.0118 | 0.0068 | 0.0052 | 0.0042 | 0.0042 | 0.0054 | 0.0065 | 0.0075 | 0.0083 | 0.0042 |
| | 0.9 | 0.0171 | 0.0068 | 0.0053 | 0.0042 | 0.0043 | 0.0066 | 0.0086 | 0.0103 | 0.0115 | 0.0043 |
| 0.3 | 0 | 0.0056 | 0.0097 | 0.0071 | 0.0058 | 0.0058 | 0.0059 | 0.0060 | 0.0059 | 0.0059 | 0.0058 |
| | 0.3 | 0.0106 | 0.0097 | 0.0072 | 0.0059 | 0.0060 | 0.0066 | 0.0073 | 0.0079 | 0.0083 | 0.0059 |
| | 0.6 | 0.0190 | 0.0097 | 0.0072 | 0.0062 | 0.0065 | 0.0084 | 0.0103 | 0.0120 | 0.0132 | 0.0064 |
| | 0.9 | 0.0278 | 0.0097 | 0.0074 | 0.0067 | 0.0072 | 0.0107 | 0.0140 | 0.0167 | 0.0188 | 0.0071 |
| 0.6 | 0 | 0.0096 | 0.0167 | 0.0122 | 0.0099 | 0.0100 | 0.0102 | 0.0102 | 0.0101 | 0.0100 | 0.0099 |
| | 0.3 | 0.0199 | 0.0169 | 0.0128 | 0.0120 | 0.0116 | 0.0122 | 0.0135 | 0.0146 | 0.0155 | 0.0117 |
| | 0.6 | 0.0364 | 0.0173 | 0.0141 | 0.0171 | 0.0157 | 0.0170 | 0.0204 | 0.0234 | 0.0258 | 0.0160 |
| | 0.9 | 0.0536 | 0.0180 | 0.0152 | 0.0232 | 0.0208 | 0.0229 | 0.0284 | 0.0333 | 0.0370 | 0.0213 |
| 0.9 | 0 | 0.0322 | 0.0578 | 0.0397 | 0.0332 | 0.0337 | 0.0351 | 0.0352 | 0.0349 | 0.0345 | 0.0336 |
| | 0.3 | 0.0725 | 0.0691 | 0.0618 | 0.0633 | 0.0610 | 0.0567 | 0.0576 | 0.0597 | 0.0619 | 0.0616 |
| | 0.6 | 0.1338 | 0.0946 | 0.1012 | 0.1128 | 0.1073 | 0.0959 | 0.0976 | 0.1030 | 0.1084 | 0.1087 |
| | 0.9 | 0.1975 | 0.1260 | 0.1427 | 0.1651 | 0.1566 | 0.1384 | 0.1410 | 0.1494 | 0.1578 | 0.1587 |

Table 1.6: Finite sample bias and RMSE of the various estimator of $\beta_1$, $N = 25$, $T = 100$, Bartlett kernel

| $\rho_1$ | $\rho_2$ | P-OLS | P-IM | P-DOLS | Panel FM-OLS | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | b=0.06 | 0.1 | 0.3 | 0.5 | 0.7 | 0.9 | AND |
| | | | | | Panel A: Bias | | | | | | |
| 0 | 0 | 0.0000 | -0.0001 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | -0.0001 | -0.0001 | 0.0000 |
| | 0.3 | 0.0027 | -0.0002 | 0.0000 | 0.0000 | 0.0002 | 0.0007 | 0.0011 | 0.0015 | 0.0017 | 0.0000 |
| | 0.6 | 0.0054 | -0.0002 | 0.0001 | 0.0000 | 0.0004 | 0.0015 | 0.0023 | 0.0030 | 0.0034 | 0.0000 |
| | 0.9 | 0.0081 | -0.0003 | 0.0000 | 0.0001 | 0.0006 | 0.0023 | 0.0035 | 0.0045 | 0.0052 | 0.0000 |
| 0.3 | 0 | -0.0001 | -0.0001 | 0.0000 | 0.0000 | 0.0000 | -0.0001 | -0.0001 | -0.0001 | -0.0001 | -0.0001 |
| | 0.3 | 0.0044 | -0.0001 | 0.0001 | 0.0004 | 0.0005 | 0.0013 | 0.0019 | 0.0024 | 0.0028 | 0.0004 |
| | 0.6 | 0.0089 | -0.0001 | 0.0002 | 0.0008 | 0.0011 | 0.0027 | 0.0040 | 0.0050 | 0.0057 | 0.0008 |
| | 0.9 | 0.0134 | -0.0001 | 0.0002 | 0.0013 | 0.0017 | 0.0040 | 0.0060 | 0.0075 | 0.0086 | 0.0012 |
| 0.6 | 0 | -0.0001 | -0.0002 | -0.0001 | -0.0001 | -0.0001 | -0.0001 | -0.0001 | -0.0001 | -0.0001 | -0.0001 |
| | 0.3 | 0.0087 | 0.0004 | 0.0008 | 0.0023 | 0.0020 | 0.0029 | 0.0040 | 0.0050 | 0.0057 | 0.0024 |
| | 0.6 | 0.0175 | 0.0009 | 0.0011 | 0.0047 | 0.0041 | 0.0059 | 0.0082 | 0.0101 | 0.0115 | 0.0049 |
| | 0.9 | 0.0264 | 0.0015 | 0.0012 | 0.0071 | 0.0062 | 0.0089 | 0.0123 | 0.0151 | 0.0173 | 0.0074 |
| 0.9 | 0 | -0.0003 | -0.0009 | -0.0001 | -0.0003 | -0.0002 | -0.0003 | -0.0003 | -0.0003 | -0.0003 | -0.0003 |
| | 0.3 | 0.0360 | 0.0130 | 0.0142 | 0.0263 | 0.0235 | 0.0194 | 0.0212 | 0.0238 | 0.0261 | 0.0269 |
| | 0.6 | 0.0722 | 0.0270 | 0.0271 | 0.0530 | 0.0472 | 0.0391 | 0.0427 | 0.0479 | 0.0525 | 0.0540 |
| | 0.9 | 0.1084 | 0.0409 | 0.0390 | 0.0796 | 0.0709 | 0.0588 | 0.0642 | 0.0720 | 0.0789 | 0.0811 |
| | | | | | Panel B: RMSE | | | | | | |
| 0 | 0 | 0.0020 | 0.0035 | 0.0026 | 0.0020 | 0.0021 | 0.0021 | 0.0021 | 0.0021 | 0.0021 | 0.0020 |
| | 0.3 | 0.0034 | 0.0035 | 0.0027 | 0.0020 | 0.0021 | 0.0023 | 0.0025 | 0.0026 | 0.0027 | 0.0020 |
| | 0.6 | 0.0059 | 0.0035 | 0.0028 | 0.0021 | 0.0021 | 0.0027 | 0.0033 | 0.0038 | 0.0041 | 0.0021 |
| | 0.9 | 0.0085 | 0.0035 | 0.0029 | 0.0021 | 0.0023 | 0.0033 | 0.0043 | 0.0051 | 0.0058 | 0.0021 |
| 0.3 | 0 | 0.0028 | 0.0050 | 0.0041 | 0.0029 | 0.0029 | 0.0030 | 0.0030 | 0.0030 | 0.0030 | 0.0029 |
| | 0.3 | 0.0053 | 0.0050 | 0.0042 | 0.0029 | 0.0030 | 0.0033 | 0.0037 | 0.0040 | 0.0042 | 0.0029 |
| | 0.6 | 0.0095 | 0.0050 | 0.0043 | 0.0031 | 0.0032 | 0.0042 | 0.0052 | 0.0060 | 0.0066 | 0.0030 |
| | 0.9 | 0.0140 | 0.0050 | 0.0044 | 0.0032 | 0.0035 | 0.0053 | 0.0070 | 0.0083 | 0.0094 | 0.0032 |
| 0.6 | 0 | 0.0049 | 0.0087 | 0.0081 | 0.0050 | 0.0051 | 0.0052 | 0.0053 | 0.0052 | 0.0051 | 0.0050 |
| | 0.3 | 0.0102 | 0.0087 | 0.0083 | 0.0056 | 0.0055 | 0.0061 | 0.0068 | 0.0073 | 0.0078 | 0.0056 |
| | 0.6 | 0.0186 | 0.0087 | 0.0085 | 0.0070 | 0.0067 | 0.0081 | 0.0100 | 0.0116 | 0.0129 | 0.0072 |
| | 0.9 | 0.0273 | 0.0088 | 0.0086 | 0.0089 | 0.0083 | 0.0107 | 0.0139 | 0.0165 | 0.0185 | 0.0092 |
| 0.9 | 0 | 0.0181 | 0.0328 | 0.0297 | 0.0186 | 0.0189 | 0.0196 | 0.0197 | 0.0195 | 0.0192 | 0.0186 |
| | 0.3 | 0.0407 | 0.0354 | 0.0333 | 0.0326 | 0.0304 | 0.0279 | 0.0293 | 0.0311 | 0.0327 | 0.0330 |
| | 0.6 | 0.0753 | 0.0426 | 0.0412 | 0.0568 | 0.0515 | 0.0445 | 0.0479 | 0.0526 | 0.0567 | 0.0578 |
| | 0.9 | 0.1113 | 0.0527 | 0.0504 | 0.0827 | 0.0743 | 0.0631 | 0.0684 | 0.0759 | 0.0824 | 0.0843 |

Table 1.7: Empirical null rejection probabilities, 0.05 level, t-tests for $h_0 : \beta_1 = 1$, $N = 5$, data dependent bandwidths and lag lengths.

| $\rho_1, \rho_2$ | P-OLS | Bartlett kernel | | | | | QS kernel | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | P-DOLS | P-FM | P-IM(O) | P-IM(D) | P-IM(Fb) | P-DOLS | P-FM | P-IM(O) | P-IM(D) | P-IM(Fb) |
| Panel A: $T = 50$ | | | | | | | | | | | |
| 0 | 0.0514 | 0.1294 | 0.0762 | 0.0724 | 0.0544 | 0.0744 | 0.1264 | 0.0836 | 0.0808 | 0.0624 | 0.0422 |
| 0.3 | 0.2178 | 0.1792 | 0.1032 | 0.0974 | 0.0796 | 0.132 | 0.1632 | 0.0984 | 0.0936 | 0.0746 | 0.101 |
| 0.6 | 0.6064 | 0.2452 | 0.2214 | 0.1462 | 0.1212 | 0.2694 | 0.2176 | 0.1908 | 0.1246 | 0.109 | 0.1924 |
| 0.9 | 0.9366 | 0.6068 | 0.777 | 0.4804 | 0.422 | 0.7628 | 0.5736 | 0.7336 | 0.4462 | 0.4042 | 0.7052 |
| Panel B: $T = 100$ | | | | | | | | | | | |
| 0 | 0.0472 | 0.1416 | 0.0588 | 0.0586 | 0.0516 | 0.0524 | 0.1402 | 0.062 | 0.0626 | 0.0524 | 0.0458 |
| 0.3 | 0.2082 | 0.227 | 0.0794 | 0.0774 | 0.0678 | 0.1038 | 0.2158 | 0.0756 | 0.0726 | 0.0602 | 0.0866 |
| 0.6 | 0.5914 | 0.3256 | 0.1584 | 0.111 | 0.0952 | 0.1776 | 0.2976 | 0.1326 | 0.094 | 0.0784 | 0.1308 |
| 0.9 | 0.934 | 0.4804 | 0.7266 | 0.3376 | 0.286 | 0.6198 | 0.4494 | 0.6974 | 0.3112 | 0.2626 | 0.561 |

Table 1.8: Empirical null rejection probabilities, 0.05 level, t-tests for $h_0 : \beta_1 = 1$, $N = 10$, data dependent bandwidths and lag lengths.

| $\rho_1, \rho_2$ | P-OLS | Bartlett kernel | | | | | QS kernel | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | P-DOLS | P-FM | P-IM(O) | P-IM(D) | P-IM(Fb) | P-DOLS | P-FM | P-IM(O) | P-IM(D) | P-IM(Fb) |
| Panel A: $T = 50$ | | | | | | | | | | | |
| 0 | 0.051 | 0.1142 | 0.0654 | 0.0708 | 0.0642 | 0.0762 | 0.1116 | 0.0712 | 0.08 | 0.072 | 0.0516 |
| 0.3 | 0.2846 | 0.1652 | 0.0902 | 0.0982 | 0.0908 | 0.1418 | 0.1454 | 0.0848 | 0.0912 | 0.0852 | 0.1054 |
| 0.6 | 0.7924 | 0.2188 | 0.2422 | 0.1482 | 0.1352 | 0.3306 | 0.1898 | 0.1994 | 0.1256 | 0.1222 | 0.2892 |
| 0.9 | 0.9924 | 0.7518 | 0.9332 | 0.5832 | 0.5276 | 0.8576 | 0.7094 | 0.908 | 0.5464 | 0.5094 | 0.8362 |
| Panel B: $T = 100$ | | | | | | | | | | | |
| 0 | 0.0476 | 0.139 | 0.0542 | 0.0582 | 0.0534 | 0.0536 | 0.137 | 0.0558 | 0.0628 | 0.057 | 0.046 |
| 0.3 | 0.2882 | 0.2252 | 0.0758 | 0.0794 | 0.075 | 0.1096 | 0.21 | 0.0662 | 0.0728 | 0.0688 | 0.0908 |
| 0.6 | 0.787 | 0.308 | 0.1802 | 0.1098 | 0.0984 | 0.1776 | 0.281 | 0.1466 | 0.0938 | 0.0848 | 0.1366 |
| 0.9 | 0.9896 | 0.4982 | 0.9002 | 0.3888 | 0.3396 | 0.6724 | 0.4598 | 0.8742 | 0.3588 | 0.316 | 0.6254 |

Table 1.9: Empirical null rejection probabilities, 0.05 level, t-tests for $h_0 : \beta_1 = 1$, $N = 25$, data dependent bandwidths and lag lengths.

| $\rho_1, \rho_2$ | P-OLS | Bartlett kernel | | | | | QS kernel | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | P-DOLS | P-FM | P-IM(O) | P-IM(D) | P-IM(Fb) | P-DOLS | P-FM | P-IM(O) | P-IM(D) | P-IM(Fb) |
| Panel A: $T = 50$ | | | | | | | | | | | |
| 0 | 0.0482 | 0.1022 | 0.0566 | 0.0622 | 0.0592 | 0.0664 | 0.099 | 0.0622 | 0.0706 | 0.0682 | 0.0404 |
| 0.3 | 0.5116 | 0.1534 | 0.0822 | 0.0848 | 0.0826 | 0.1292 | 0.1358 | 0.0724 | 0.0796 | 0.079 | 0.0972 |
| 0.6 | 0.9786 | 0.2362 | 0.3852 | 0.1428 | 0.1328 | 0.335 | 0.2042 | 0.2952 | 0.1204 | 0.1202 | 0.2966 |
| 0.9 | 1 | 0.962 | 0.9976 | 0.8058 | 0.7674 | 0.9542 | 0.9448 | 0.9966 | 0.7804 | 0.7474 | 0.9476 |
| Panel B: $T = 100$ | | | | | | | | | | | |
| 0 | 0.0476 | 0.1354 | 0.0554 | 0.0612 | 0.0588 | 0.0558 | 0.1332 | 0.0562 | 0.0638 | 0.0612 | 0.046 |
| 0.3 | 0.5124 | 0.2032 | 0.0766 | 0.08 | 0.079 | 0.1106 | 0.188 | 0.067 | 0.0726 | 0.0716 | 0.0954 |
| 0.6 | 0.9744 | 0.2752 | 0.2708 | 0.1128 | 0.1056 | 0.1844 | 0.2478 | 0.2064 | 0.0988 | 0.095 | 0.1424 |
| 0.9 | 1 | 0.6172 | 0.9958 | 0.5474 | 0.5132 | 0.7944 | 0.5736 | 0.993 | 0.5224 | 0.4876 | 0.7528 |

Table 1.10: Empirical null rejection probabilities, 0.05 level, Wald-tests for $h_0 : \beta_1 = 1, \beta_2 = 1$, $N = 5$, data dependent bandwidths and lag lengths.

| $\rho_1, \rho_2$ | P-OLS | Bartlett kernel | | | | | QS kernel | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | P-DOLS | P-FM | P-IM(O) | P-IM(D) | P-IM(Fb) | P-DOLS | P-FM | P-IM(O) | P-IM(D) | P-IM(Fb) |
| Panel A: $T = 50$ | | | | | | | | | | | |
| 0 | 0.05 | 0.1552 | 0.0794 | 0.0804 | 0.0574 | 0.0794 | 0.1514 | 0.0938 | 0.1006 | 0.0696 | 0.04 |
| 0.3 | 0.3026 | 0.2344 | 0.1254 | 0.1188 | 0.0952 | 0.1746 | 0.2086 | 0.1168 | 0.1156 | 0.088 | 0.122 |
| 0.6 | 0.822 | 0.3362 | 0.3098 | 0.1996 | 0.161 | 0.3802 | 0.297 | 0.2578 | 0.167 | 0.1408 | 0.2664 |
| 0.9 | 0.9972 | 0.8098 | 0.9486 | 0.7018 | 0.6244 | 0.9414 | 0.7728 | 0.9202 | 0.6646 | 0.6006 | 0.9068 |
| Panel B: $T = 100$ | | | | | | | | | | | |
| 0 | 0.049 | 0.185 | 0.0626 | 0.0636 | 0.0478 | 0.056 | 0.1832 | 0.0686 | 0.071 | 0.052 | 0.0418 |
| 0.3 | 0.2898 | 0.3258 | 0.0908 | 0.0912 | 0.0772 | 0.1302 | 0.3048 | 0.085 | 0.0812 | 0.067 | 0.105 |
| 0.6 | 0.8276 | 0.4574 | 0.2068 | 0.1372 | 0.1142 | 0.2356 | 0.4216 | 0.1696 | 0.1168 | 0.0952 | 0.169 |
| 0.9 | 0.9982 | 0.6786 | 0.9156 | 0.5056 | 0.4282 | 0.8274 | 0.6352 | 0.8962 | 0.4684 | 0.3962 | 0.7724 |

Table 1.11: Empirical null rejection probabilities, 0.05 level, Wald-tests for $h_0 : \beta_1 = 1, \beta_2 = 1$, $N = 10$, data dependent bandwidths and lag lengths.

| $\rho_1, \rho_2$ | P-OLS | Bartlett kernel | | | | | QS kernel | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | P-DOLS | P-FM | P-IM(O) | P-IM(D) | P-IM(Fb) | P-DOLS | P-FM | P-IM(O) | P-IM(D) | P-IM(Fb) |
| Panel A: $T = 50$ | | | | | | | | | | | |
| 0 | 0.0486 | 0.1396 | 0.0708 | 0.0798 | 0.0704 | 0.0874 | 0.134 | 0.0766 | 0.0928 | 0.078 | 0.049 |
| 0.3 | 0.4166 | 0.2162 | 0.109 | 0.1148 | 0.1038 | 0.1782 | 0.1886 | 0.099 | 0.1094 | 0.0976 | 0.127 |
| 0.6 | 0.9562 | 0.3076 | 0.3512 | 0.1914 | 0.1706 | 0.4722 | 0.2586 | 0.2786 | 0.1558 | 0.1494 | 0.4046 |
| 0.9 | 1 | 0.916 | 0.9952 | 0.8002 | 0.7398 | 0.9756 | 0.8894 | 0.9888 | 0.765 | 0.7182 | 0.9672 |
| Panel B: $T = 100$ | | | | | | | | | | | |
| 0 | 0.0484 | 0.1718 | 0.0602 | 0.0636 | 0.0564 | 0.0546 | 0.169 | 0.065 | 0.068 | 0.0608 | 0.0444 |
| 0.3 | 0.4212 | 0.311 | 0.085 | 0.091 | 0.082 | 0.1342 | 0.2854 | 0.0762 | 0.0806 | 0.073 | 0.1084 |
| 0.6 | 0.9602 | 0.4354 | 0.2532 | 0.1358 | 0.1186 | 0.2432 | 0.3916 | 0.1936 | 0.1106 | 0.0964 | 0.1782 |
| 0.9 | 1 | 0.688 | 0.9916 | 0.566 | 0.5024 | 0.8694 | 0.6454 | 0.9854 | 0.528 | 0.468 | 0.8304 |

Table 1.12: Empirical null rejection probabilities, 0.05 level, Wald-tests for $h_0 : \beta_1 = 1, \beta_2 = 1$, $N = 25$, data dependent bandwidths and lag lengths.

| $\rho_1, \rho_2$ | P-OLS | Bartlett kernel | | | | | QS kernel | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | P-DOLS | P-FM | P-IM(O) | P-IM(D) | P-IM(Fb) | P-DOLS | P-FM | P-IM(O) | P-IM(D) | P-IM(Fb) |
| Panel A: $T = 50$ | | | | | | | | | | | |
| 0 | 0.0528 | 0.132 | 0.0656 | 0.073 | 0.0702 | 0.0836 | 0.1292 | 0.0702 | 0.085 | 0.0806 | 0.0472 |
| 0.3 | 0.7142 | 0.208 | 0.1128 | 0.1078 | 0.105 | 0.1686 | 0.1796 | 0.0984 | 0.1016 | 0.0982 | 0.1202 |
| 0.6 | 0.9998 | 0.3328 | 0.5454 | 0.1914 | 0.1786 | 0.484 | 0.2736 | 0.4302 | 0.1598 | 0.1608 | 0.4234 |
| 0.9 | 1 | 0.9986 | 1 | 0.9474 | 0.9264 | 0.9978 | 0.9972 | 1 | 0.9344 | 0.9138 | 0.9962 |
| Panel B: $T = 100$ | | | | | | | | | | | |
| 0 | 0.0538 | 0.1706 | 0.0606 | 0.0628 | 0.0596 | 0.058 | 0.1674 | 0.0628 | 0.069 | 0.065 | 0.046 |
| 0.3 | 0.7098 | 0.2822 | 0.0868 | 0.0876 | 0.0866 | 0.135 | 0.257 | 0.076 | 0.08 | 0.0786 | 0.1074 |
| 0.6 | 0.9998 | 0.4058 | 0.3832 | 0.1362 | 0.125 | 0.2498 | 0.3598 | 0.2922 | 0.1124 | 0.1054 | 0.1796 |
| 0.9 | 1 | 0.8108 | 1 | 0.7578 | 0.7186 | 0.9488 | 0.774 | 1 | 0.7304 | 0.6874 | 0.93 |

Table 1.13: Fixed-b asymptotic critical value for t-test of $\beta$ in regression with intercept and two regressors, $N = 25$, Bartlett kernel

| b | 0.02 | 0.04 | 0.06 | 0.08 | 0.10 | 0.12 | 0.14 | 0.16 | 0.18 | 0.20 |
|---|---|---|---|---|---|---|---|---|---|---|
| 95% | 1.7329 | 1.9363 | 2.1731 | 2.4357 | 2.7227 | 3.0220 | 3.3221 | 3.6112 | 3.8807 | 4.1140 |
| 97.5% | 2.0630 | 2.3079 | 2.5845 | 2.8934 | 3.2298 | 3.5864 | 3.9396 | 4.2836 | 4.5986 | 4.8755 |
| 99% | 2.4683 | 2.7561 | 3.0980 | 3.4713 | 3.8776 | 4.3055 | 4.7293 | 5.1508 | 5.5243 | 5.8506 |
| 99.5% | 2.7275 | 3.0599 | 3.4300 | 3.8481 | 4.3085 | 4.7773 | 5.2548 | 5.7118 | 6.1411 | 6.5021 |
| b | 0.22 | 0.24 | 0.26 | 0.28 | 0.3 | 0.32 | 0.34 | 0.36 | 0.38 | 0.40 |
| 95% | 4.3072 | 4.4702 | 4.6004 | 4.7217 | 4.8296 | 4.9306 | 5.0294 | 5.1291 | 5.2242 | 5.3160 |
| 97.5% | 5.1085 | 5.2956 | 5.4627 | 5.5963 | 5.7331 | 5.8559 | 5.9673 | 6.0789 | 6.1903 | 6.3132 |
| 99% | 6.1355 | 6.3751 | 6.5593 | 6.7345 | 6.8818 | 7.0326 | 7.1687 | 7.3157 | 7.4607 | 7.5930 |
| 99.5% | 6.8015 | 7.0752 | 7.2862 | 7.4722 | 7.6352 | 7.7980 | 7.9644 | 8.1320 | 8.2883 | 8.4333 |
| b | 0.42 | 0.44 | 0.46 | 0.48 | 0.5 | 0.52 | 0.54 | 0.56 | 0.58 | 0.60 |
| 95% | 5.4079 | 5.5070 | 5.5967 | 5.6805 | 5.7679 | 5.8447 | 5.9164 | 5.9942 | 6.0680 | 6.1366 |
| 97.5% | 6.4189 | 6.5329 | 6.6388 | 6.7461 | 6.8517 | 6.9494 | 7.0410 | 7.1285 | 7.2094 | 7.2829 |
| 99% | 7.7228 | 7.8553 | 7.9922 | 8.1188 | 8.2377 | 8.3674 | 8.4834 | 8.5767 | 8.6779 | 8.7842 |
| 99.5% | 8.5820 | 8.7258 | 8.8732 | 9.0041 | 9.1516 | 9.2895 | 9.4089 | 9.5389 | 9.6542 | 9.7574 |
| b | 0.62 | 0.64 | 0.66 | 0.68 | 0.7 | 0.72 | 0.74 | 0.76 | 0.78 | 0.80 |
| 95% | 6.2004 | 6.2597 | 6.3208 | 6.3811 | 6.4329 | 6.4885 | 6.5452 | 6.5982 | 6.6506 | 6.7020 |
| 97.5% | 7.3600 | 7.4337 | 7.5122 | 7.5804 | 7.6560 | 7.7180 | 7.7900 | 7.8533 | 7.9185 | 7.9721 |
| 99% | 8.8838 | 8.9709 | 9.0515 | 9.1402 | 9.2231 | 9.3065 | 9.3759 | 9.4616 | 9.5297 | 9.5959 |
| 99.5% | 9.8619 | 9.9699 | 10.0559 | 10.1605 | 10.2384 | 10.3414 | 10.4381 | 10.5210 | 10.6046 | 10.6947 |
| b | 0.82 | 0.84 | 0.86 | 0.88 | 0.9 | 0.92 | 0.94 | 0.96 | 0.98 | 1.00 |
| 95% | 6.7480 | 6.7906 | 6.8332 | 6.8817 | 6.9216 | 6.9658 | 7.0077 | 7.0485 | 7.0834 | 7.1205 |
| 97.5% | 8.0248 | 8.0815 | 8.1347 | 8.1868 | 8.2422 | 8.2949 | 8.3433 | 8.3876 | 8.4348 | 8.4781 |
| 99% | 9.6626 | 9.7342 | 9.8024 | 9.8573 | 9.9243 | 9.9824 | 10.0446 | 10.1029 | 10.1630 | 10.2177 |
| 99.5% | 10.7659 | 10.8418 | 10.9134 | 10.9985 | 11.0667 | 11.1332 | 11.1945 | 11.2611 | 11.3375 | 11.4049 |

Table 1.14: Fixed-b asymptotic critical value for t-test of $\beta$ in regression with intercept and two regressors, $N = 25$, QS kernel

| b | 0.02 | 0.04 | 0.06 | 0.08 | 0.10 | 0.12 | 0.14 | 0.16 | 0.18 | 0.20 |
|---|---|---|---|---|---|---|---|---|---|---|
| 95% | 1.7870 | 2.0829 | 2.4752 | 2.9884 | 3.6602 | 4.5278 | 5.5699 | 6.6972 | 7.7744 | 8.6329 |
| 97.5% | 2.1270 | 2.4815 | 2.9371 | 3.5436 | 4.3513 | 5.3846 | 6.6572 | 8.0320 | 9.3128 | 10.3492 |
| 99% | 2.5440 | 2.9713 | 3.5281 | 4.2744 | 5.2392 | 6.5058 | 8.0225 | 9.6946 | 11.2728 | 12.5370 |
| 99.5% | 2.8123 | 3.2967 | 3.9182 | 4.7454 | 5.8424 | 7.2470 | 8.9767 | 10.8163 | 12.6241 | 14.0653 |
| b | 0.22 | 0.24 | 0.26 | 0.28 | 0.3 | 0.32 | 0.34 | 0.36 | 0.38 | 0.40 |
| 95% | 9.2475 | 9.6378 | 9.8866 | 10.0412 | 10.1446 | 10.2071 | 10.2604 | 10.2899 | 10.3154 | 10.3393 |
| 97.5% | 11.1166 | 11.5874 | 11.8763 | 12.0800 | 12.2085 | 12.2906 | 12.3468 | 12.3913 | 12.4277 | 12.4530 |
| 99% | 13.4557 | 14.0671 | 14.4533 | 14.7158 | 14.8841 | 14.9774 | 15.0857 | 15.1470 | 15.1848 | 15.2096 |
| 99.5% | 15.1063 | 15.7767 | 16.2224 | 16.5170 | 16.6887 | 16.8283 | 16.9298 | 16.9828 | 17.0212 | 17.0680 |
| b | 0.42 | 0.44 | 0.46 | 0.48 | 0.5 | 0.52 | 0.54 | 0.56 | 0.58 | 0.60 |
| 95% | 10.3527 | 10.3643 | 10.3739 | 10.3828 | 10.3897 | 10.3990 | 10.4029 | 10.4086 | 10.4133 | 10.4200 |
| 97.5% | 12.4740 | 12.4893 | 12.5025 | 12.5136 | 12.5235 | 12.5329 | 12.5409 | 12.5420 | 12.5483 | 12.5543 |
| 99% | 15.2323 | 15.2485 | 15.2564 | 15.2734 | 15.2875 | 15.2977 | 15.3124 | 15.3196 | 15.3307 | 15.3435 |
| 99.5% | 17.0886 | 17.1114 | 17.1343 | 17.1512 | 17.1645 | 17.1798 | 17.1817 | 17.1869 | 17.2133 | 17.2270 |
| b | 0.62 | 0.64 | 0.66 | 0.68 | 0.7 | 0.72 | 0.74 | 0.76 | 0.78 | 0.80 |
| 95% | 10.4231 | 10.4271 | 10.4279 | 10.4262 | 10.4278 | 10.4319 | 10.4349 | 10.4357 | 10.4382 | 10.4410 |
| 97.5% | 12.5603 | 12.5697 | 12.5720 | 12.5781 | 12.5785 | 12.5849 | 12.5881 | 12.5913 | 12.5929 | 12.5952 |
| 99% | 15.3526 | 15.3495 | 15.3554 | 15.3513 | 15.3575 | 15.3655 | 15.3730 | 15.3769 | 15.3827 | 15.3920 |
| 99.5% | 17.2313 | 17.2400 | 17.2423 | 17.2405 | 17.2307 | 17.2383 | 17.2374 | 17.2485 | 17.2494 | 17.2503 |
| b | 0.82 | 0.84 | 0.86 | 0.88 | 0.9 | 0.92 | 0.94 | 0.96 | 0.98 | 1.00 |
| 95% | 10.4437 | 10.4459 | 10.4456 | 10.4464 | 10.4469 | 10.4463 | 10.4480 | 10.4482 | 10.4474 | 10.4481 |
| 97.5% | 12.5982 | 12.5994 | 12.6013 | 12.6013 | 12.6019 | 12.6023 | 12.6019 | 12.6021 | 12.6028 | 12.6043 |
| 99% | 15.3955 | 15.3959 | 15.3964 | 15.3969 | 15.3977 | 15.4018 | 15.4060 | 15.4099 | 15.4127 | 15.4150 |
| 99.5% | 17.2619 | 17.2727 | 17.2749 | 17.2764 | 17.2739 | 17.2792 | 17.2822 | 17.2859 | 17.2892 | 17.2920 |

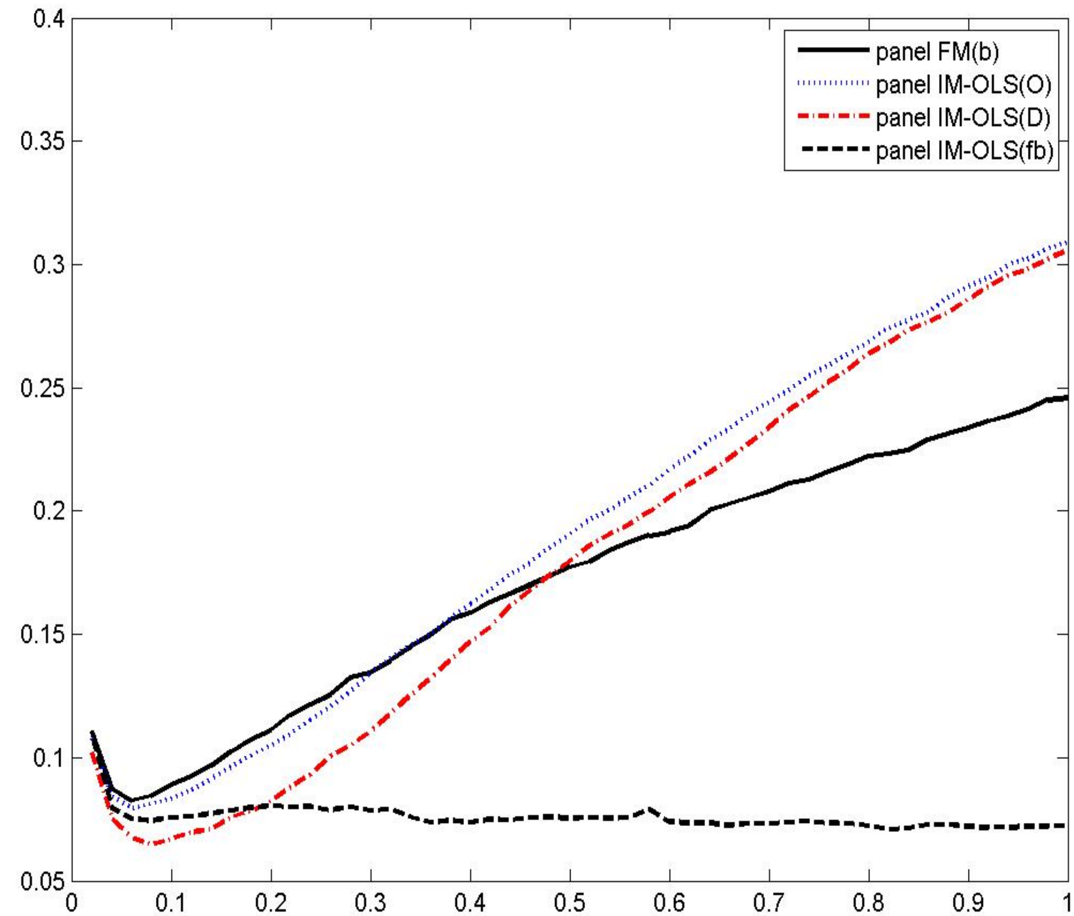


Figure 1.1: Empirical null rejections, t-test, $N = 5$, $T = 100$, $\rho_1 = \rho_2 = 0.3$, Bartlett kernel

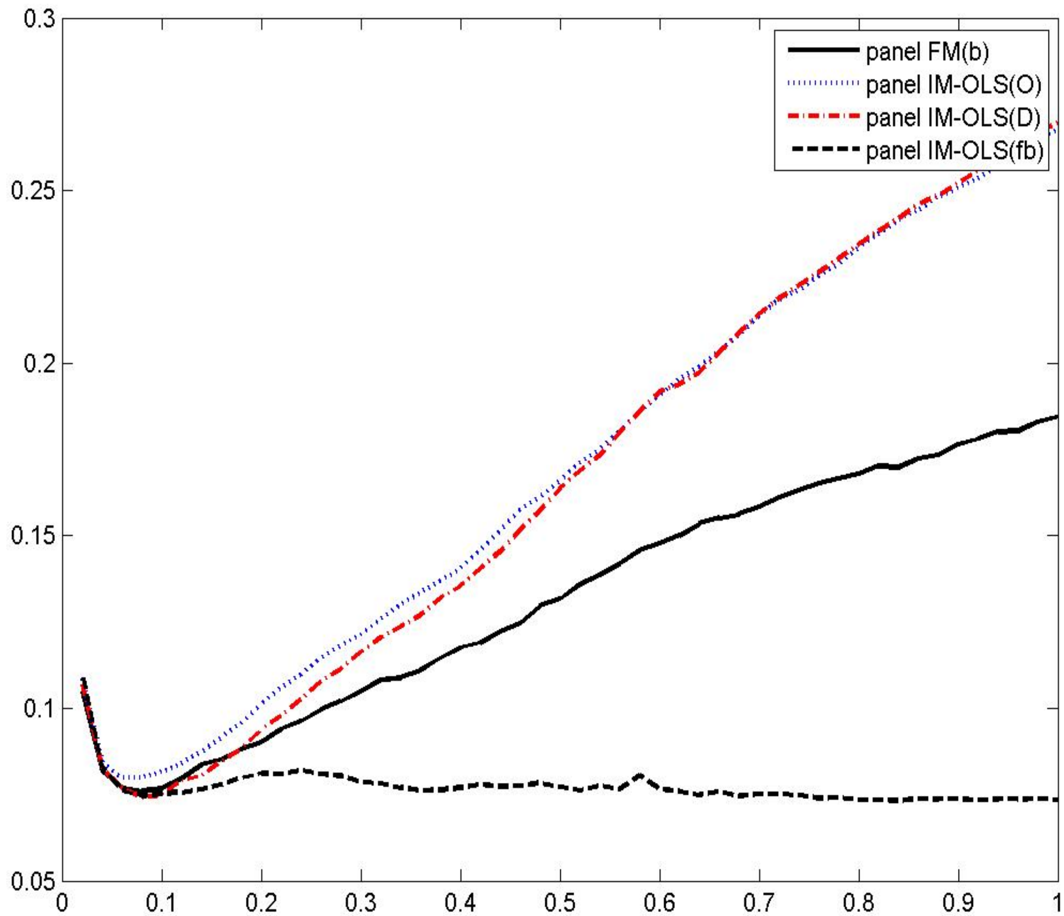


Figure 1.2: Empirical null rejections, t-test, $N = 10$, $T = 100$, $\rho_1 = \rho_2 = 0.3$, Bartlett kernel

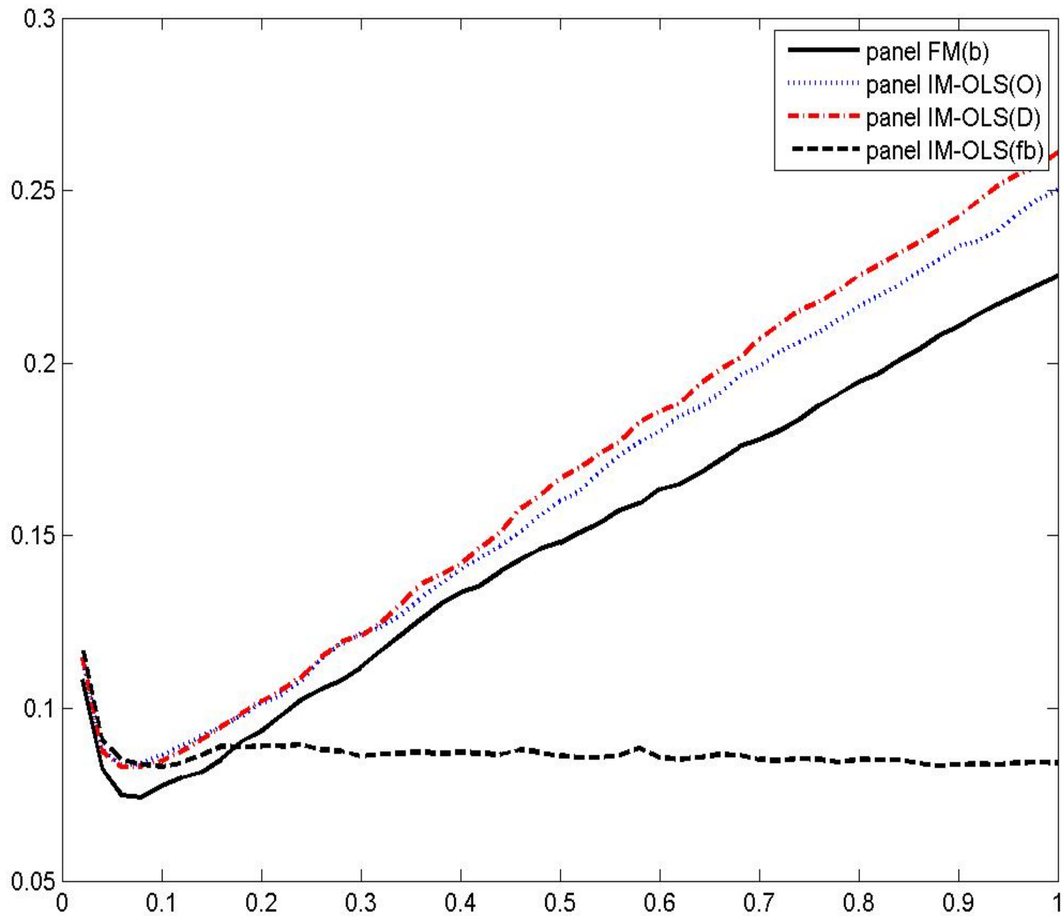


Figure 1.3: Empirical null rejections, t-test, $N = 25$, $T = 100$, $\rho_1 = \rho_2 = 0.3$, Bartlett kernel

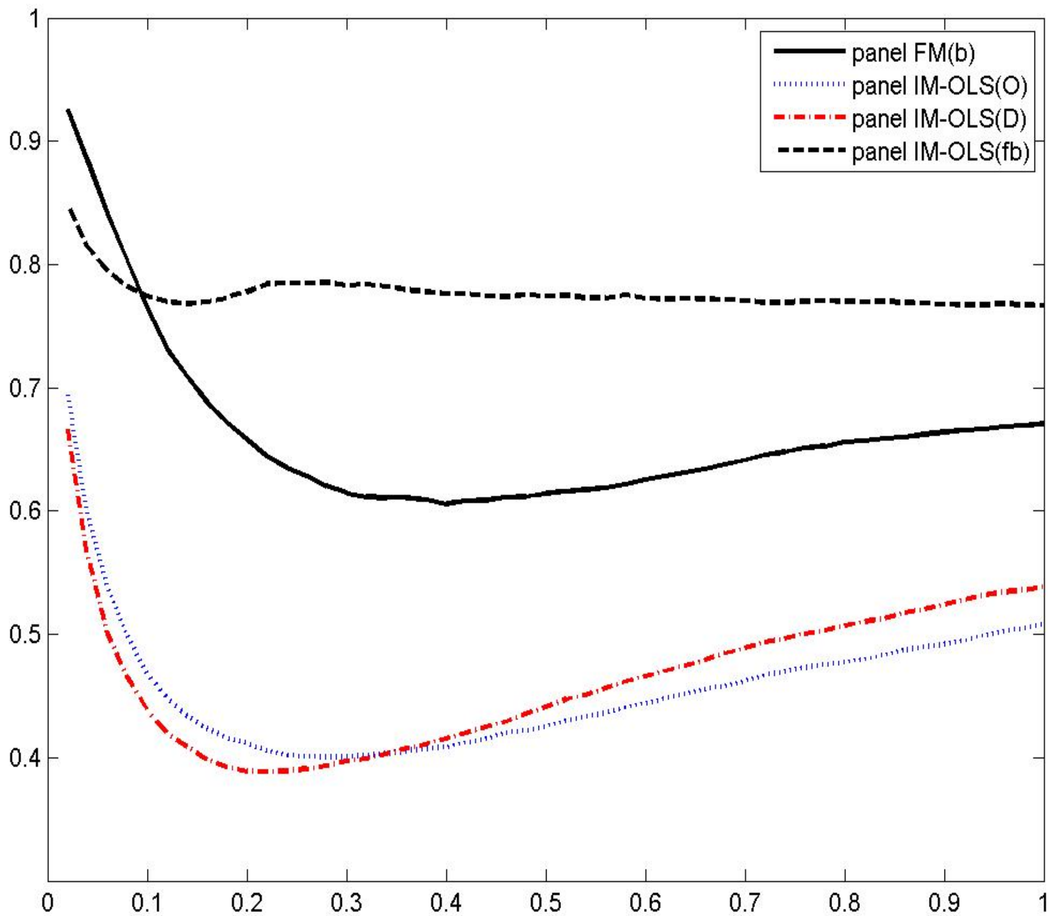


Figure 1.4: Empirical null rejections, t-test, $N = 5$, $T = 50$, $\rho_1 = \rho_2 = 0.9$, Bartlett kernel

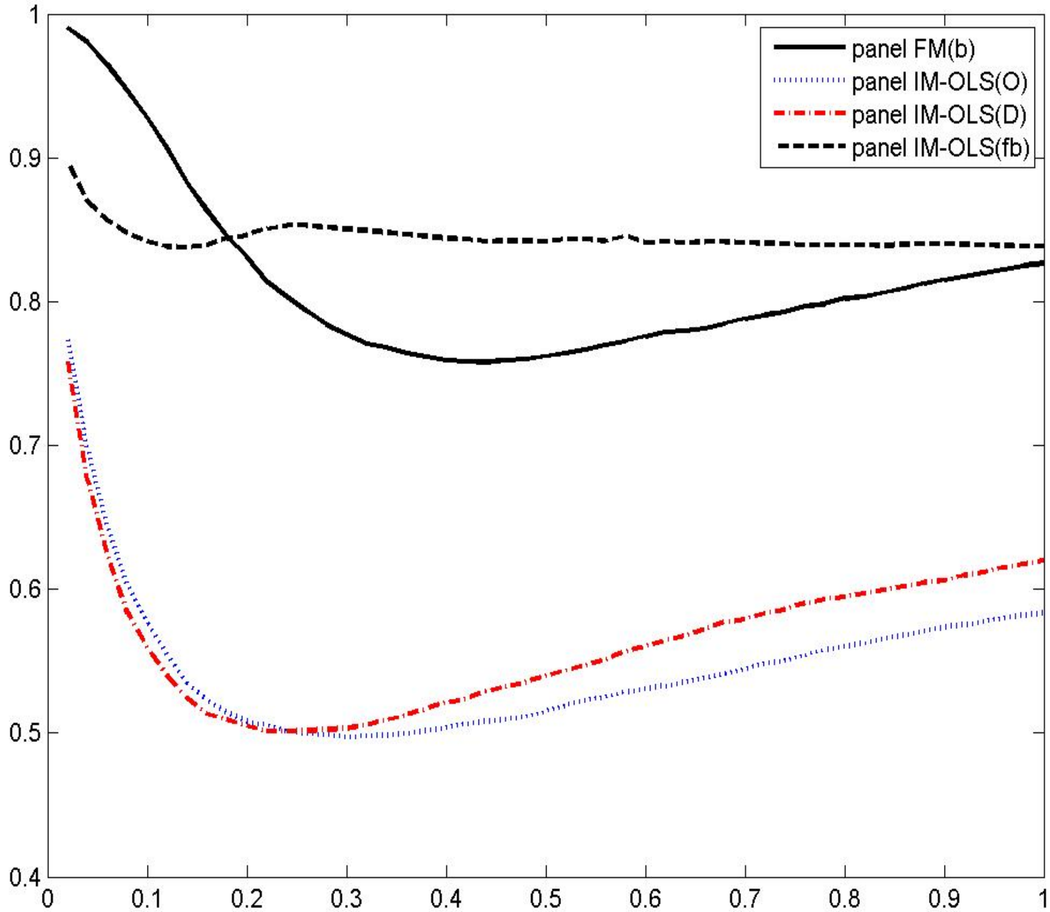


Figure 1.5: Empirical null rejections, t-test, $N = 10$, $T = 50$, $\rho_1 = \rho_2 = 0.9$, Bartlett kernel

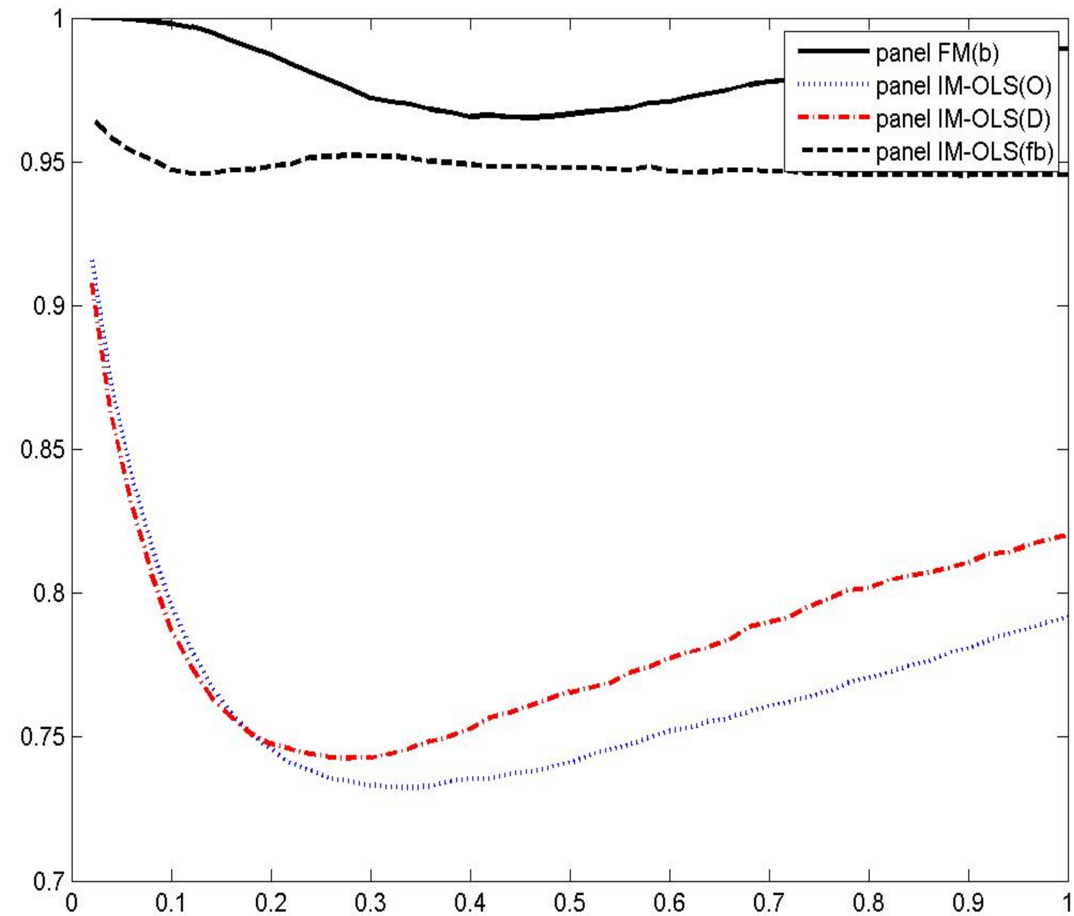


Figure 1.6: Empirical null rejections, t-test, $N = 25$, $T = 50$, $\rho_1 = \rho_2 = 0.9$, Bartlett kernel

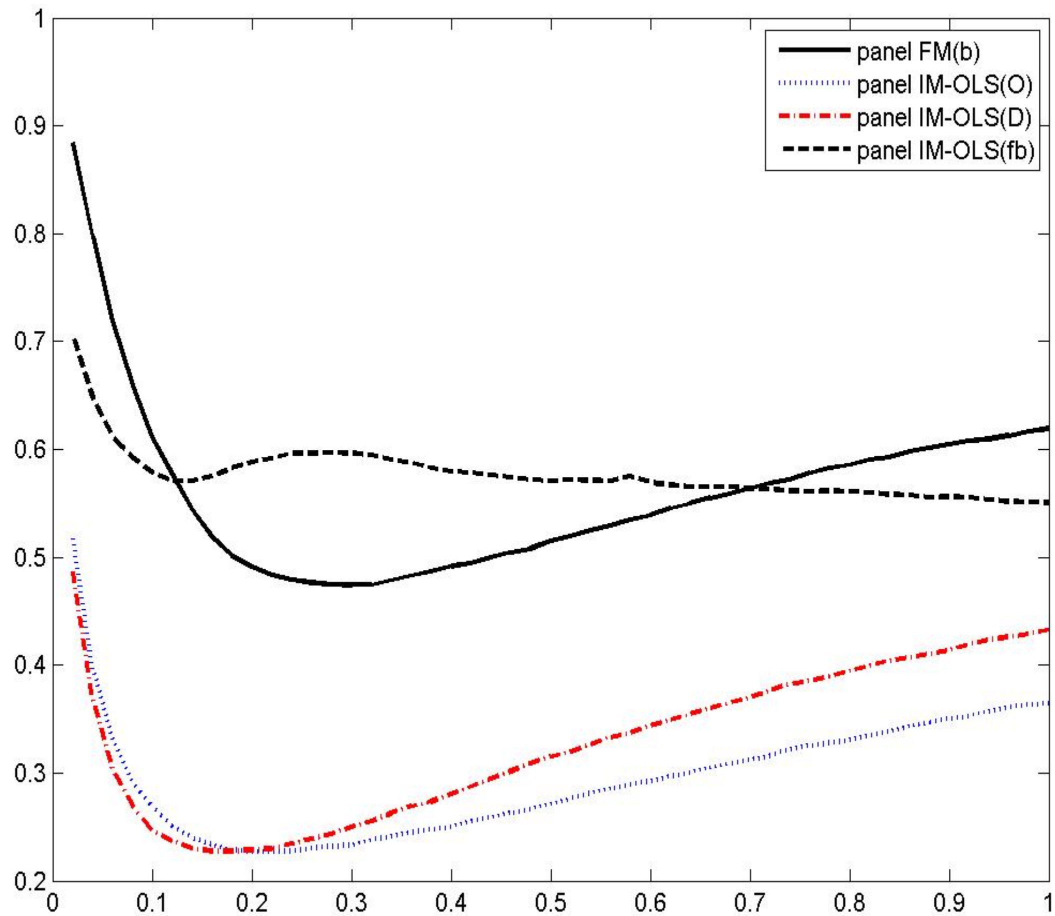


Figure 1.7: Empirical null rejections, t-test, $N = 5$, $T = 100$, $\rho_1 = \rho_2 = 0.9$, Bartlett kernel

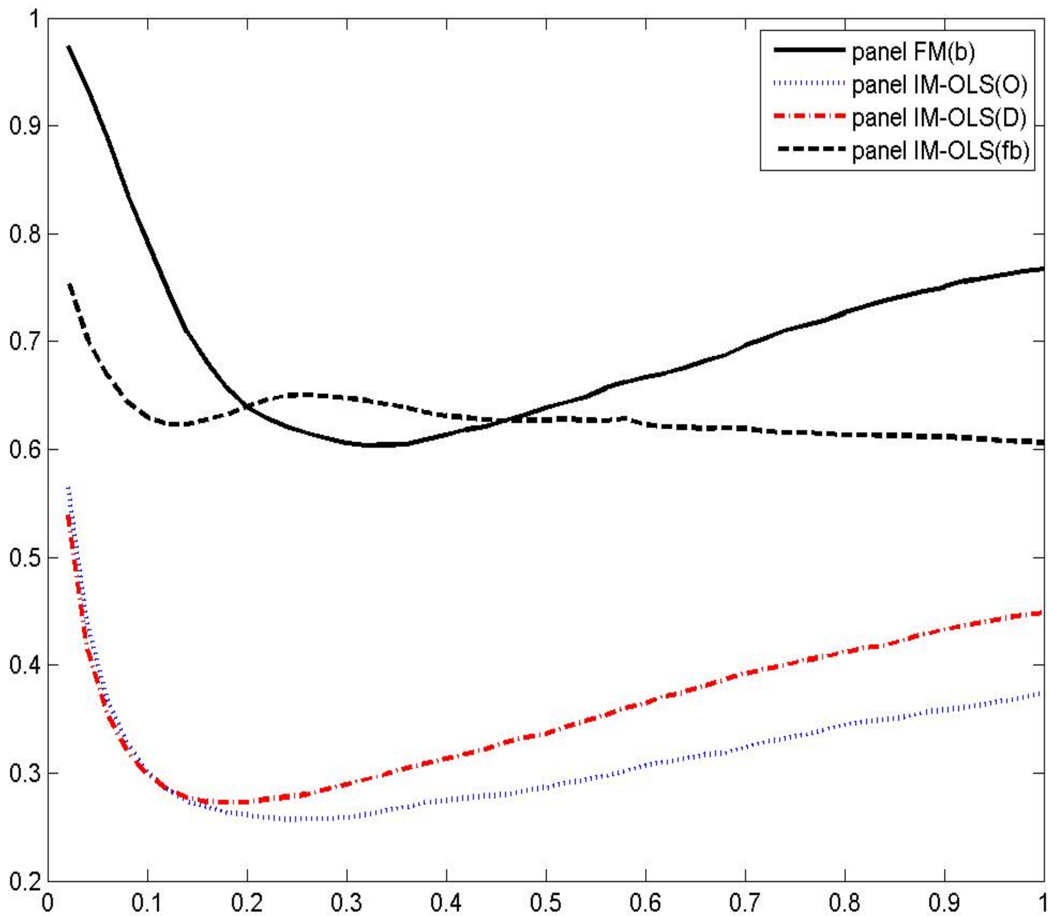


Figure 1.8: Empirical null rejections, t-test, $N = 10$, $T = 100$, $\rho_1 = \rho_2 = 0.9$, Bartlett kernel

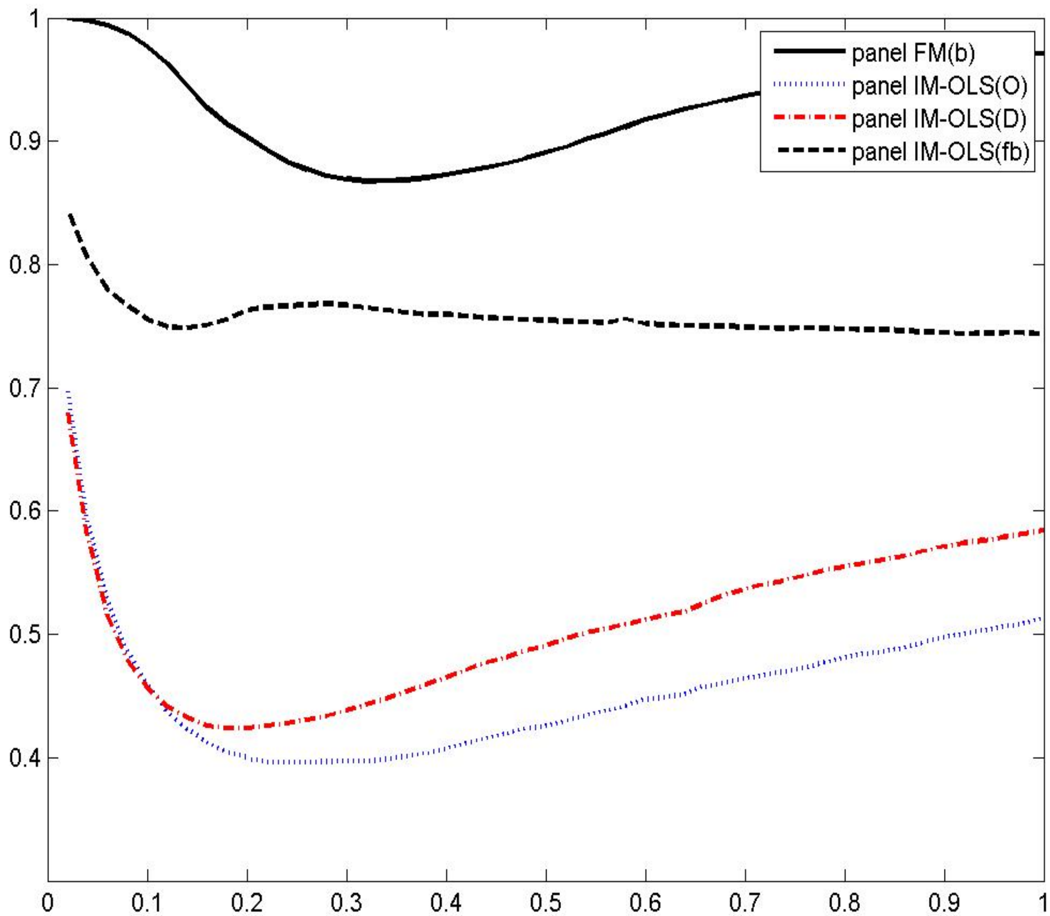


Figure 1.9: Empirical null rejections, t-test, $N = 25$, $T = 100$, $\rho_1 = \rho_2 = 0.9$, Bartlett kernel

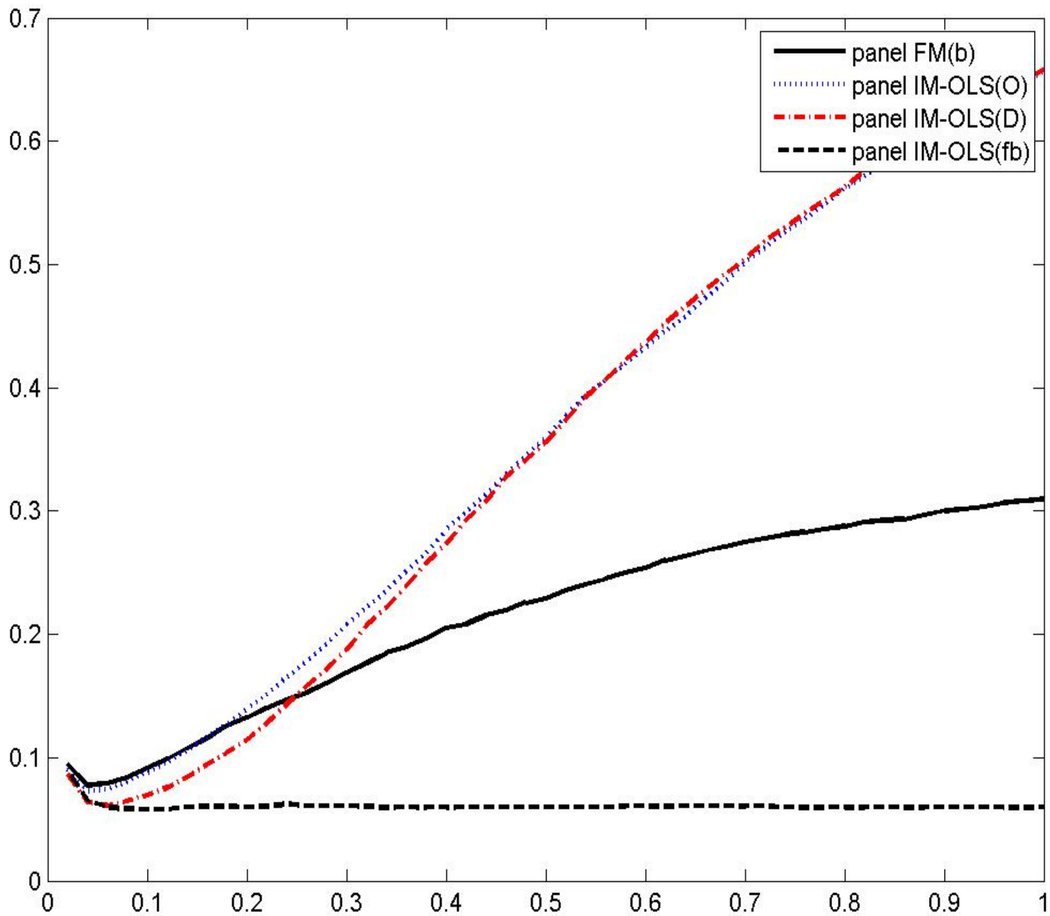


Figure 1.10: Empirical null rejections, t-test, $N = 5$, $T = 100$, $\rho_1 = \rho_2 = 0.3$, QS kernel

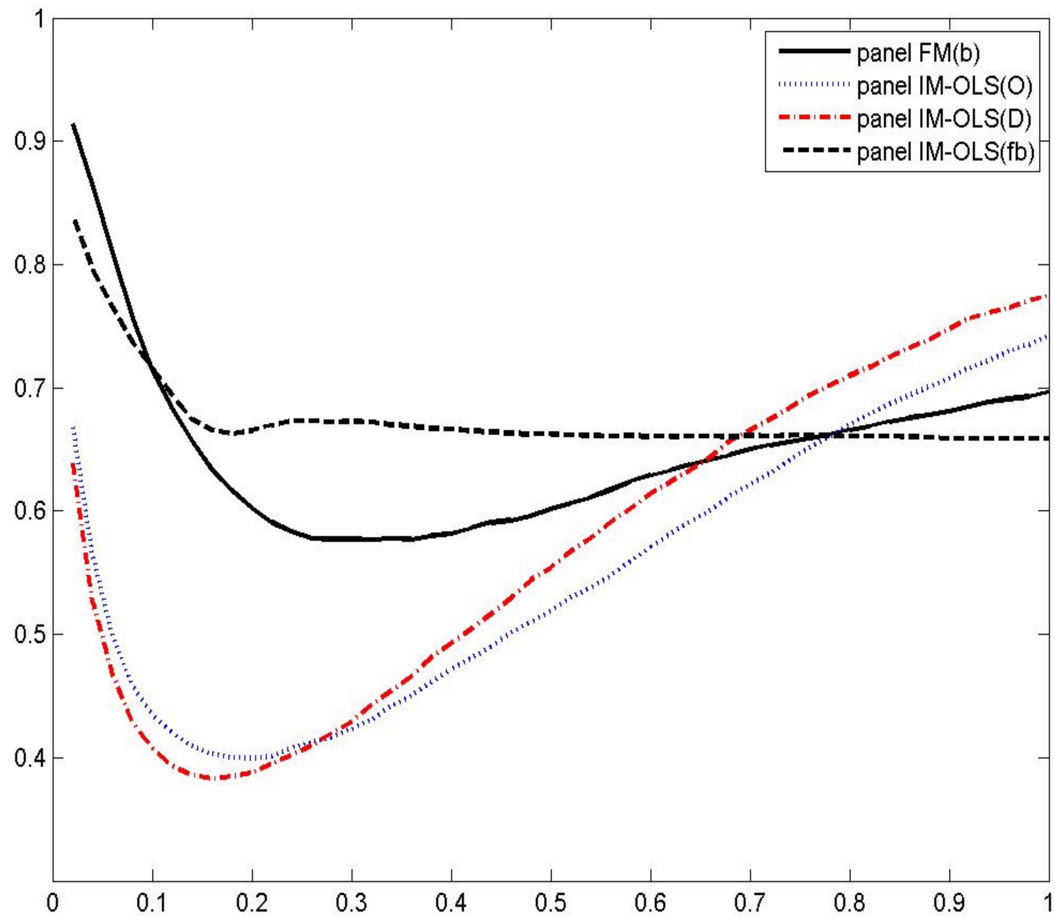


Figure 1.11: Empirical null rejections, t-test, $N = 5$, $T = 50$, $\rho_1 = \rho_2 = 0.9$, QS kernel

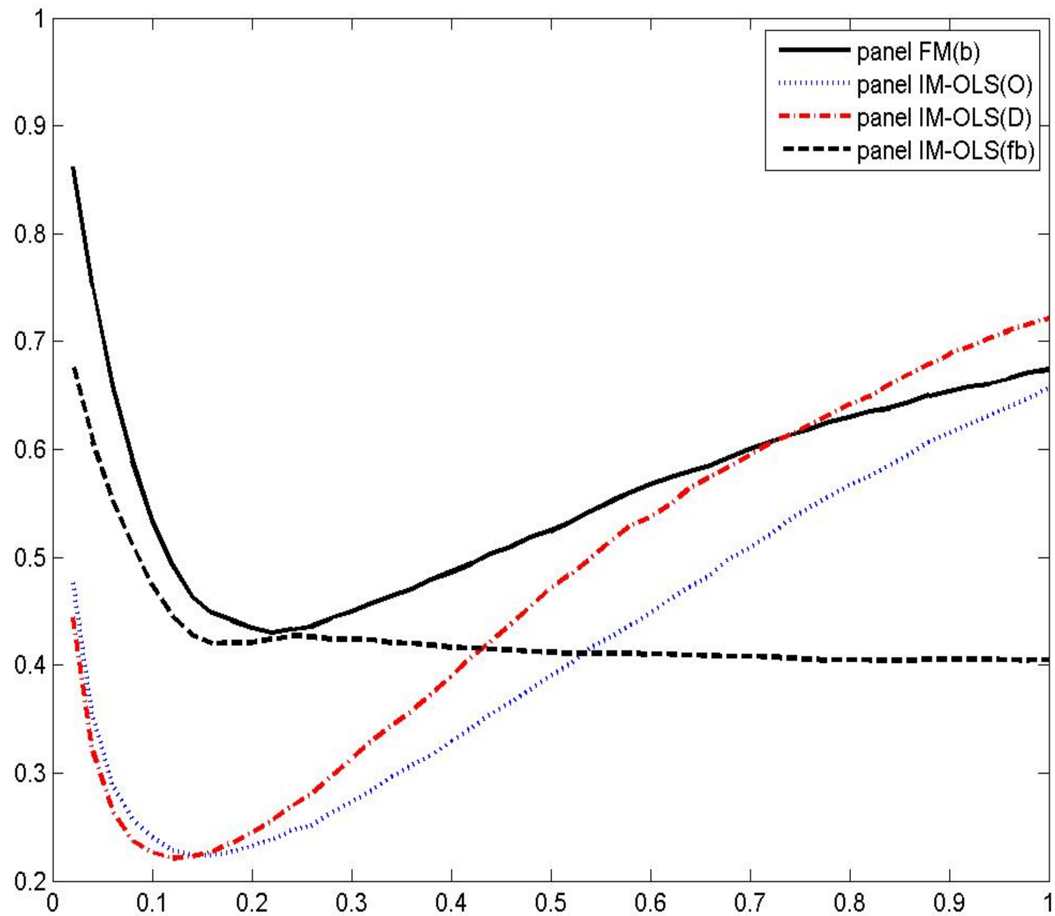


Figure 1.12: Empirical null rejections, t-test, $N = 5$, $T = 100$, $\rho_1 = \rho_2 = 0.9$, QS kernel

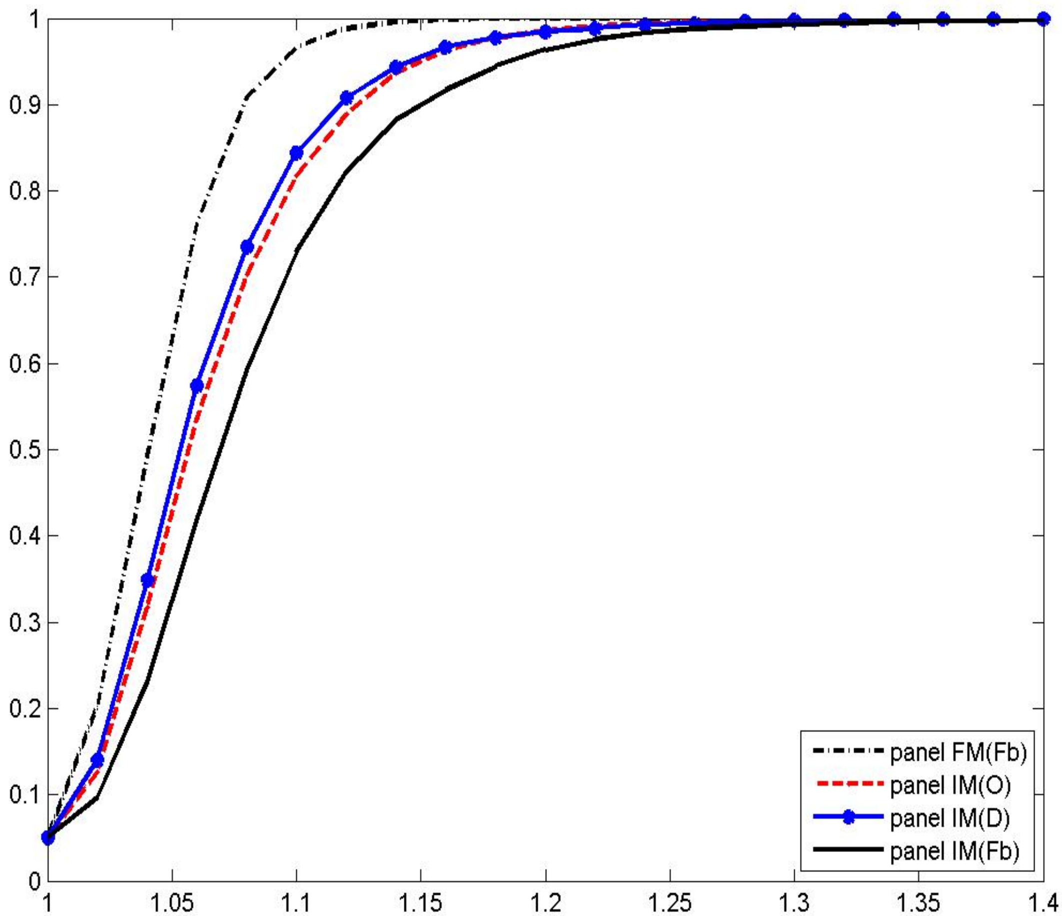


Figure 1.13: Size adjusted power, Wald test, $N = 5$, $T = 50$, $\rho_1 = \rho_2 = 0.6$, $b = 0.3$, QS kernel

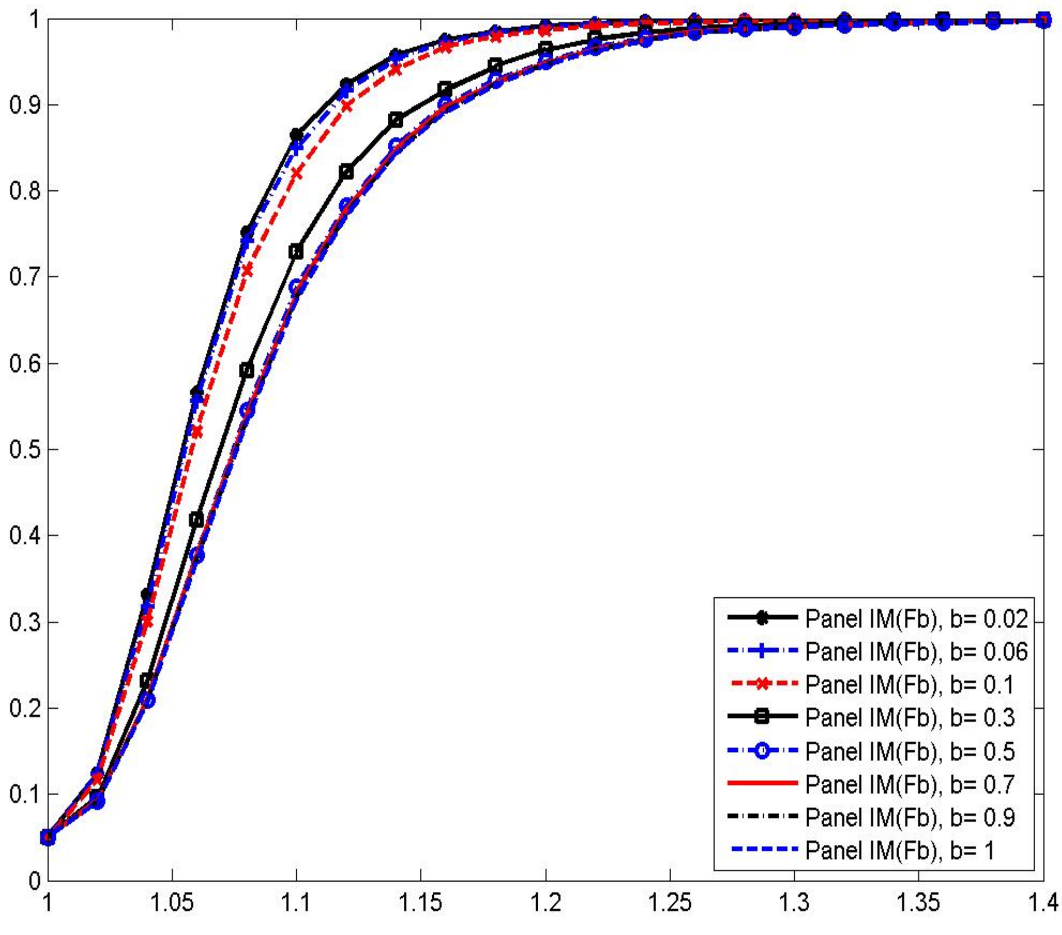


Figure 1.14: Size adjusted power of panel IM, Wald test, $N = 5$, $T = 50$, $\rho_1 = \rho_2 = 0.6$, QS kernel

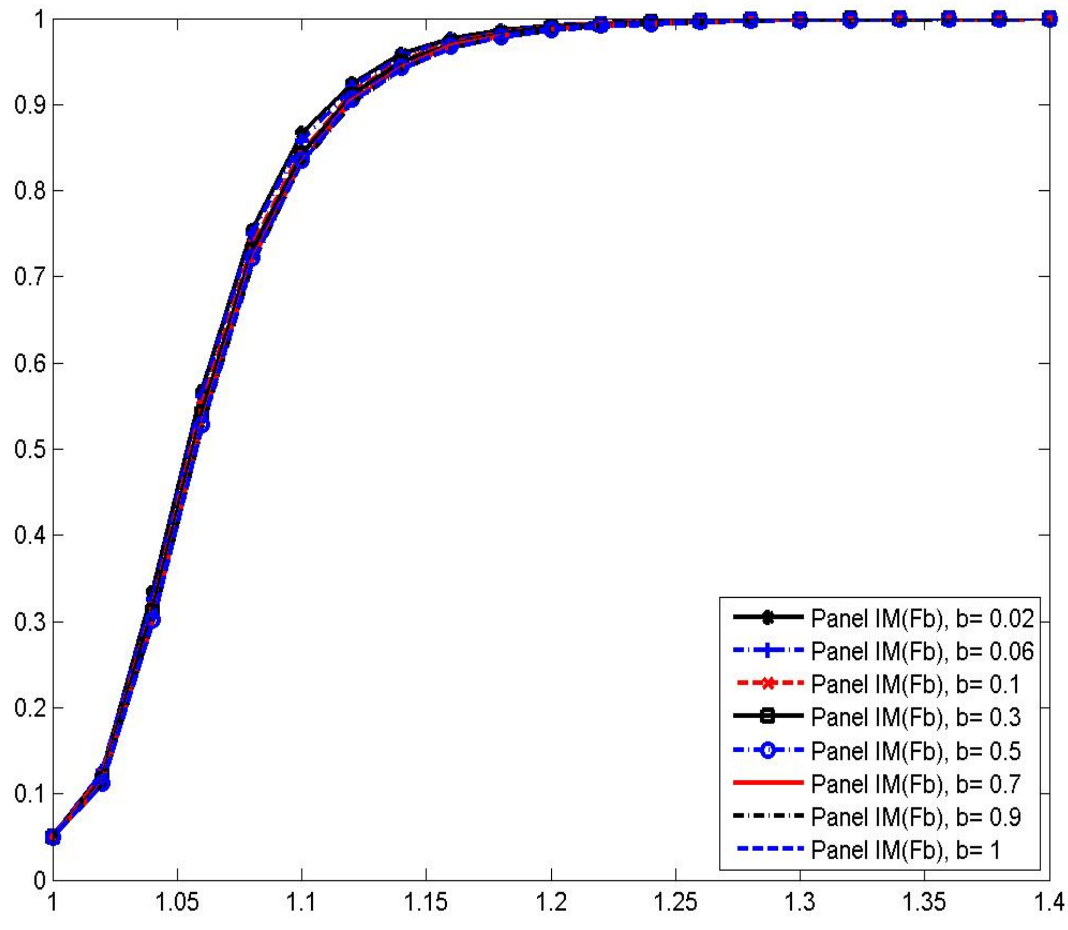


Figure 1.15: Size adjusted power of panel IM, Wald test, $N = 5$, $T = 50$, $\rho_1 = \rho_2 = 0.6$, Bartlett kernel

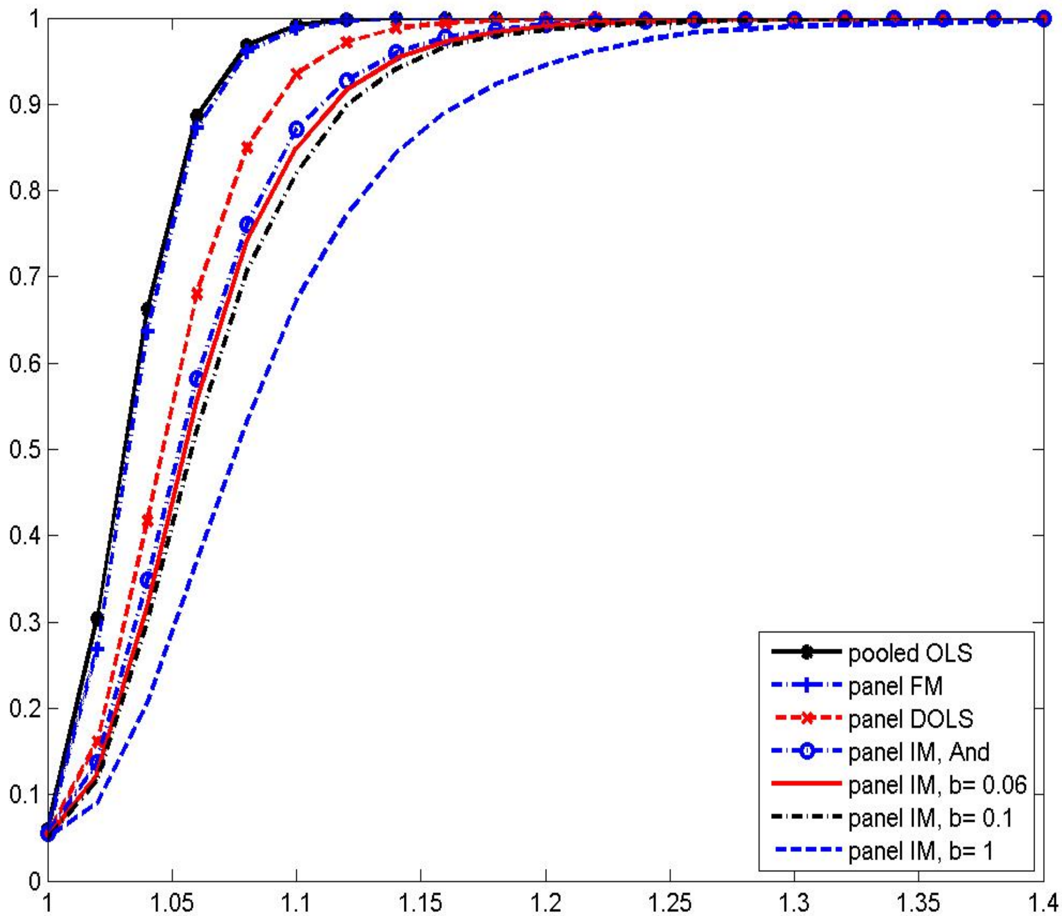


Figure 1.16: Size adjusted power, Wald test, $N = 5$, $T = 50$, $\rho_1 = \rho_2 = 0.6$, QS kernel

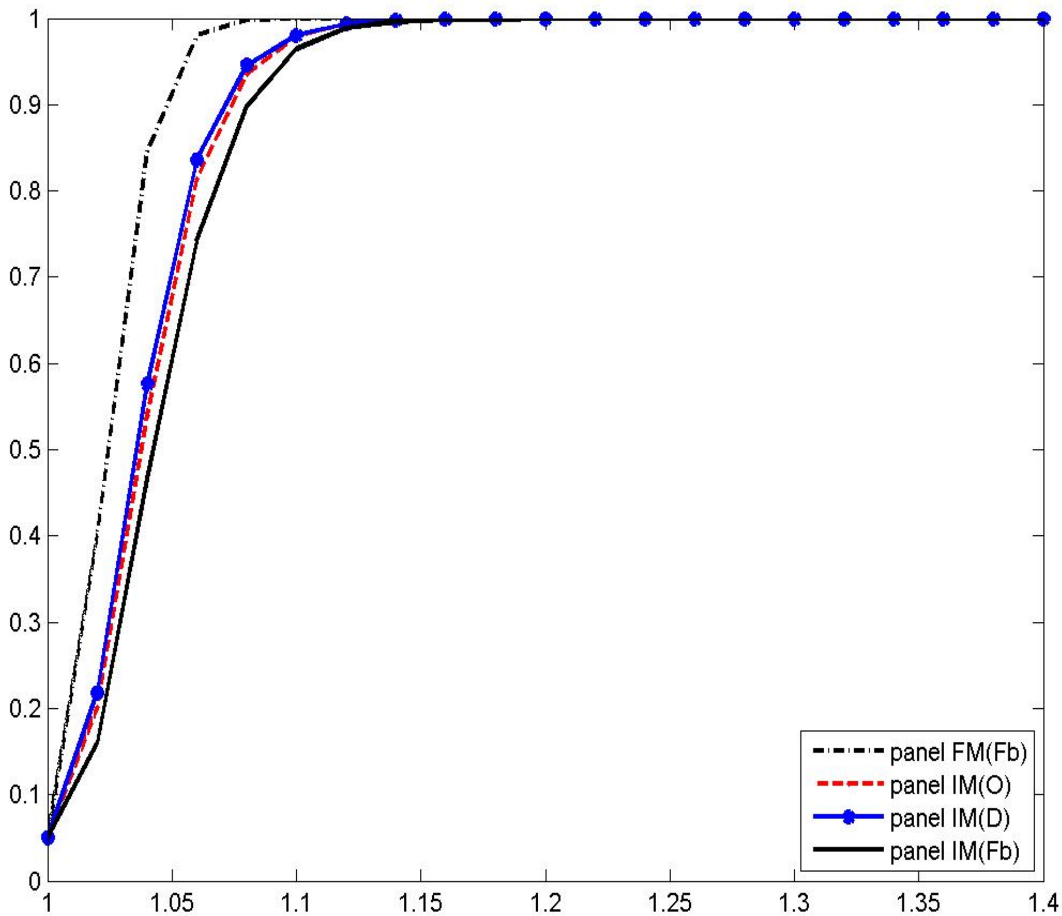


Figure 1.17: Size adjusted power, Wald test, $N = 10$, $T = 50$, $\rho_1 = \rho_2 = 0.6$, $b = 0.3$, QS kernel

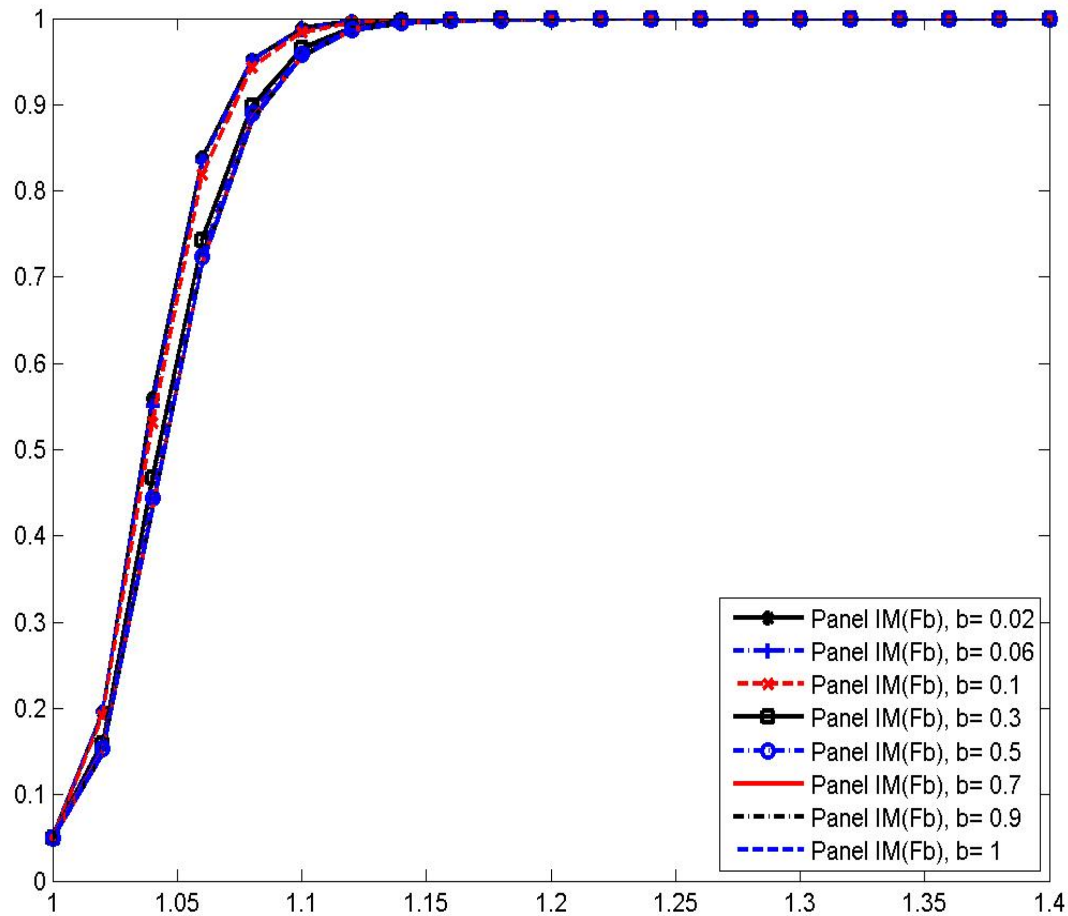


Figure 1.18: Size adjusted power of panel IM, Wald test, $N = 10$, $T = 50$, $\rho_1 = \rho_2 = 0.6$, QS kernel

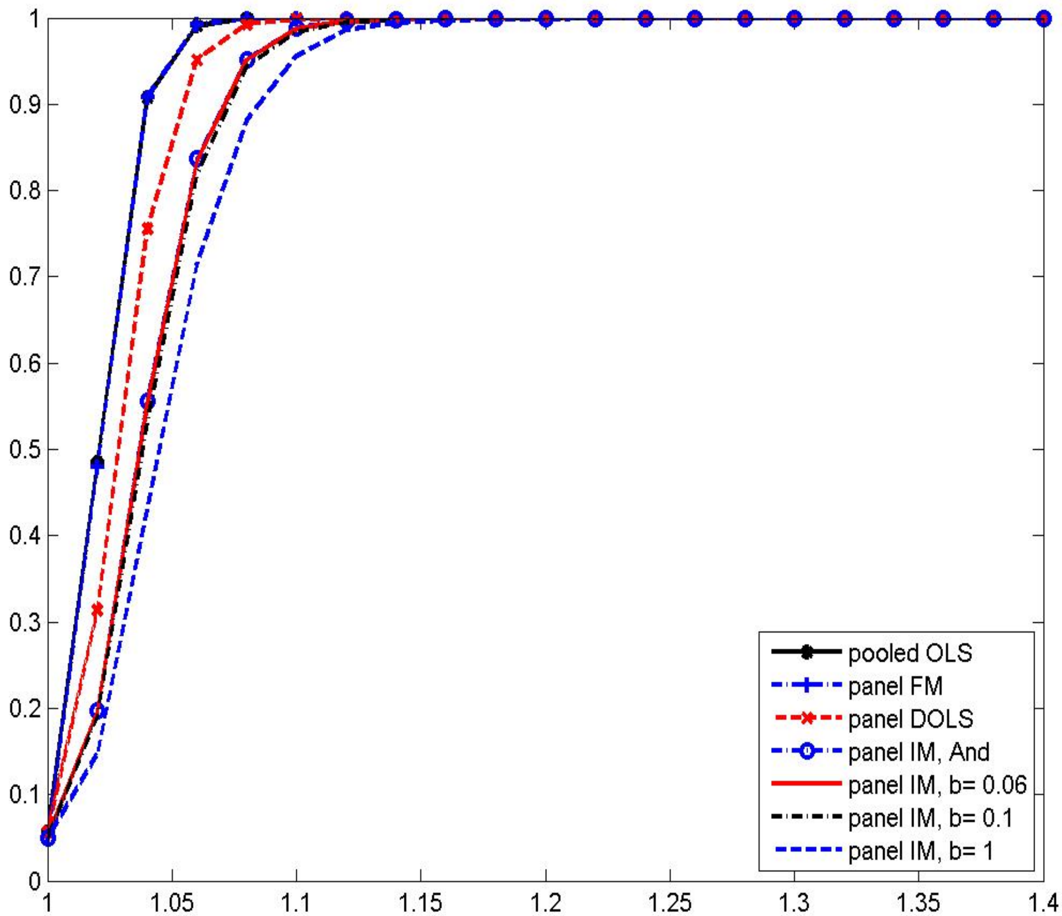


Figure 1.19: Size adjusted power, Wald test, $N = 10$, $T = 50$, $\rho_1 = \rho_2 = 0.6$, QS kernel

## Proof of Theorem 1

By Assumption 2 above, we can define stacked innovation vector

$$\eta_t = \begin{bmatrix} u_{1t}, & \cdots & u_{Nt}, & v'_{1t}, & \cdots & v'_{Nt} \end{bmatrix}',$$

which dimension is $(N + Nk) \times 1$, and assume that

$$T^{-1/2} \sum_{t=1}^{[rT]} \eta_t \Rightarrow \Omega_\eta^{1/2} W(r) = \begin{bmatrix} B_u(r) \\ B_v(r) \end{bmatrix} = \begin{bmatrix} \Lambda_{11} W_1(r) + \Lambda_{12} W_2(r) \\ \Lambda_{22} W_2(r) \end{bmatrix},$$

where

$$\Omega_\eta^{1/2} = \begin{bmatrix} \Lambda_{11} & \Lambda_{12} \\ 0_{Nk \times N} & \Lambda_{22} \end{bmatrix}$$
$$W(r) = \begin{bmatrix} W_1(r) \\ W_2(r) \end{bmatrix}.$$

The dimension of the above matrix are as follows: $\Lambda_{11}$ is $N \times N$, $\Lambda_{12}$ is $N \times Nk$, $0_{Nk \times N}$ is $Nk \times N$ zero matrix, $\Lambda_{22}$ is $Nk \times Nk$; $W_1(r) = \begin{bmatrix} w_{u,1}(r), & \cdots, & w_{u,N}(r) \end{bmatrix}'$ is $N \times 1$ vector, $W_2(r) = \begin{bmatrix} W_{v,1}(r)', & \cdots, & W_{v,N}(r)' \end{bmatrix}'$ is $Nk \times 1$ vector.

Long run variance of $\eta_t$ is:

$$\Omega_\eta = \Omega_\eta^{1/2} \left( \Omega_\eta^{1/2} \right)' = \sum_{j=-\infty}^{\infty} E\left( \eta_t \eta'_{t-j} \right).$$

Also we have:

$$\Omega_\eta = \begin{bmatrix} \Omega_{11}^\eta & \Omega_{12}^\eta \\ \Omega_{21}^\eta & \Omega_{22}^\eta \end{bmatrix} = \begin{bmatrix} \Lambda_{11} & \Lambda_{12} \\ 0_{Nk \times N} & \Lambda_{22} \end{bmatrix} \begin{bmatrix} \Lambda'_{11} & 0_{N \times Nk} \\ \Lambda'_{12} & \Lambda'_{22} \end{bmatrix} = \begin{bmatrix} \Lambda_{11}\Lambda'_{11} + \Lambda_{12}\Lambda'_{12} & \Lambda_{12}\Lambda'_{22} \\ \Lambda_{22}\Lambda'_{12} & \Lambda_{22}\Lambda'_{22} \end{bmatrix},$$

where $\Omega_{11}^\eta = \Lambda_{11}\Lambda'_{11} + \Lambda_{12}\Lambda'_{12}$ is long run variance of $u_t = \begin{bmatrix} u_{1t} & \cdots & u_{Nt} \end{bmatrix}'$, $\Omega_{22}^\eta = \Lambda_{22}\Lambda'_{22}$ is long run variance of $v_t = \begin{bmatrix} v'_{1t} & \cdots & v'_{Nt} \end{bmatrix}'$, $\Omega_{12}^\eta = \left( \Omega_{21}^\eta \right)' = \Lambda_{12}\Lambda'_{22}$ is long run covariance of $u_t$ and $v_t$. From assumption 1, we know that $\Lambda_{11} = \begin{bmatrix} I_N \otimes \sigma_{u \cdot v} \end{bmatrix}_{N \times N}$ is $N \times N$ diagonal matrix, $\Lambda_{12} = \begin{bmatrix} I_N \otimes \lambda_{uv} \end{bmatrix}_{N \times Nk}$ is $N \times Nk$ diagonal matrix, $\Lambda_{22} = \begin{bmatrix} I_N \otimes \Omega_{vv}^{1/2} \end{bmatrix}_{Nk \times Nk}$ is $Nk \times Nk$ diagonal matrix, where $\sigma_{u \cdot v}$ is scalar, $\lambda_{uv}$ is $1 \times k$ vector, and $\Omega_{vv}^{1/2}$ is $k \times k$ matrix.

Using diagonal scaling matrix

$$A_{1T} = \begin{bmatrix} T^{3/2} I_k & & 0 \\ & T^{1/2} I_k & \\ 0 & & T^{1/2} I_N \otimes G_D \end{bmatrix},$$

then we have:

$$A_{PIM}^{-1}\left(\tilde{\theta} - \theta\right) = T^{-1/2} A_{1T}\left(\tilde{\theta} - \theta\right) = \begin{bmatrix} T\left(\tilde{\beta} - \beta\right) \\ \left(\tilde{\gamma} - \Omega_{vv}^{-1}\Omega_{vu}\right) \\ G_D\left(\tilde{\delta}_1 - \delta_1\right) \\ \vdots \\ G_D\left(\tilde{\delta}_N - \delta_N\right) \end{bmatrix}$$

$$= \left(T^{-1}\sum_{t=1}^{T}\sum_{i=1}^{N} A_{1T}^{-1} q_{it} q_{it}' A_{1T}^{-1}\right)^{-1} \left(T^{-1}\sum_{t=1}^{T}\sum_{i=1}^{N} A_{1T}^{-1} q_{it} T^{-1/2}\left(S_{it}^{u} - x_{it}'\gamma\right)\right).$$

By our assumptions,

$$A_{1T}^{-1} q_{it} = \begin{bmatrix} T^{-\frac{3}{2}} S_{i[rT]}^{x} \\ T^{-\frac{1}{2}} x_{i[rT]} \\ 0_{p\times 1} \\ \vdots \\ T^{-\frac{1}{2}} G_D^{-1} S_t^{D} \\ \vdots \\ 0_{p\times 1} \end{bmatrix} \Rightarrow \begin{bmatrix} \int_0^r B_{v,i}(s)ds \\ B_{v,i}(r) \\ 0_{p\times 1} \\ \vdots \\ \int_0^r D(s)ds \\ \vdots \\ 0_{p\times 1} \end{bmatrix} = \begin{bmatrix} \Omega_{vv}^{1/2}\int_0^r W_{v,i}(s)ds \\ \Omega_{vv}^{1/2} W_{v,i}(r) \\ 0_{p\times 1} \\ \vdots \\ \int_0^r D(s)ds \\ \vdots \\ 0_{p\times 1} \end{bmatrix} = \Pi g_{1,i}(r),$$

and it follows that

$$T^{-1}\sum_{t=1}^{T}\sum_{i=1}^{N} A_{1T}^{-1} q_{it} q_{it}' A_{1T}^{-1} \Rightarrow \int_0^1 \sum_{i=1}^{N} \Pi g_{1,i}(r) g_{1,i}(r)' \Pi' dr.$$

Also, by previous assumptions,

$$
\begin{aligned}
T^{-1/2}\left(S_{it}^{u}-x_{it}'\gamma\right) &= T^{-1/2}S_{it}^{u}-T^{-1/2}x_{it}'\gamma \\
&\Rightarrow B_{u,i}(r)-B_{v,i}'(r)\gamma \\
&= \sigma_{u\cdot v}w_{u,i}(r)+\lambda_{uv}W_{v,i}(r)-\left[\Omega_{vv}^{1/2}W_{v,i}(r)\right]'\gamma \\
&= \sigma_{u\cdot v}w_{u,i}(r)-W_{v,i}'(r)\left(\Omega_{vv}^{1/2}\right)'\left[\gamma-\left(\left(\Omega_{vv}^{1/2}\right)'\right)^{-1}\lambda_{uv}'\right] \\
&= \sigma_{u\cdot v}w_{u,i}(r)-W_{v,i}'(r)\left(\Omega_{vv}^{1/2}\right)'\left[\gamma-\left(\left(\Omega_{vv}^{1/2}\right)'\right)^{-1}\left(\Omega_{vv}^{-1/2}\right)\Omega_{vu}\right] \\
&= \sigma_{u\cdot v}w_{u,i}(r)-W_{v,i}'(r)\left(\Omega_{vv}^{1/2}\right)'\left[\gamma-\Omega_{vv}^{-1}\Omega_{vu}\right]
\end{aligned}
$$

so if $\gamma=\Omega_{vv}^{-1}\Omega_{vu}$, then

$$
T^{-1/2}\left(S_{it}^{u}-x_{it}'\gamma\right)\Rightarrow\sigma_{u\cdot v}w_{u,i}(r).
$$

Combining above results, we have:

$$
\begin{bmatrix}
T\left(\tilde{\beta}-\beta\right) \\
\left(\tilde{\gamma}-\Omega_{vv}^{-1}\Omega_{vu}\right) \\
G_D\left(\tilde{\delta}_1-\delta_1\right) \\
\vdots \\
G_D\left(\tilde{\delta}_N-\delta_N\right)
\end{bmatrix}
= A_{PIM}^{-1}\left(\tilde{\theta}-\theta\right)
$$

$$
\Rightarrow\left(\int_0^1\sum_{i=1}^{N}\Pi g_{1,i}(s)g_{1,i}(s)'\Pi'dr\right)^{-1}\left(\int_0^1\sum_{i=1}^{N}\Pi g_{1,i}(s)\sigma_{u\cdot v}w_{u,i}(s)ds\right)
$$

$$
=\sigma_{u\cdot v}\left(\Pi'\right)^{-1}\left(\int_0^1\sum_{i=1}^{N}g_{1,i}(s)g_{1,i}(s)'ds\right)^{-1}\left(\int_0^1\sum_{i=1}^{N}g_{1,i}(s)w_{u,i}(s)ds\right)
$$

$$
=\sigma_{u\cdot v}\left(\Pi'\right)^{-1}\left(\sum_{i=1}^{N}\int_0^1 g_{1,i}(s)g_{1,i}(s)'ds\right)^{-1}\left(\sum_{i=1}^{N}\int_0^1\left[G_{1,i}(1)-G_{1,i}(s)\right]dw_{u,i}(s)\right)=\Psi.
$$

For the sequential limit of $\left(\tilde{\beta}-\beta\right)$, we first let $T\to\infty$, then let $N\to\infty$, so we have

$$
\begin{aligned}
\sqrt{N}T\left(\tilde{\beta}-\beta\right) =& \sqrt{N}\left[I_k \quad 0_{k\times k} \quad 0_{k\times p} \quad \cdots \quad 0_{k\times p}\right] A_{PIM}^{-1}\left(\tilde{\theta}-\theta\right) \\
=& \left[I_k \quad 0_{k\times k} \quad 0_{k\times p} \quad \cdots \quad 0_{k\times p}\right]\left(\frac{1}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N} A_{1T}^{-1} q_{it} q_{it}' A_{1T}^{-1}\right)^{-1} \times \\
& \left(\frac{1}{\sqrt{N}T}\sum_{t=1}^{T}\sum_{i=1}^{N} A_{1T}^{-1} q_{it}\frac{1}{\sqrt{T}}\left(S_{it}^{u}-x_{it}'\gamma\right)\right) \\
=& \left[I_k \quad 0_{k\times k} \quad 0_{k\times p} \quad \cdots \quad 0_{k\times p}\right]\left(\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T}\sum_{t=1}^{T} A_{1T}^{-1} q_{it} q_{it}' A_{1T}^{-1}\right)^{-1} \times \\
& \left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{1}{T}\sum_{t=1}^{T} A_{1T}^{-1} q_{it}\frac{1}{\sqrt{T}}\left(S_{it}^{u}-x_{it}'\gamma\right)\right) \\
\overset{T\to\infty}{\Longrightarrow}& \sigma_{u\cdot v}\left[I_k \quad 0_{k\times k} \quad 0_{k\times p} \quad \cdots \quad 0_{k\times p}\right]\left(\Pi'\right)^{-1} \times \\
& \left(\frac{1}{N}\sum_{i=1}^{N}\int_0^1 g_{1,i}(r) g_{1,i}(r)' dr\right)^{-1}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\int_0^1 [G_{1,i}(1)-G_{1,i}(r)] dw_{u,i}(r)\right) \\
\overset{N\to\infty}{\Longrightarrow}& \Phi
\end{aligned}
$$

In order to get the distribution of $\Phi$, we need to know the limit of the upper and left $k\times k$ block of $\frac{1}{N}\sum_{i=1}^{N}\int_0^1 g_{1,i}(r)g_{1,i}(r)'dr$ and the distribution of the upper $k\times 1$ block of $\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\int_0^1[G_{1,i}(1)-G_{1,i}(r)]dw_{u,i}(r)$ as $N\to\infty$.

First, consider the limit of the upper and left $k\times k$ block of $\frac{1}{N}\sum_{i=1}^{N}\int_0^1 g_{1,i}(r)g_{1,i}(r)'dr$.

Note that, $\frac{1}{N}\sum_{i=1}^{N}\int_0^1 g_{1,i}(r)g_{1,i}(r)'dr = \int_0^1 \frac{1}{N}\sum_{i=1}^{N}\left[g_{1,i}(r)g_{1,i}(r)'\right]dr$. The related components for the sequential limits is the integral of the upper and left $k\times k$ block of the limit of $\frac{1}{N}\sum_{i=1}^{N}\left[g_{1,i}(r)g_{1,i}(r)'\right]$, which is given by

$$
\begin{aligned}
&\frac{1}{N}\sum_{i=1}^{N}\int_0^r W_{v,i}(s)ds\int_0^r W_{v,i}(s)'ds\\
=&\frac{1}{N}\sum_{i=1}^{N}\int_0^r \left(\int_0^r W_{v,i}(u)du\right) W_{v,i}(s)'ds\\
=&\frac{1}{N}\sum_{i=1}^{N}\int_0^r \int_0^r W_{v,i}(u)W_{v,i}(s)'duds\\
=&\frac{1}{N}\sum_{i=1}^{N}\int_0^r \int_0^s W_{v,i}(u)W_{v,i}(s)'duds+\\
&\frac{1}{N}\sum_{i=1}^{N}\int_0^r \int_s^r W_{v,i}(u)W_{v,i}(s)'duds\\
\rightarrow&\int_0^r\int_0^s ududs\cdot I_k+\int_0^r\int_s^r sduds\cdot I_k\\
=&\frac{r^3}{3}I_k=A_1(r)
\end{aligned}
$$

Therefore, $\int_0^1 A_1(r)dr=(1/12)I_k$.

Second, consider the distribution of the upper $k\times 1$ block of

$$
\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\int_0^1[G_{1,i}(1)-G_{1,i}(r)]dw_{u,i}(r).
$$

It has an asymptotic normal distribution, with zero mean, when conditional on $G_{1,i}(r)$ for all $i=1,2,\ldots,N$. So we only need to find its asymptotic variance. Also, recall that the units are cross-sectional independent, so we have

$$
\begin{aligned}
&var\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\int_0^1[G_{1,i}(1)-G_{1,i}(r)]dw_{u,i}(r)\right)\\
=&\int_0^1\frac{1}{N}\sum_{i=1}^{N}[G_{1,i}(1)-G_{1,i}(r)][G_{1,i}(1)-G_{1,i}(r)]'dr
\end{aligned}
$$

as $N$ fixed.

Note that the variance of the upper $k \times 1$ block of

$$\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \int_0^1 [G_{1,i}(1) - G_{1,i}(r)] dw_{u,i}(r)$$

is just the upper and left $k \times k$ block of

$$\int_0^1 \frac{1}{N} \sum_{i=1}^{N} [G_{1,i}(1) - G_{1,i}(r)][G_{1,i}(1) - G_{1,i}(r)]' dr,$$

which is given by

$$\begin{aligned}
&\frac{1}{N} \sum_{i=1}^{N} \int_r^1 \int_0^s W_{v,i}(u) du ds \int_r^1 \int_0^s W_{v,i}(u)' du ds \\
=&\frac{1}{N} \sum_{i=1}^{N} \int_r^1 \int_0^s \int_r^1 \int_0^s W_{v,i}(v) W_{v,i}(u)' dv dt du ds \\
=&\frac{1}{N} \sum_{i=1}^{N} \int_r^1 \int_0^s \int_r^1 \int_0^u W_{v,i}(v) W_{v,i}(u)' dv dt du ds + \\
&\frac{1}{N} \sum_{i=1}^{N} \int_r^1 \int_0^s \int_r^1 \int_u^s W_{v,i}(v) W_{v,i}(u)' dv dt du ds \\
\to &\left( \int_r^1 \int_0^s \int_r^1 \int_0^u v dv dt du ds \right) I_k + \\
&\left( \int_r^1 \int_0^s \int_r^1 \int_u^s u dv dt du ds \right) I_k \\
=&\frac{1}{12} (1 - r) \left( 1 - r^4 \right) I_k = A_2(r)
\end{aligned}$$

So, we have $\int_0^1 A_2(r) dr = (7/180) I_k$.

Using above notations, the sequential asymptotic distribution $\Phi$ is given by

$$\Phi \sim N \left( 0, \sigma_{u \cdot v}^2 \left[ \left( \Omega_{vv}^{1/2} \right)' \right]^{-1} \left( \int_0^1 A_1(r) dr \right)^{-1} \left( \int_0^1 A_2(r) dr \right) \left( \int_0^1 A_1(r) dr \right)^{-1} \left( \Omega_{vv}^{1/2} \right)^{-1} \right).$$

We denote its variance as

$$\begin{aligned}
V_{seq}^{\beta} &= \sigma_{u \cdot v}^2 \left[ \left( \Omega_{vv}^{1/2} \right)' \right]^{-1} \left( \int_0^1 A_1(r) dr \right)^{-1} \left( \int_0^1 A_2(r) dr \right) \left( \int_0^1 A_1(r) dr \right)^{-1} \left( \Omega_{vv}^{1/2} \right)^{-1} \\
&= 5.6 \cdot \sigma_{u \cdot v}^2 \Omega_{vv}^{-1}.
\end{aligned}$$

## Proof of Theorem 2

In Theorem 2, the Wald statistics considered was $\breve{W}$, with $\breve{W} \in \{\hat{W}, \tilde{W}, \tilde{W}^*\}$. Those statistics only differ with respect to the used estimator of the long run variance parameter, $\breve{\sigma}^2_{u \cdot v} \in \{\hat{\sigma}^2_{u \cdot v}, \tilde{\sigma}^2_{u \cdot v}, \tilde{\sigma}^{2*}_{u \cdot v}\}$. As in the proof of Theorem 2, $\tilde{\theta}$ represents the vector of panel IM-OLS estimators $\left(\tilde{\delta}', \ \tilde{\beta}', \ \tilde{\gamma}'\right)'$, and $\theta$ denotes the vector $\left(\delta', \ \beta', \ \Omega'_{vu}\Omega^{-1}_{vv}\right)'$. The estimator for $V_{PIM}$ is given by

$$
\begin{aligned}
\breve{V}_{PIM} &= \breve{\sigma}^2_{u \cdot v} \left( T^{-2} \sum_{t=1}^{T} \sum_{i=1}^{N} A_{PIM} q_{it} q'_{it} A_{PIM} \right)^{-1} \times \\
& \left( T^{-4} \sum_{t=1}^{T} \sum_{i=1}^{N} A_{PIM} \left[ S^q_{iT} - S^q_{i,t-1} \right] \left[ S^q_{iT} - S^q_{i,t-1} \right]' A_{PIM} \right) \times \\
& \left( T^{-2} \sum_{t=1}^{T} \sum_{i=1}^{N} A_{PIM} q_{it} q'_{it} A_{PIM} \right)^{-1} \\
&= \breve{\sigma}^2_{u \cdot v} \cdot \breve{V},
\end{aligned}
$$

where $\breve{V}$ is the estimator for

$$
\begin{aligned}
V &= (\Pi')^{-1} \left( \int_0^1 \sum_{i=1}^{N} g_{1,i}(s) g_{1,i}(s)' ds \right)^{-1} \times \\
& \left( \int_0^1 \sum_{i=1}^{N} [G_{1,i}(1) - G_{1,i}(s)][G_{1,i}(1) - G_{1,i}(s)]' ds \right) \times \\
& \left( \int_0^1 \sum_{i=1}^{N} g_{1,i}(s) g_{1,i}(s)' ds \right)^{-1} \Pi^{-1}.
\end{aligned}
$$

Under the null hypothesis the Wald statistics and t statistics can be written as

$$
\begin{aligned}
\breve{W} &= \left(R\tilde{\theta} - r\right)' \left[RA_{PIM}\breve{V}_{PIM}A_{PIM}R'\right]^{-1} \left(R\tilde{\theta} - r\right) \\
&= \left(R\left(\tilde{\theta} - \theta\right)\right)' \left[RA_{PIM}\breve{V}_{PIM}A_{PIM}R'\right]^{-1} \left(R\left(\tilde{\theta} - \theta\right)\right) \\
&= \left(A_R^{-1}RA_{PIM}A_{PIM}^{-1}\left(\tilde{\theta} - \theta\right)\right)' \left[A_R^{-1}RA_{PIM}\breve{V}_{PIM}A_{PIM}R'\left(A_R^{-1}\right)'\right]^{-1} \times \\
&\quad \left(A_R^{-1}RA_{PIM}A_{PIM}^{-1}\left(\tilde{\theta} - \theta\right)\right),
\end{aligned}
$$

and

$$
\begin{aligned}
\breve{t} &= \left(R\tilde{\theta} - r\right) / \left(\sqrt{\left(RA_{PIM}\breve{V}_{PIM}A_{PIM}R'\right)}\right) \\
&= \left[A_R^{-1}RA_{PIM}A_{PIM}^{-1}\left(\tilde{\theta} - \theta\right)\right] / \left(\sqrt{\left(A_R^{-1}RA_{PIM}\breve{V}_{PIM}A_{PIM}R'\left(A_R^{-1}\right)'\right)}\right).
\end{aligned}
$$

Now, by assumption the restriction matrix fulfills

$$
\underset{T\to\infty}{lim} A_R^{-1}RA_{PIM} = R^*,
$$

and

$$
A_{PIM}^{-1}\left(\tilde{\theta} - \theta\right) \Rightarrow \Psi(V_{PIM})
$$

under the null hypothesis. Therefore, in case of consistent estimation of the conditional long run variance $\sigma_{u\cdot v}^2$ using $\hat{V}_{PIM}$ it follows that

$$
\begin{aligned}
\hat{W} &= \left(A_R^{-1}RA_{PIM}A_{PIM}^{-1}\left(\tilde{\theta} - \theta\right)\right)' \left[A_R^{-1}RA_{PIM}\hat{V}_{PIM}A_{PIM}R'\left(A_R^{-1}\right)'\right]^{-1} \\
&\quad \times \left(A_R^{-1}RA_{PIM}A_{PIM}^{-1}\left(\tilde{\theta} - \theta\right)\right) \\
&= \left(A_R^{-1}RA_{PIM}A_{PIM}^{-1}\left(\tilde{\theta} - \theta\right)\right)' \left[A_R^{-1}RA_{PIM}\sigma_{u\cdot v}^2\breve{V}A_{PIM}R'\left(A_R^{-1}\right)'\right]^{-1} \\
&\quad \times \left(A_R^{-1}RA_{PIM}A_{PIM}^{-1}\left(\tilde{\theta} - \theta\right)\right) \times \frac{\sigma_{u\cdot v}^2}{\hat{\sigma}_{u\cdot v}^2} \\
&\Rightarrow \left(R^*\Psi(V_{PIM})\right)' \left(R^*V_{PIM}R^{*\prime}\right)^{-1} \left(R^*\Psi(V_{PIM})\right) \\
&\sim \chi_q^2
\end{aligned}
$$

and for $q = 1$, we have

$$\hat{t} \Rightarrow \frac{R^{*}\Psi(V_{PIM})}{\sqrt{R^{*}V_{PIM}R^{*\prime}}} \sim Z.$$

Next, we consider the asymptotic behavior of the test statistic $\tilde{W}$ using $\tilde{\sigma}^2_{u\cdot v}$. From the construction of $\tilde{\sigma}^2_{u\cdot v}$, we know that it is an estimator based on $\Delta\tilde{S}^u_{it}$, which is the difference of $\tilde{S}^u_{it}$, where $\tilde{S}^u_{it} = S^y_{it} - S^{D\prime}_t\tilde{\delta}_i - S^{x\prime}_{it}\tilde{\beta} - x^{\prime}_{it}\tilde{\gamma}$. Then, we have

$$\begin{aligned}
\Delta\tilde{S}^u_{it} &= \Delta S^y_{it} - \Delta S^{D\prime}_t\tilde{\delta}_i - \Delta S^{x\prime}_{it}\tilde{\beta} - \Delta x^{\prime}_{it}\tilde{\gamma} \\
&= y_{it} - D^{\prime}_t\tilde{\delta}_i - x^{\prime}_{it}\tilde{\beta} - v^{\prime}_{it}\tilde{\gamma} \\
&= D^{\prime}_t\delta_i + x^{\prime}_{it}\beta + u_{it} - D^{\prime}_t\tilde{\delta}_i - x^{\prime}_{it}\tilde{\beta} - v^{\prime}_{it}\tilde{\gamma} \\
&= \left(u_{it} - v^{\prime}_{it}\gamma\right) - v^{\prime}_{it}\left(\tilde{\gamma} - \gamma\right) - D^{\prime}_t\left(\tilde{\delta}_i - \delta_i\right) - x^{\prime}_{it}\left(\tilde{\beta} - \beta\right) \\
&= u^+_{it} - v^{\prime}_{it}\left(\tilde{\gamma} - \gamma\right) - D^{\prime}_t\left(\tilde{\delta}_i - \delta_i\right) - x^{\prime}_{it}\left(\tilde{\beta} - \beta\right).
\end{aligned}$$

It can be shown that the last two parts of the formula can be neglected for long run variance estimation of $\Delta\tilde{S}^u_{it}$. Thus, the long run variance estimator based on $\Delta\tilde{S}^u_{it}$, that is $\tilde{\sigma}^2_{u\cdot v}$, asymptotically coincides with long run variance estimator of $u^+_{it} - v^{\prime}_{it}\left(\tilde{\gamma} - \gamma\right)$.

Let's define $\eta^+_{it} = \begin{bmatrix} u^+_{it}, & v^{\prime}_{it} \end{bmatrix}^{\prime}$, and then its long run variance is $\Omega^+_i = \begin{bmatrix} \sigma^2_{u\cdot v} & \\ & \Omega_{22} \end{bmatrix}$, so an infeasible long run variance estimator $\widehat{\Omega^+_i}$, using unobserved $\eta^+_{it}$ is consistent: $\widehat{\Omega^+_i} \xrightarrow{p} \Omega^+_i$.

Note that: $u^+_{it} - v^{\prime}_{it}\left(\tilde{\gamma} - \gamma\right) = \eta^{+\prime}_{it}\begin{bmatrix} 1 \\ -\left(\tilde{\gamma} - \gamma\right) \end{bmatrix}$, then HAC estimator, $\tilde{\Omega}^+_i$, for $u^+_{it} - v^{\prime}_{it}\left(\tilde{\gamma} - \gamma\right)$ can be written as:

$$\begin{bmatrix} 1 & -\left(\tilde{\gamma} - \gamma\right)^{\prime} \end{bmatrix}\widehat{\Omega^+_i}\begin{bmatrix} 1 \\ -\left(\tilde{\gamma} - \gamma\right) \end{bmatrix}$$

with

$$\begin{aligned}
\left(\tilde{\gamma} - \gamma\right) &\Rightarrow \begin{bmatrix} 0_{k\times k} & I_k & 0_{k\times p} & \cdots & 0_{k\times p} \end{bmatrix}\sigma_{u\cdot v}\left(\Pi^{\prime}\right)^{-1} \times \\
&\quad \left(\sum_{i=1}^{N}\int_0^1 g_{1,i}(s)g_{1,i}(s)^{\prime}ds\right)^{-1}\left(\sum_{i=1}^{N}\int_0^1 [G_{1,i}(1) - G_{1,i}(s)]dw_{u,i}(s)\right) \\
&= \sigma_{u\cdot v}\left(\Omega^{-1/2}_{vv}\right)^{\prime}d_{\gamma},
\end{aligned}$$

where $d_\gamma$ is the second $k \times 1$ block of

$$\left(\sum_{i=1}^{N}\int_0^1 g_{1,i}(s)g_{1,i}(s)'ds\right)^{-1}\left(\sum_{i=1}^{N}\int_0^1 [G_{1,i}(1)-G_{1,i}(s)]dw_{u,i}(s)\right).$$

This implies that $\tilde{\Omega}_i^{+}$ will converge to:

$$\begin{bmatrix}1 & -\left[\sigma_{u\cdot v}\left(\Omega_{vv}^{-1/2}\right)' d_\gamma\right]\end{bmatrix}\begin{bmatrix}\sigma_{u\cdot v}^2 & 0\\ 0 & \Omega_{vv}\end{bmatrix}\begin{bmatrix}1\\ -\sigma_{u\cdot v}\left(\Omega_{vv}^{-1/2}\right)' d_\gamma\end{bmatrix}$$

$$= \sigma_{u\cdot v}^2 + \sigma_{u\cdot v}^2\left(d_\gamma\right)'\left(\Omega_{vv}^{1/2}\right)^{-1}\Omega_{vv}^{1/2}\left(\Omega_{vv}^{1/2}\right)'\left(\left(\Omega_{vv}^{1/2}\right)'\right)^{-1} d_\gamma = \sigma_{u\cdot v}^2\left(1+d_\gamma' d_\gamma\right).$$

So we have: $\tilde{\sigma}_{u\cdot v}^2 \Rightarrow \frac{1}{N}\sum\limits_{i=1}^{N}\left[\sigma_{u\cdot v}^2\left(1+d_\gamma' d_\gamma\right)\right] = \sigma_{u\cdot v}^2\left(1+d_\gamma' d_\gamma\right)$. This implies that

$$\begin{aligned}
\tilde{W} &= \left(A_R^{-1}RA_{PIM}A_{PIM}^{-1}\left(\tilde{\theta}-\theta\right)\right)'\left[A_R^{-1}RA_{PIM}\tilde{V}_{PIM}A_{PIM}R'\left(A_R^{-1}\right)'\right]^{-1}\\
&\quad\times\left(A_R^{-1}RA_{PIM}A_{PIM}^{-1}\left(\tilde{\theta}-\theta\right)\right)\\
&= \left(A_R^{-1}RA_{PIM}A_{PIM}^{-1}\left(\tilde{\theta}-\theta\right)\right)'\left[A_R^{-1}RA_{PIM}\sigma_{u\cdot v}^2\breve{V}A_{PIM}R'\left(A_R^{-1}\right)'\right]^{-1}\\
&\quad\times\left(A_R^{-1}RA_{PIM}A_{PIM}^{-1}\left(\tilde{\theta}-\theta\right)\right)\times\frac{\sigma_{u\cdot v}^2}{\tilde{\sigma}_{u\cdot v}^2}\\
&\Rightarrow \frac{\left(R^*\Psi(V_{PIM})\right)'\left(R^*V_{PIM}R^{*\prime}\right)^{-1}\left(R^*\Psi(V_{PIM})\right)}{\left(1+d_\gamma' d_\gamma\right)}\\
&\sim \frac{\chi_q^2}{\left(1+d_\gamma' d_\gamma\right)}
\end{aligned}$$

and when $q=1$, we have

$$\tilde{t} \Rightarrow \frac{Z}{\sqrt{1+d_\gamma' d_\gamma}}.$$

For the result of the fixed-b test statistic, using $\tilde{\sigma}^{2*}_{u\cdot v}$, we have

$$\begin{aligned}
\tilde{W}^* &= \left(A_R^{-1} R A_{PIM} A_{PIM}^{-1}\left(\tilde{\theta}-\theta\right)\right)' \left[A_R^{-1} R A_{PIM} \tilde{V}^*_{PIM} A_{PIM} R' \left(A_R^{-1}\right)'\right]^{-1} \\
&\quad \times \left(A_R^{-1} R A_{PIM} A_{PIM}^{-1}\left(\tilde{\theta}-\theta\right)\right) \\
&= \left(A_R^{-1} R A_{PIM} A_{PIM}^{-1}\left(\tilde{\theta}-\theta\right)\right)' \left[A_R^{-1} R A_{PIM} \sigma^2_{u\cdot v} \check{V} A_{PIM} R' \left(A_R^{-1}\right)'\right]^{-1} \\
&\quad \times \left(A_R^{-1} R A_{PIM} A_{PIM}^{-1}\left(\tilde{\theta}-\theta\right)\right) \times \frac{\sigma^2_{u\cdot v}}{\tilde{\sigma}^{2*}_{u\cdot v}} \\
&\Rightarrow \frac{\left(R^*\Psi(V_{PIM})\right)' \left(R^* V_{PIM} R^{*\prime}\right)^{-1} \left(R^*\Psi(V_{PIM})\right)}{\frac{1}{N}\sum\limits_{i=1}^{N} Q_i^*(b)} \\
&\sim \frac{\chi^2_q}{\frac{1}{N}\sum\limits_{i=1}^{N} Q_i^*(b)}
\end{aligned}$$

and when $q = 1$, we have

$$\tilde{t}^* \Rightarrow \frac{Z}{\sqrt{\frac{1}{N}\sum\limits_{i=1}^{N} Q_i^*(b)}}.$$

Note that numerator and the denominator of the limiting distribution are independent, because Vogelsang and Wagner [2014] have proved that the numerator is independent with $Q_i^*(b)$ for all $i = 1, 2, \cdots, N$, then it follows that the numerator is independent with the sum $\frac{1}{N}\sum\limits_{i=1}^{N} Q_i^*(b)$.

Due to the independence of numerator and denominator in above limiting distribution, if we know $\mu_Q$, which given by $\mu_Q = E\left[Q_i^*(b)\right]$, then as $T \to \infty$ followed by $N \to \infty$, we have the following sequentially limit results:

$$\begin{aligned}
\tilde{W}^*_{\mu_Q} &= \left[T\left(R\tilde{\beta}-r\right)\right]' \left[\frac{\tilde{\sigma}^{2*}_{u.v}}{\mu_Q} R^* \hat{V} R^{*\prime}\right]^{-1} \left[T\left(R\tilde{\beta}-r\right)\right] \\
&= \mu_Q \cdot \tilde{W}^* \\
&\Rightarrow \frac{\chi^2_q}{\left[\frac{1}{N}\sum\limits_{i=1}^{N} \frac{1}{\mu_Q} Q_i^*(b)\right]} \\
&\xrightarrow{P} \chi^2_q \quad as\ N \to \infty.
\end{aligned}$$

Also, when $q = 1$, similarly, we have:

$$\begin{aligned}
\tilde{t}^*_{\mu_Q} &= \sqrt{\mu_Q} \cdot \tilde{t}^* \\
&\Rightarrow \frac{Z}{\sqrt{\frac{1}{N}\sum\limits_{i=1}^{N}\frac{1}{\mu_Q}Q^*_i(b)}} \\
&\xrightarrow{P} Z \quad as\ N \to \infty.
\end{aligned}$$

# Chapter 2

# Hypothesis testing in cointegrated panels: Asymptotic and Bootstrap method

This paper compares asymptotic and bootstrap hypothesis tests in cointegrated panels with cross-sectional uncorrelated units and endogenous regressors. All the tests are based on the panel integrated modified ordinary least square (panel IM-OLS) estimator from Vogelsang et al. [2016]. The aim of using the bootstrap tests is to deal with the size distortion problems in the finite samples of fixed-b tests. Finite sample simulations show that the bootstrap method is better than the asymptotic method in terms of having lower size distortions. In general, the stationary bootstrap is better than the conditional-on-regressors bootstrap, although in some cases, the conditional-on-regressors bootstrap has less size distortions. The improvement in size comes with only minor power losses, which can be ignored when the sample size is large.

## 2.1 Introduction

The bootstrap has become common in econometric analysis, especially in performing hypothesis tests. The basic idea of hypothesis tests is to compare the observed value of a test statistic with the distribution that it would follow if the null hypothesis were true. If the distribution is known, then we can perform exact tests. However, in many cases of interest, the distribution of the test statistic is only known asymptotically or is dependent upon nuisance parameters. In many cases, bootstrap hypothesis testing works well since the bootstrap statistics converge to the same asymptotic distributions as the sample statistics do. Therefore, the nuisance parameter dependent limit distributions can be approximated by the bootstrap simulations, which makes inference available.

The purpose of the present paper is to compare the fixed-b asymptotic hypothesis test with two bootstrap hypothesis tests, conditional-on-regressors bootstrap test and stationary bootstrap test, for panel cointegrated regressions with endogenous regressors. When the regressors are endogenous, it is well known that a variety of different methods, such as panel fully modified Ordinary Least Square (panel FM-OLS), panel dynamic Ordinary Least Square (panel DOLS) and panel integrated modified Ordinary Least Square (panel IM-OLS), will deliver estimators that have zero mean Gaussian mixture limiting distributions, which in turn allow asymptotic inference to be carried out (see Kao and Chiang [2000], Pedroni [2000], Bai et al. [2009], Mark and Sul [2003], Vogelsang et al. [2016]). Among those methodologies, panel IM-OLS relies on the fixed-b asymptotic theory. Compared with the traditional asymptotic theory, the fixed-b asymptotic theory can capture the impact of kernel and bandwidth choices on the sampling distributions of HAC-type test statistics. However, both of those asymptotic theories often provide poor approximations to the distributions of associated test statistics in finite samples, which leads to size distortion problems. To improve the quality of finite sample inference, in terms of decreasing size distortions, the bootstrap method is considered in this paper.

Although bootstrap methods are widely employed for analyzing nonstationary time series data, a surprisingly small proportion are devoted to bootstrap inference in cointegrated regressions. Li and Maddala [1997] investigated the usefulness of bootstrap methods for small sample inference in cointegrated regression models. Their simulation results showed that the substantial size distortions of the asymptotic tests can be corrected by properly implemented bootstrap methods. Psaradakis [2001] applied the sieve bootstrap procedure to cointegrated regressions, and his simulation study demonstrated the small-sample superiority of the sieve bootstrap over both the traditional asymptotic approximation and the blockwise bootstrap. Chang et al. [2006] considered the sieve bootstrap based on a VAR model for the cointegrated regressions. They established the bootstrap consistency for both OLS and DOLS, which leads to valid bootstrap inference. Shin and Hwang [2013] applied the stationary bootstrap to cointegrated regressions. They established the limiting distribution of the bootstrap ordinary least square estimator (OLSE) as well as the limiting null distribution of the bootstrap Wald-type test regarding the cointegration parameter. Also, finite sample size and power properties of the bootstrap test were studied by a Monte Carlo simulation. Note that in the above literature, the bootstrap methods are applied in pure time series setting.

The contribution of this paper is twofold. First, the results complement the existing literature by applying bootstrap inference to panel cointegrated regressions, and second, comparisons are made between bootstrap and fixed-b methods for inference using panel IM-OLS. Bootstrap methods are applied to a cointegrated panel with uncorrelated cross sectional units and homogeneous 2nd order moments. Finite sample size and power properties of the bootstrap test are studied by a Monte Carlo simulation. The bootstrap methods applied in this paper are the conditional-on-regressors bootstrap and stationary bootstrap. We do not consider the sieve bootstrap for two reasons. First, even though the sieve bootstrap can be applied in fairly general models and performs well in pure time series setting, Smeekes and Urbain [2014] questioned the validity of the use of VAR sieve bootstrap in panels with a moderate cross-sectional dimension. In addition, when estimating the models and carrying out inference, we do not assume the error terms follow AR or VAR models.

The rest of the paper is organized as follows. Section 2.2 introduces the model, assumptions and asymptotic inference based on the panel IM-OLS estimators. In Section 2.3, the conditional-on-regressors bootstrap and stationary bootstrap procedures are presented. Section 2.4 provides a simulation study to compare the size and the power of the bootstrap tests with the fixed-b asymptotic test. Section 2.5 summarizes the results and concludes the paper.

## 2.2 The model, assumptions and asymptotic inference

### 2.2.1 The model and assumptions

Consider the panel data model given by

$$y_{it} = D_t'\delta + x_{it}'\beta + u_{it} \tag{2.1}$$

$$x_{it} = x_{it-1} + v_{it} \tag{2.2}$$

where $i = 1, 2, \cdots, N$ and $t = 1, 2, \cdots, T$ index the cross-sectional and time series units respectively; $y_{it}$ and $u_{it}$ are scalars; $D_t$ is the deterministic component, and $\delta$ is a $p \times 1$ vector; $x_{it}$, $v_{it}$ and $\beta$ are $k \times 1$ vectors. Suppose that $\eta_{it} = \begin{bmatrix} u_{it} & v_{it}' \end{bmatrix}'$ is a $(k+1)$ dimensional stationary vector process across $i$, then the model introduced in (2.1) describes a system of panel cointegrated regressions, i.e. $y_{it}$ is cointegrated with $x_{it}$.

In the above system, we are interested in inference about $\beta$ based on the panel IM-OLS estimator. Before we define the panel IM-OLS estimator of $\beta$, we make following assumptions.

**Assumption 3.** *Assume that* $\{\eta_{it}\}_{i=1}^{N}$ *are cross-sectionally uncorrelated and 2nd order moments are constant across* $i$.

Note that the Assumption 3 only requires that the panels are homogeneous in the 2nd order moment; it's possible the higher order moment structures are heterogeneous across $i$.

**Assumption 4.** *Assume that for all* $i$, $\eta_{it}$ *is a stationary process and it satisfies a functional central limit theorem (FCLT) of the form*

$$T^{-1/2} \sum_{t=1}^{T} \eta_{it} \Rightarrow B_i(r) = \Omega^{1/2} W_i(r), \quad r \in (0, 1].$$

In Assumption 4, $[rT]$ represents the integer part of $rT$, and $W_i(r)$ is a $(k+1) \times 1$ vector of independent standard Brownian motions. $\Omega^{1/2}$ is a $(k+1) \times (k+1)$ matrix that satisfies $\Omega = \Omega^{1/2} \left(\Omega^{1/2}\right)'$, and

$$\Omega = \sum_{j=-\infty}^{\infty} \mathbb{E}\left(\eta_{it}\eta_{it-j}'\right) = \begin{bmatrix} \Omega_{uu} & \Omega_{uv} \\ \Omega_{vu} & \Omega_{vv} \end{bmatrix} > 0,$$

where it is clear that $\Omega_{uv} = \Omega_{vu}'$. The assumption $\Omega_{vv} > 0$ rules out cointegration in $x_{it}$. Partition $B_i(r)$ as $B_i(r) = \begin{bmatrix} B_{u,i}(r) & B_{v,i}'(r) \end{bmatrix}'$, and likewise partition $W_i(r)$ as $W_i(r) = \begin{bmatrix} w_{u,i}(r) & W_{v,i}'(r) \end{bmatrix}'$, where $w_{u,i}(r)$ and $W_{v,i}(r)$ are a scalar and a $k$-dimensional standard Brownian motion respectively. Using the Cholesky form of $\Omega^{1/2}$,

$$\Omega^{1/2} = \begin{bmatrix} \sigma_{u\cdot v} & \lambda_{uv} \\ 0_{k\times 1} & \Omega_{vv}^{1/2} \end{bmatrix},$$

it can be shown that $\sigma_{u\cdot v}^2 = \Omega_{uu} - \Omega_{uv}\Omega_{vv}^{-1}\Omega_{vu}$, and $\lambda_{uv} = \Omega_{uv}\left(\Omega_{vv}^{-1/2}\right)^{-1}$. By this Cholesky decomposition, we can write

$$B_i(r) = \begin{bmatrix} B_{u,i}(r) \\ B_{v,i}(r) \end{bmatrix} = \begin{bmatrix} \sigma_{u\cdot v} w_{u,i}(r) + \lambda_{uv} W_{v,i}(r) \\ \Omega_{vv}^{1/2} W_{v,i}(r) \end{bmatrix}.$$

**Assumption 5.** *For the deterministic component, $D_t$, assume that there is a $p \times p$ matrix $G_D$ and a vector of functions, $D(s)$, such that*

$$\lim_{T\to\infty} \sqrt{T} G_D^{-1} D_{[sT]} = D(s) \quad with \quad 0 < \int_0^r D(s)D(s)'ds < \infty, \quad 0 < r \leqslant 1$$

The deterministic component $D_t$ could include an intercept, time trend and polynomial of time and other functions of time.

### 2.2.2 Inference based on panel IM-OLS

This section provides some key results about the panel IM-OLS estimator and inference based on it. To conserve space, we don't provide all the details. After applying a partial sum transformation and adding the original regressors, $x_{it}$, to regression (2.1) , the system becomes

$$S_{it}^{y} = S_{t}^{D\prime}\delta + S_{it}^{x\prime}\beta + x_{it}'\gamma + S_{it}^{\tilde{u}}. \tag{2.3}$$

Vogelsang et al. [2016] pointed out that the parameter $\beta$ in panel cointegrated regression (2.1) can be consistently estimated by the panel IM-OLS estimator $\hat{\theta}$, which is the OLS estimator of regression (2.3). Its limiting distribution is given by

$$\begin{aligned}
A_{PIM}^{-1}\left(\hat{\theta}-\theta\right) &= \begin{bmatrix} G_D\left(\hat{\delta}-\delta\right) \\ T\left(\hat{\beta}-\beta\right) \\ \left(\hat{\gamma}-\Omega_{vv}^{-1}\Omega_{vu}\right) \end{bmatrix} \\
&\Rightarrow \sigma_{u\cdot v}\left(\Pi'\right)^{-1}\left(\int_0^1 \sum_{i=1}^{N} g_{1,i}(s)g_{1,i}'(s)ds\right)^{-1} \times \\
&\quad \left(\sum_{i=1}^{N}\int_0^1 \left[G_{1,i}(1)-G_{1,i}(s)\right]dw_{u,i}(s)\right) \\
&= \Psi_1
\end{aligned}$$

with $N$ fixed and $T \to \infty$. The scaling matrix $A_{PIM}^{-1}$ is a diagonal matrix given by

$$A_{PIM}^{-1} = \begin{bmatrix} G_D & & 0 \\ & T \cdot I_k & \\ 0 & & I_k \end{bmatrix},$$

and $\Pi$ is a diagonal matrix given by

$$\Pi = \begin{bmatrix} I_p & & 0 \\ & \Omega_{vv}^{1/2} & \\ 0 & & \Omega_{vv}^{1/2} \end{bmatrix}.$$

In addition,

$$G_{1,i}(r) = \int_0^r g_{1,i}(s)ds$$

where $g_{1,i}(r)$ is defined as

$$g_{1,i}(r) = \begin{bmatrix} \int_0^r D(s)ds \\ \int_0^r W_{v,i}(s)ds \\ W_{v,i}(r) \end{bmatrix}.$$

Conditional on $g_{1,i}(r)$ for all $i$, it can be shown that $\Psi_1 \sim N\left(0, V_{PIM}\right)$ where $V_{PIM}$ is given by

$$V_{PIM} = \sigma^2_{u \cdot v} \left(\Pi'\right)^{-1} \left( \int_0^1 \sum_{i=1}^N g_{1,i}(s) g'_{1,i}(s) ds \right)^{-1} \times$$
$$\left( \int_0^1 \sum_{i=1}^N \left[G_{1,i}(1) - G_{1,i}(s)\right] \left[G_{1,i}(1) - G_{1,i}(s)\right]' ds \right) \times$$
$$\left( \int_0^1 \sum_{i=1}^N g_{1,i}(s) g'_{1,i}(s) ds \right)^{-1} \Pi^{-1}.$$

Consider the null hypothesis $H_0: \ R\theta = r$, where $R \in \mathbb{R}^{q \times (p+2k)}$ with full rank $q$ and $r \in \mathbb{R}^q$. Define $q_{it} = \begin{bmatrix} S_t^{D\prime} & S_{it}^{x\prime} & x'_{it} \end{bmatrix}'$, $S_{it}^q = \sum_{j=1}^t q_{ij}$, and $S_{i0}^q$ is a zero vector for all $i$. Suppose that $\breve{\sigma}^2_{u \cdot v}$ is an estimator for $\sigma^2_{u \cdot v}$, then an estimator for $V_{PIM}$ is given by

$$\breve{V}_{PIM} = \breve{\sigma}^2_{u \cdot v} \left( T^{-2} \sum_{t=1}^T \sum_{i=1}^N A_{PIM} q_{it} q'_{it} A_{PIM} \right)^{-1} \times$$
$$\left( T^{-4} \sum_{t=1}^T \sum_{i=1}^N A_{PIM} \left[S_{iT}^q - S_{i,t-1}^q\right] \left[S_{iT}^q - S_{i,t-1}^q\right]' A_{PIM} \right) \times$$
$$\left( T^{-2} \sum_{t=1}^T \sum_{i=1}^N A_{PIM} q_{it} q'_{it} A_{PIM} \right)^{-1}.$$

Here, two potential candidates for $\breve{\sigma}^2_{u \cdot v}$ are considered. The first candidate, $\tilde{\sigma}^2_{u \cdot v}$, is based on the first differences of the residuals of the augmented partial sum regression (2.3), i.e.

$$\tilde{S}_{it}^u = S_{it}^y - S_t^{D\prime} \hat{\delta} - S_{it}^{x\prime} \hat{\beta} - x'_{it} \hat{\gamma},$$

where $\hat{\delta}$, $\hat{\beta}$ and $\hat{\gamma}$ are the panel IM-OLS estimators. Define a HAC estimator using the first difference of $\tilde{S}_{it}^{u}$

$$\tilde{\sigma}_{u\cdot v}^{2}=\frac{1}{N}\sum_{i=1}^{N}\left[T^{-1}\sum_{j=2}^{T}\sum_{h=2}^{T}k\left(\frac{|j-h|}{M}\right)\Delta\tilde{S}_{ij}^{u}\Delta\tilde{S}_{ih}^{u}\right].$$

Another HAC-type estimator is based on the first difference of the residuals from the further augmented partial sum regression

$$\tilde{\sigma}_{u\cdot v}^{2*}=\frac{1}{N}\sum_{i=1}^{N}\left[T^{-1}\sum_{j=2}^{T}\sum_{h=2}^{T}k\left(\frac{|j-h|}{M}\right)\Delta\tilde{S}_{ij}^{u*}\Delta\tilde{S}_{ih}^{u*}\right],$$

where

$$\tilde{S}_{it}^{u*}=S_{it}^{y}-S_{t}^{D\prime}\hat{\delta}_{i}-S_{it}^{x\prime}\hat{\beta}_{i}-x_{it}^{\prime}\hat{\gamma}_{i}-z_{it}^{\prime}\hat{\lambda}_{i},$$

$z_{it}$ is given by

$$z_{it}=t\sum_{j=1}^{T}q_{ij}-\sum_{j=1}^{t-1}\sum_{s=1}^{j}q_{is},$$

and $\hat{\delta}_i$, $\hat{\beta}_i$, $\hat{\gamma}_i$ and $\hat{\lambda}_i$ $(i=1,2,\cdots,N)$ are the OLS estimators from the further augmented regressions given by

$$S_{it}^{y}=S_{t}^{D\prime}\delta_{i}+S_{it}^{x\prime}\beta_{i}+x_{it}^{\prime}\gamma_{i}+z_{it}^{\prime}\lambda_{i}. \tag{2.4}$$

Note that the residual $\tilde{S}_{it}^{u*}$ is obtained by estimating regression (2.4) individual by individual. As discussed in Vogelsang and Wagner [2014], $\tilde{\sigma}_{u\cdot v}^{2*}$ has a fixed-b limit that is proportional to $\sigma_{u\cdot v}^{2}$, independent of $\hat{\theta}$, and does not depend upon additional nuisance parameters.

Let $\tilde{t}$ and $\tilde{W}$ denote statistics defined using $\tilde{\sigma}_{u\cdot v}^{2}$ to construct $\tilde{V}_{PIM}$, and likewise $\tilde{t}^{*}$ and $\tilde{W}^{*}$ denote statistics defined using $\tilde{\sigma}_{u\cdot v}^{2*}$ to construct $\tilde{V}_{PIM}^{*}$. Letting $\breve{t}$ and $\breve{W}$ denote either

$\tilde{t}$ and $\tilde{W}$ or $\tilde{t}^*$ and $\tilde{W}^*$, define the $t$ and $Wald$ statistics as:

$$\breve{t} = \frac{\left(R\hat{\theta} - r\right)}{\sqrt{RA_{PIM}\breve{V}_{PIM}A_{PIM}R'}}$$

$$\breve{W} = \left(R\hat{\theta} - r\right)' \left[RA_{PIM}\breve{V}_{PIM}A_{PIM}R'\right]^{-1} \left(R\hat{\theta} - r\right).$$

The limiting distributions of above test statistics are discussed in Vogelsang et al. [2016].

1. Under traditional bandwidth and kernel assumptions, with $N$ fixed as $T \to \infty$,

$$\tilde{W} \Rightarrow \frac{\chi_q^2}{1 + d_\gamma' d_\gamma}$$

and when $q = 1$,

$$\tilde{t} \Rightarrow \frac{Z}{\sqrt{1 + d_\gamma' d_\gamma}}$$

where $\chi_q^2$ is a chi-square random variable with q degrees of freedom that is correlated with $d_\gamma$, $Z$ is distributed standard normal and is correlated with $d_\gamma$, and $d_\gamma$ denotes the second $k \times 1$ block of

$$\left(\int_0^1 \sum_{i=1}^N g_{1,i}(s) g_{1,i}'(s) ds\right)^{-1} \left(\sum_{i=1}^N \int_0^1 \left[G_{1,i}(1) - G_{1,i}(s)\right] dw_{u,i}(s)\right).$$

2. Under fixed-b asymptotics where $M = bT$, $b \in (0, 1]$ is held fixed as $T \to \infty$, the

fixed-b limits of $\tilde{W}$ and $\tilde{t}$ are given by

$$\tilde{W} \Rightarrow \frac{\chi_q^2}{\frac{1}{N}\sum_{i=1}^{N} Q_b\left(\tilde{P}_i(r)\right)}$$

$$\tilde{t} \Rightarrow \frac{Z}{\sqrt{\frac{1}{N}\sum_{i=1}^{N} Q_b\left(\tilde{P}_i(r)\right)}}$$

where

$$\begin{aligned} Q_b\left(\tilde{P}_i(r)\right) &= \frac{2}{b}\int_0^1 \tilde{P}_i^2(r)dr - \frac{2}{b}\int_0^{1-b} \tilde{P}_i(r)\tilde{P}_i(r+b)dr \\ &\quad -\frac{2}{b}\int_{1-b}^1 \tilde{P}_i(r)\tilde{P}_i(1)dr + \tilde{P}_i^2(r) \end{aligned}$$

and

$$\begin{aligned} \tilde{P}_i(r) &= w_{u,i}(r) - g_{1,i}'(r)\left(\int_0^1 \sum_{i=1}^{N} g_{1,i}(s)g_{1,i}'(s)ds\right)^{-1} \times \\ &\quad \left(\sum_{i=1}^{N}\int_0^1 \left[G_{1,i}(1) - G_{1,i}(s)\right] dw_{u,i}(s)\right) \end{aligned}$$

3. Under fixed-b asymptotics where $M = bT$, $b \in (0,1]$ is held fixed as $T \to \infty$, the fixed-b limits of $\tilde{W}^*$ and $\tilde{t}^*$ are given by

$$\tilde{W}^* \Rightarrow \frac{\chi_q^2}{\frac{1}{N}\sum_{i=1}^{N} Q_b\left(\tilde{P}_i^*(r)\right)}$$

$$\tilde{t}^* \Rightarrow \frac{Z}{\sqrt{\frac{1}{N}\sum_{i=1}^{N} Q_b\left(\tilde{P}_i^*(r)\right)}}$$

where $Q_b(\cdot)$ is the same as above, and $\tilde{P}_i^*(r)$ is similar as $\tilde{P}_i(r)$ but its component is from the further augmented regression (2.4), which is a complicated stochastic process depend on the kernel function, bandwidth, and $W_i(r)$[1]. In addition, $Q_b\left(\tilde{P}_i^*(r)\right)$ is independent of $\chi_q^2$ and $Z$.

## 2.3 Bootstrap hypothesis tests

In this section, we introduce two different bootstrap procedures based on the panel IM-OLS estimator. For each of these bootstrap procedures, bootstrap test statistics are computed using the same formulas as the original test statistics but with resampled data.

### 2.3.1 Conditional-on-regressors bootstrap

A formal description of the conditional-on-regressors bootstrap is given below.

1. Calculate the residuals as

$$\hat{S}_{it}^u = S_{it}^y - S_t^{D\prime}\hat{\delta} - S_{it}^{x\prime}\hat{\beta} - x_{it}^{\prime}\hat{\gamma}$$

   where $\hat{\delta}$, $\hat{\beta}$ and $\hat{\gamma}$ are the panel IM-OLS estimators.

2. Obtain the bootstrap resamples $\left(u_{i1}^*, u_{i2}^*, \cdots, u_{iT}^*\right)$ from $\left(\Delta\hat{S}_{i1}^u, \Delta\hat{S}_{i2}^u, \cdots, \Delta\hat{S}_{iT}^u\right)$ by $i.i.d.$ sampling with replacement.

[1]For more details about the $Q_b\left(\tilde{P}_i^*(r)\right)$ function, please see Vogelsang and Wagner [2014].

3. Define $S_{it}^{u*} = \sum_{j=1}^{t} u_{ij}^{*}$. And generate the bootstrap samples $S_{it}^{y*}$ from

$$S_{it}^{y*} = S_t^{D\prime}\hat{\delta} + S_{it}^{x\prime}\hat{\beta} + x_{it}^{\prime}\hat{\gamma} + S_{it}^{u*}. \tag{2.5}$$

4. Define the bootstrap statistics as

$$\breve{t}_{CBS}^{*} = \frac{\left(R\hat{\theta}^{*} - R\hat{\theta}\right)}{\sqrt{RA_{PIM}\breve{V}_{PIM}^{*}A_{PIM}R'}}$$
$$\breve{W}_{CBS}^{*} = \left(R\hat{\theta}^{*} - R\hat{\theta}\right)' \left[RA_{PIM}\breve{V}_{PIM}^{*}A_{PIM}R'\right]^{-1} \left(R\hat{\theta}^{*} - R\hat{\theta}\right)$$

where $\hat{\theta}^{*}$ is the bootstrap panel IM-OLS estimator for regression (2.5), $\breve{V}_{PIM}^{*}$ is constructed exactly as $\breve{V}_{PIM}$ but using the bootstrap data.

5. Repeat above steps 2-4 independently $B$ times to obtain samples $\left\{\breve{t}_{CBS,j}^{*}\right\}_{j=1}^{B}$ and $\left\{\breve{W}_{CBS,j}^{*}\right\}_{j=1}^{B}$.

6. Compute the equal tail bootstrap p-value as

$$p^{*}\left(\breve{t}\right) = 2\min\left(\frac{1}{B}\sum_{j=1}^{B} I\left(\breve{t}_{CBS,j}^{*} \leq \breve{t}\right), \frac{1}{B}\sum_{b=1}^{B} I\left(\breve{t}_{CBS,j}^{*} > \breve{t}\right)\right)$$
$$p^{*}\left(\breve{W}\right) = \frac{1}{B}\sum_{j=1}^{B} I\left(\breve{W}_{CBS,j}^{*} > \breve{W}\right)$$

where $I(\cdot)$ is the indicator function. Reject the null hypothesis if the equal tail bootstrap p-value is less than 5%.

### 2.3.2 Stationary bootstrap

The stationary bootstrap, proposed by Politis and Romano [1994], is a special type of block bootstrap where the block size follows a geometric distribution instead of a fixed number. For a geometric distribution with parameter $p_T$, the expected block size of the stationary bootstrap is $1/p_T$. The stationary bootstrap has been used in the literature of unit root tests, cointegration tests and cointegrated regression inference (see Swensen [2003], Paparoditis and Politis [2005], Parker et al. [2006], Shin [2015] and Shin and Hwang [2013]). It can capture the serial correlation structure in the original sample by block resampling, and it produces stationary bootstrap samples. A formal description of the stationary bootstrap inference procedure is given below.

1. Calculate the residuals as

$$\hat{S}_{it}^{u} = S_{it}^{y} - S_{t}^{D\prime}\hat{\delta} - S_{it}^{x\prime}\hat{\beta} - x_{it}^{\prime}\hat{\gamma}$$

   where $\hat{\delta}$, $\hat{\beta}$ and $\hat{\gamma}$ are the panel IM-OLS estimators.

2. Define $\hat{\eta}_{it} = \left(\Delta\hat{S}_{it}^{u}, \Delta x_{it}\right)$ for $t = 1, 2, \cdots, T$, where $\hat{S}_{i0}^{u} = 0$, and $x_{i0}$ is zero vector for all $i$.

3. Resample the series $\{\hat{\eta}_{it}\}_{t=1}^{T}$ via the stationary bootstrap, obtaining $\{\hat{\eta}_{it}^{*}\}_{t=1}^{T}$.

4. Partition $\hat{\eta}_{it}^{*} = \begin{bmatrix} u_{it}^{*} & v_{it}^{*\prime} \end{bmatrix}^{\prime}$ analogously as $\eta_{it} = \begin{bmatrix} u_{it} & v_{it}^{\prime} \end{bmatrix}^{\prime}$. Obtain the bootstrap samples $\{x_{it}^{*}\}_{t=1}^{T}$ by

$$x_{it}^{*} = \sum_{j=1}^{t} v_{ij}^{*},$$

and generate the bootstrap samples $\{y_{it}^*\}_{t=1}^T$ from

$$y_{it}^* = D_t'\hat{\delta} + x_{it}^{*\prime}\hat{\beta} + u_{it}^*. \tag{2.6}$$

5. Define the bootstrap statistics as

$$\breve{t}_{SBS}^* = \frac{\left(R\tilde{\theta}^* - R\hat{\theta}\right)}{\sqrt{RA_{PIM}\breve{V}_{PIM}^* A_{PIM}R'}}$$
$$\breve{W}_{SBS}^* = \left(R\tilde{\theta}^* - R\hat{\theta}\right)' \left[RA_{PIM}\breve{V}_{PIM}^* A_{PIM}R'\right]^{-1} \left(R\tilde{\theta}^* - R\hat{\theta}\right)$$

   where $\tilde{\theta}^*$ is the bootstrap panel IM-OLS estimator from regression (2.6), $\breve{V}_{PIM}^*$ is constructed exactly as $\breve{V}_{PIM}$ but using the bootstrapping data.

6. Repeat above steps 3-5 independently $B$ times to obtain samples $\left\{\breve{t}_{SBS,j}^*\right\}_{j=1}^B$ and $\left\{\breve{W}_{SBS,j}^*\right\}_{j=1}^B$.

7. Compute the equal tail bootstrap p-value as

$$p^*\left(\breve{t}\right) = 2\min\left(\frac{1}{B}\sum_{j=1}^B I\left(\breve{t}_{SBS,j}^* \le \breve{t}\right), \frac{1}{B}\sum_{j=1}^B I\left(\breve{t}_{SBS,j}^* > \breve{t}\right)\right)$$
$$p^*\left(\breve{W}\right) = \frac{1}{B}\sum_{j=1}^B I\left(\breve{W}_{SBS,j}^* > \breve{W}\right)$$

   where $I(\cdot)$ is the indicator function. Reject the null hypothesis if the equal tail bootstrap p-value is less than 5%.

Note that the step 1 of Section 2.3.1 and Section 2.3.2 are both based on the regression

(2.3), which includes the augmented regressor $x_{it}$ in the regression. It might worth exploring how the bootstrap works if the residuals are calculated from non-augmented partial sum regression. That is, for both stationary and conditional-on-regressors bootstrap, the residuals are obtained from

$$\hat{S}_{it}^{u} = S_{it}^{y} - S_{t}^{D\prime}\hat{\delta} - S_{it}^{x\prime}\hat{\beta} \tag{2.7}$$

where $\hat{\delta}$ and $\hat{\beta}$ are the panel IM-OLS estimators, and all other steps are the same as its corresponding procedures. Using these residuals, the stationary bootstrap resampling and the conditional-on-regressors bootstrap resampling will be more comparable, and it could capture some of endogeneity in the bootstrap resamples. In next section, we will provide the bootstrap results based on Section 2.3.1 and Section 2.3.2 as well as the bootstrap results based on the residuals from regression (2.7).

## 2.4 Finite sample simulations

In this section, we compare finite sample size and power performance of the bootstrap tests with the asymptotic tests based on the panel IM-OLS estimators. The data generating process is the same as in Vogelsang et al. [2016], which is given by

$$\begin{aligned}
y_{it} &= \mu + x1_{it}\beta_1 + x2_{it}\beta_2 + u_{it} \\
x1_{it} &= x1_{i,t-1} + v1_{it} \\
x2_{it} &= x2_{i,t-1} + v2_{it}
\end{aligned}$$

where for all $i = 1, 2, \cdots, N$, $u_{i0} = 0$, $x1_{i0}$ and $x2_{i0}$ are zero vectors, and

$$\begin{aligned}
u_{it} &= \rho_1 u_{i,t-1} + \epsilon_{it} + \rho_2 \left( e1_{it} + e2_{it} \right) \\
v1_{it} &= e1_{it} + 0.5 e1_{i,t-1} \\
v2_{it} &= e2_{it} + 0.5 e2_{i,t-1}
\end{aligned}$$

where $\epsilon_{it}$, $e1_{it}$ and $e2_{it}$ are $i.i.d.$ standard normal random variables independent of each other. The parameter values are $\mu = 3$, $\beta_1 = \beta_2 = 1$. In addition, we use $\rho_1, \rho_2 \in \{0.6, 0.9\}$. The parameter $\rho_1$ controls serial correlation in the regression error, and $\rho_2$ determines the endogeneity of the regressors. In this paper, we only provide results where both $\rho_1$ and $\rho_2$ are relatively large because according to the findings in Vogelsang et al. [2016], if $\rho_1$ and $\rho_2$ are relatively small ($\rho_1 = \rho_2 = 0.3$), there are only minor size distortions for fixed-b asymptotic tests. Therefore, the bootstrap method is not necessary when $\rho_1$ and $\rho_2$ are small. The kernel function used in this simulation study is the Bartlett kernel, and the bandwidths are given by $M = bT$ with $b \in \{0.06, 0.1, 0.3, 0.5, 0.7, 0.9, 1\}$. We use $p_T = 0.02(T/50)^{-1/3}$ as the block length parameter in the stationary bootstrap[2]. The sample sizes are $N = 5$, $T \in \{50, 100\}$. The number of bootstrap replications is $B = 399$, and the number of simulation replications is 1000.

Using the simulation designed above, we only report results for cases where $\rho_1 = \rho_2$. The results include t-statistics for testing the null hypothesis $H_0 : \beta_1 = 1$ and Wald statistics

[2] Politis and White (2004, 2009) considered estimators constructed via stationary bootstrap to obtain an approximation to the sampling distribution of the mean of a finite sample from the (strictly) stationary real-valued sequence. They showed that the optimal block length parameter minimizing MSE of the stationary bootstrap sample mean is $cT^{-1/3}$ for some constant $c$. In addition, Shin and Hwang (2013) considered the bootstrap ordinary least square estimator for cointegrating regressions. They established large sample validity of a bootstrap test regarding cointegration parameters and showed that the block length parameter $0.02(T/50)^{-1/3}$ would provide stable size performance for the stationary bootstrap test.

for testing the joint null hypothesis $H_0 : \beta_1 = 1, \beta_2 = 1$. The asymptotic panel IM-OLS statistics were implemented in two ways. The first uses $\tilde{\sigma}^2_{u \cdot v}$ and is labeled panel IMOLS(D), and the second uses $\tilde{\sigma}^{2*}_{u \cdot v}$ and is labeled panel IMOLS(fb). The bootstrap panel IM-OLS statistics were implemented in four ways. The first two statistics are based on the bootstrapped $\tilde{\sigma}^2_{u \cdot v}$ and are labeled Cond-BS IMOLS (D) and Stat-BS IMOLS(D) respectively for the conditional-on-regressors bootstrap and the stationary bootstrap. The second two statistics are based on the bootstrapped $\tilde{\sigma}^{2*}_{u \cdot v}$ and are labeled Cond-BS IMOLS(fb) and Stat-BS IMOLS(fb) respectively. Rejections for panel IMOLS(D) are carried out using $N(0, 1)$ critical values for the t test and $\chi^2_2$ critical values for the Wald test. Rejections for panel IMOLS(fb) are carried out using fixed-b asymptotic critical values. In contrast, rejections for the bootstrap statistics are carried out by comparing the bootstrap p-value with the nominal level, which is 5% in this simulation.

In order to see if the bootstrap methods can help solve the over-rejection problem of the asymptotic tests in finite sample, we plot in Figures 2.1-2.8 null rejection probabilities of the t and Wald tests as a function of $b \in (0, 1]$. The first two figures give the results for $N = 5$, $T = 50$ using the Bartlett kernel and $\rho_1 = \rho_2 = 0.6$. In Figure 2.1, all t-tests have some over-rejection problems, and there is no test that dominates the others in this scenario. When the bandwidth is small ($b = 0.1$), panel IMOLS(D) is better than the other tests because it is conservative. But when bandwidth is relative large ($b > 0.2$), it turns out that Cond-BS IMOLS(D) is the best. Even though it is better than all other tests, the Cond-BS IMOLS(D) rejection probabilities are close to 15%, which is much larger than nominal level 5%. In Figure 2.2, for Wald tests, the stationary bootstrap tests dominate the other tests for all values of $b$. Its rejection probabilities are close to 10% for all values of $b$, which is much better than the asymptotic tests. Also, Cond-BS IMOLS(D) has rejection probabilities around 12% as long as the bandwidth is not very small ($b > 0.1$).

In Figures 2.3 and 2.4, all the settings are the same as in Figure 2.1 and 2.2 except that the time series sample size increases from $T = 50$ to $T = 100$. Comparing Figures 2.1 and 2.3, Figures 2.2 and 2.4, we see that both t and Wald tests have less size distortion when sample size increases. In Figure 2.3, the patterns of the rejection probabilities of all t tests are similar as those in Figure 2.1. And still, there is no test that dominates the others for t tests. For Wald tests, the pattern is very clear. As we can see from Figure 2.4, for all values of $b$, the asymptotic tests have the highest size distortions. But the rejection probabilities of the stationary bootstrap tests are stable and close to 5%, which implies that stationary bootstrap successfully solves the over-rejection problem in this scenario. The null rejection probabilities of the conditional-on-regressors bootstrap tests are higher than 5% but less than those of the asymptotic tests.

As the values of $\rho_1$, $\rho_2$ increase to 0.9, there exists strong serial correlation and endogeneity. We can see from Figures 2.5-2.8 that all the tests have serious over-rejection problems regardless of bandwidth. For $N = 5$, a time series sample size $T = 100$ is not large enough for the stationary bootstrap to obtain reasonable size that is close to 5%. But among all three tests, the stationary bootstrap tests are better than conditional on regressor bootstrap and asymptotic fixed-b tests. And this is true for both t and Wald tests, which is not the case when $\rho_1 = \rho_2 = 0.6$. In addition, unlike the results before, Stat-BS IMOLS(D) and Stat-BS IMOLS(fb) rejection probabilities are not that close any more. Generally speaking, when both $\rho_1$ and $\rho_2$ are very large, Stat-BS IMOLS(D) tends to have the smaller size distortion than Stat-BS IMOLS(fb). Therefore, when $\rho_1$, $\rho_2$ are large, in order to obtain reasonable size, we need a very large time series sample size and to use the Stat-BS IMOLS(D) statistics.

From the above, we see that the bootstrap tests generally have less size distortions than the asymptotic tests. However, if the power of the bootstrap testing is low, then the bootstrap methods are less useful. When the alternative is true, some bootstrap methods fail to simulate critical values that are valid under the null in which case the tests have no power. Therefore, the analysis of the power properties of the bootstrap tests is necessary. For the sake of brevity we only display results of the stationary bootstrap for the case $\rho_1 = \rho_2 = 0.6$ for the Wald test for $N = 5$, $T \in \{50, 100\}$ and using the Bartlett kernel. Starting from the null values of $\beta_1$ and $\beta_2$ equal to 1, we consider under the alternative $\beta_1 = \beta_2 = \beta \in (1, 1.25]$, using (including the null value) a total of 13 values on a grid with mesh 0.02. We focus on raw power using bootstrapped critical values.

Using $N = 5$, $T = 50$, with Bartlett kernel and $b = 0.1$, Figure 2.9 provides power comparisons between Stat-BS IMOLS(D) and Stat-BS IMOLS(fb). The power plots indicate that when the alternative is true, the stationary bootstrap is still simulating critical values that are valid under the null. Figure 2.10 displays the same power comparisons as in Figure 2.9 but with $T = 100$. The main finding is that power increases as $T$ increases. From Figures 2.9 and 2.10, we can see that the bootstrap tests have good power.

As mentioned in the end of Section 2.3, we also consider the stationary bootstrap and the conditional-on-regressors bootstrap based on the residuals from the non-augmented partial sum regression. Null rejection probabilities of the t and Wald tests as a function of $b \in (0, 1]$ are shown in Figures 2.11-2.18. The general patterns in Figures 2.11-2.18 are close to those in Figures 2.1-2.8. Overall, the bootstrap methods based on the residuals from the non-augmented partial sum regression have less size distortion problems than the asymptotic methods especially for the Wald test with large sample size. But when serial correlation and endogeneity are both large, it seems that the bootstrap results are depend little on the choice of the residuals. The power results in this case are displayed in Figure 2.19 and 2.20, which are very similar to the power results in Figures 2.9 and 2.10.

## 2.5 Summary and conclusion

This paper compares bootstrap tests with fixed-b asymptotic tests based on the panel IM-OLS estimator of Vogelsang et al. [2016] for a homogeneous panel cointegrated regression with endogenous regressors. The bootstrap methods used are the conditional-on-regressors bootstrap and the stationary bootstrap. The purpose of using the bootstrap tests is to improve the quality of finite sample inference. The Monte Carlo simulations show that the bootstrap methods can effectively reduce size distortions in finite samples. In general, the stationary bootstrap has less size distortions than the conditional-on-regressors bootstrap and asymptotic fixed-b tests, especially when there is strong serial correlation and endogeneity ($\rho_1 = \rho_2 = 0.9$). It is necessary to have a large time series sample size to obtain reasonable size of the tests. When the serial correlation and endogeneity is medium ($\rho_1 = \rho_2 = 0.6$), the bootstrap methods still have less size distortion, but t and Wald tests have different results. For Wald tests, the stationary bootstrap is always better than the other two methods. In contrast, for t-tests, Cond-BS IMOLS(D), the statistic constructed using a HAC estimator based on the first differences of the residuals from the augmented partial sum regression for $\sigma^2_{u\cdot v}$, has less size distortions when the bandwidth is relatively large ($b > 0.25$). In addition, the stationary bootstrap statistics are more robust than all other test statistics for all values of bandwidth. Finally, the power plots from the simulation show that the bootstrap tests have good power.

Further research will study the panel IM-OLS method for estimation and inference in a heterogeneous cointegrating panel with endogenous regressors. In that more general scenario, finding a fixed-b asymptotic pivotal statistic based on panel IM-OLS will be challenging. However, the results in this paper indicate that the bootstrap method could be an alternative solution for hypothesis tests. In addition, if the panel consists of cross-sectional dependent units, then the bootstrap procedure will need to be modified to resample all individuals together rather than resample individual by individual. Another topic of future research is to establish the consistency of the bootstrap for panel IM-OLS tests.

APPENDIX

# Figures

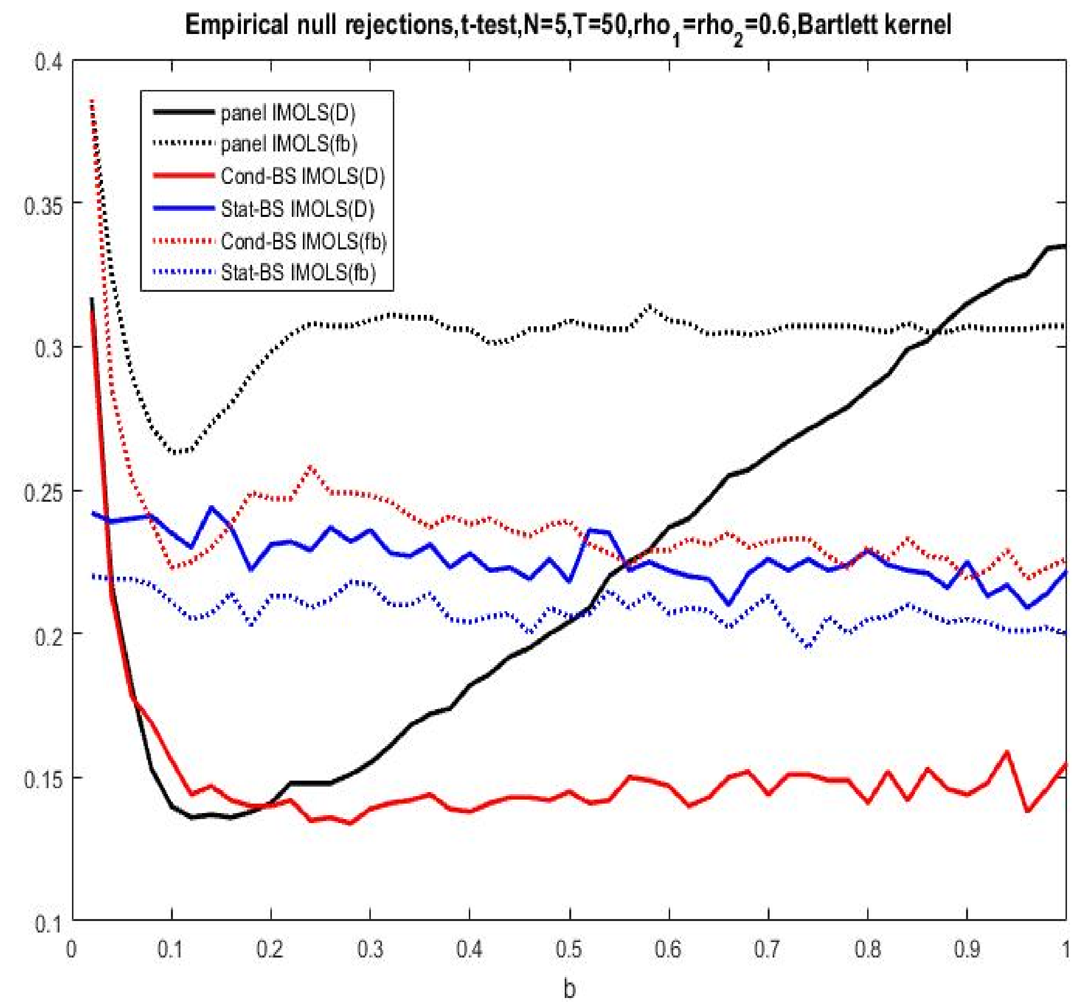


Figure 2.1: Empirical null rejections, t-test, $N = 5$, $T = 50$, $\rho_1 = \rho_2 = 0.6$, Bartlett kernel

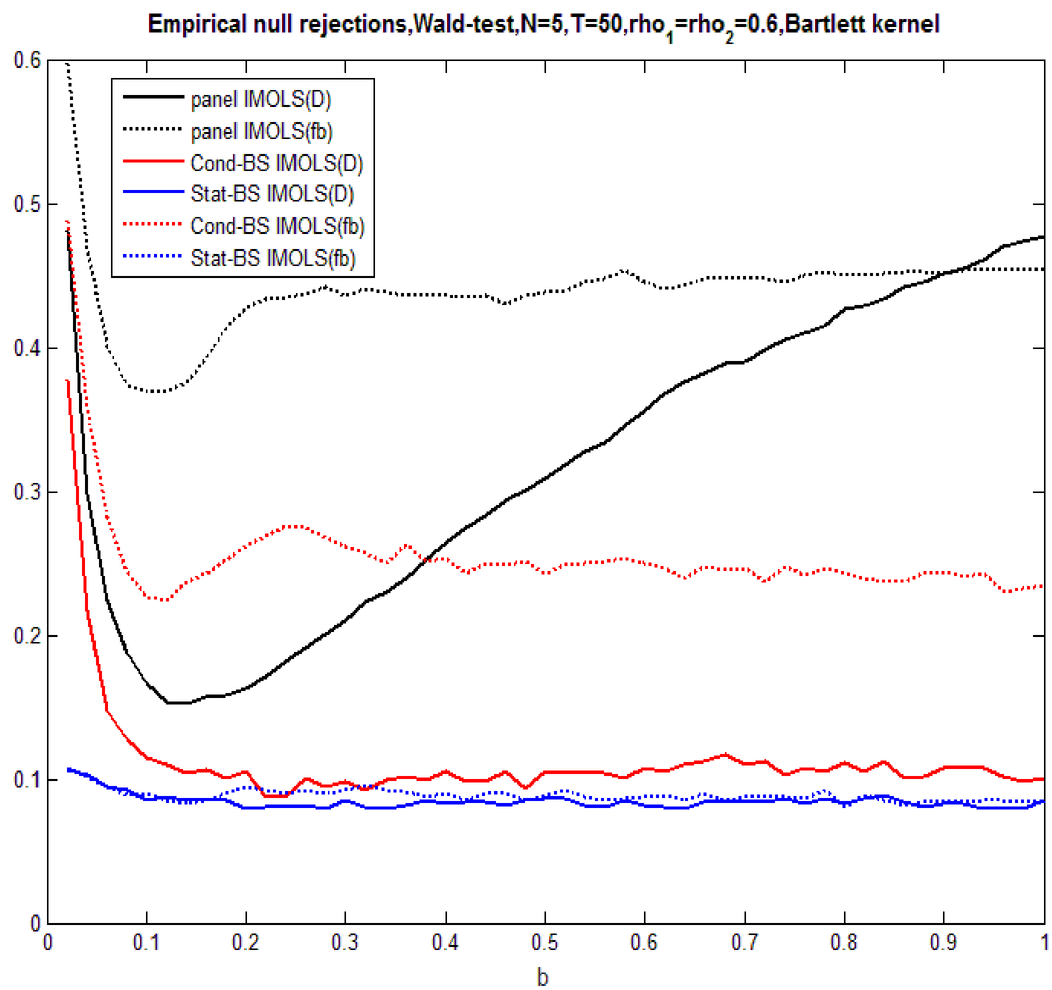


Figure 2.2: Empirical null rejections, Wald test, $N = 5$, $T = 50$, $\rho_1 = \rho_2 = 0.6$, Bartlett kernel

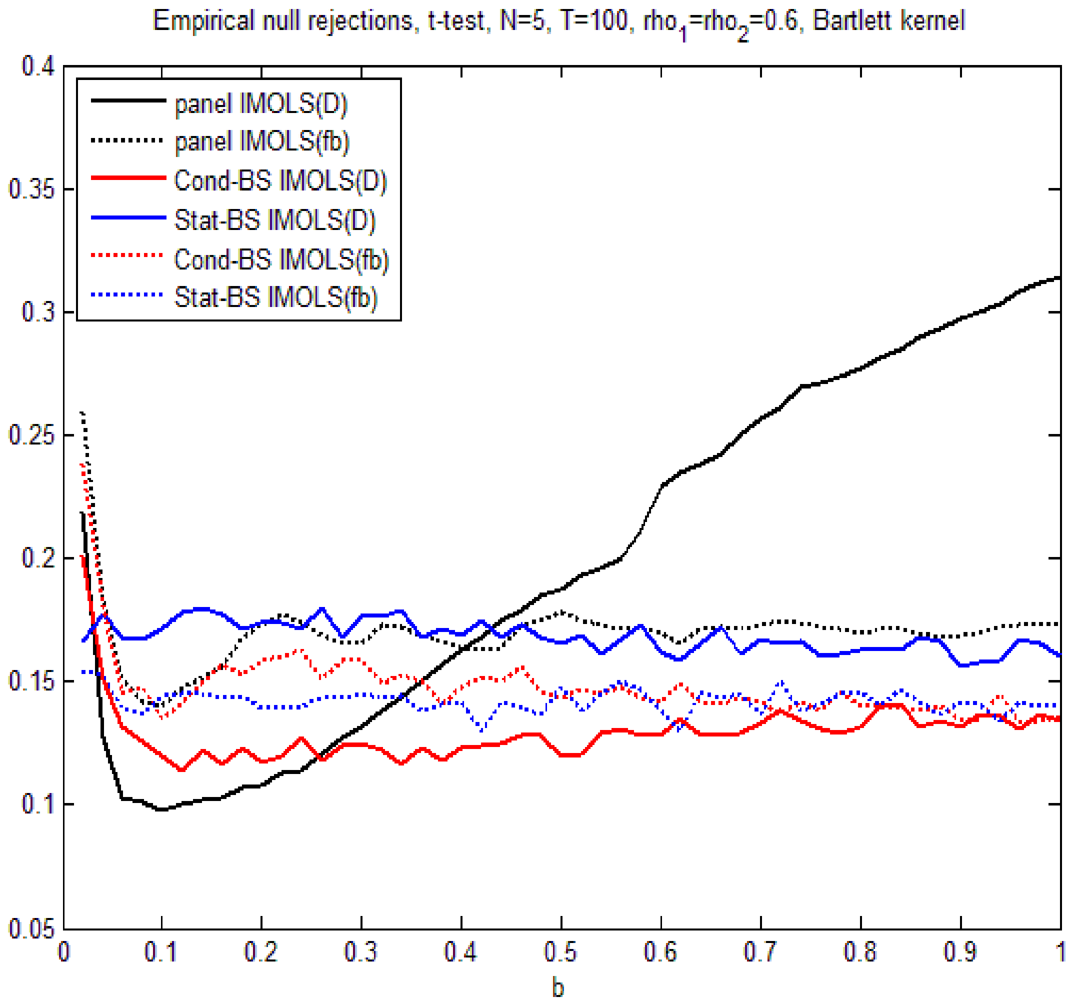


Figure 2.3: Empirical null rejections, t-test, $N = 5$, $T = 100$, $\rho_1 = \rho_2 = 0.6$, Bartlett kernel

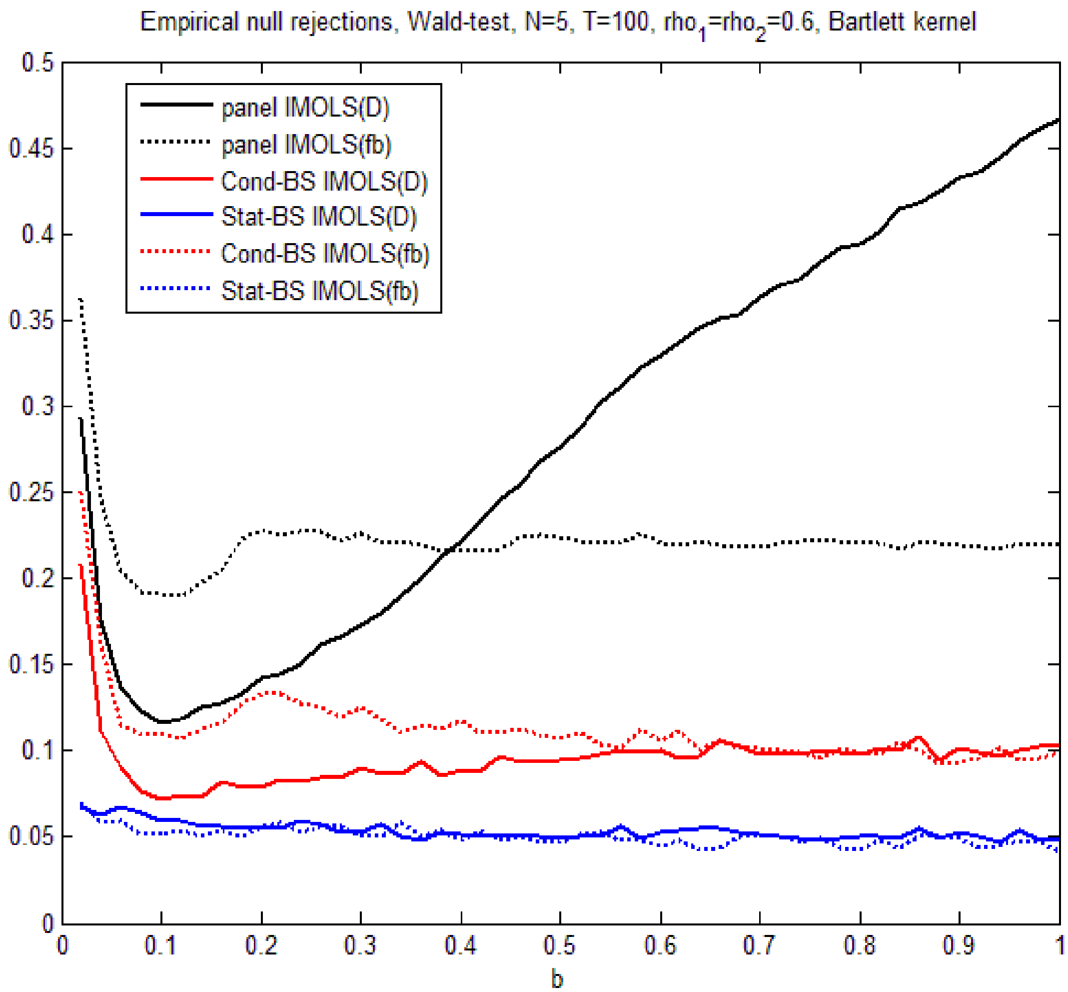


Figure 2.4: Empirical null rejections, Wald test, $N = 5$, $T = 100$, $\rho_1 = \rho_2 = 0.6$, Bartlett kernel

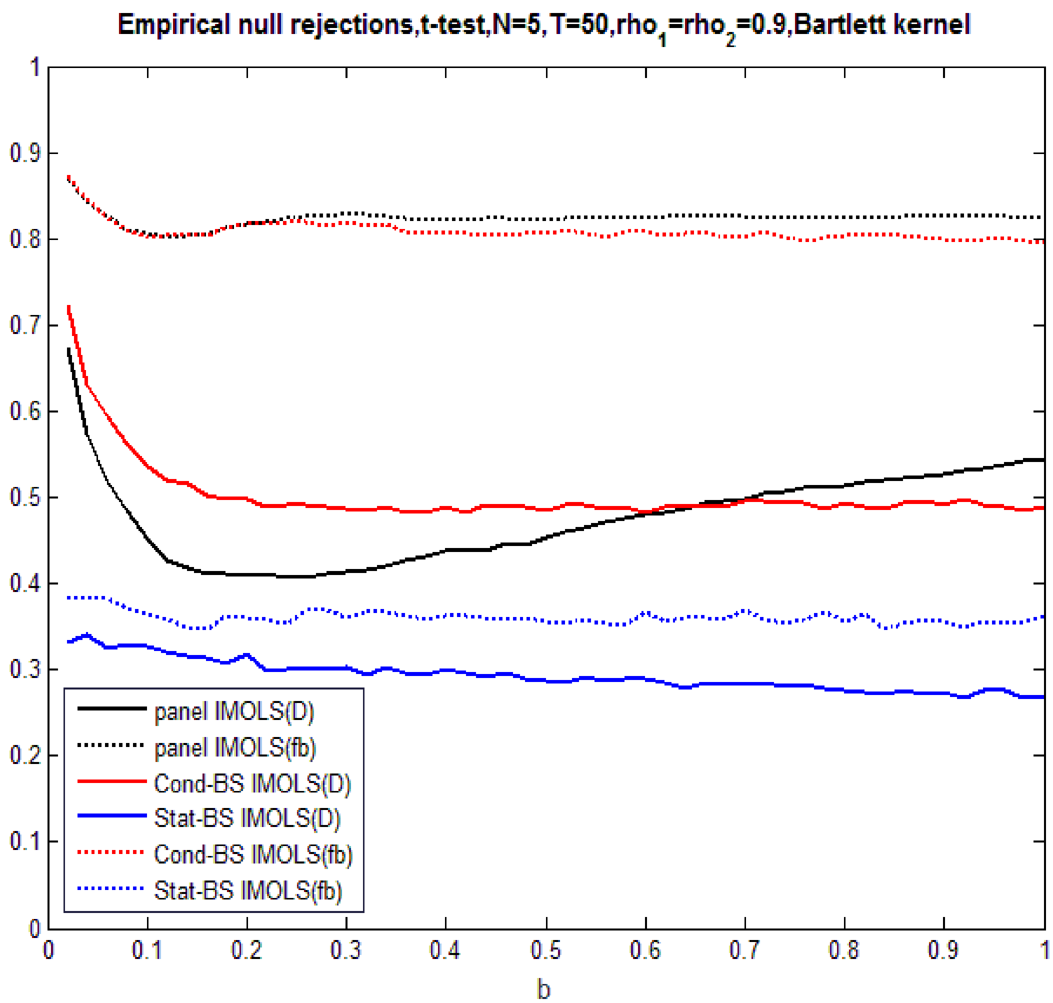


Figure 2.5: Empirical null rejections, t-test, $N = 5$, $T = 50$, $\rho_1 = \rho_2 = 0.9$, Bartlett kernel

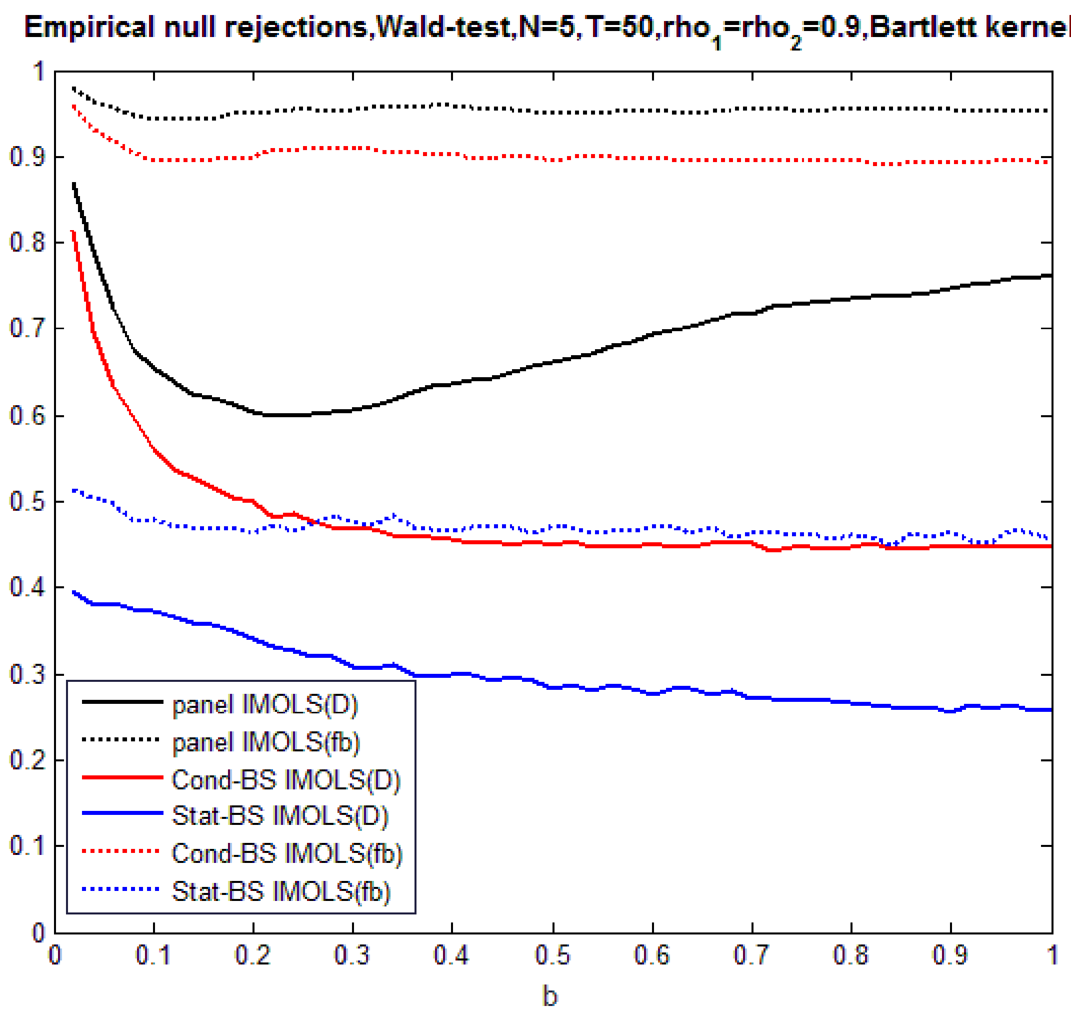


Figure 2.6: Empirical null rejections, Wald test, $N = 5$, $T = 50$, $\rho_1 = \rho_2 = 0.9$, Bartlett kernel

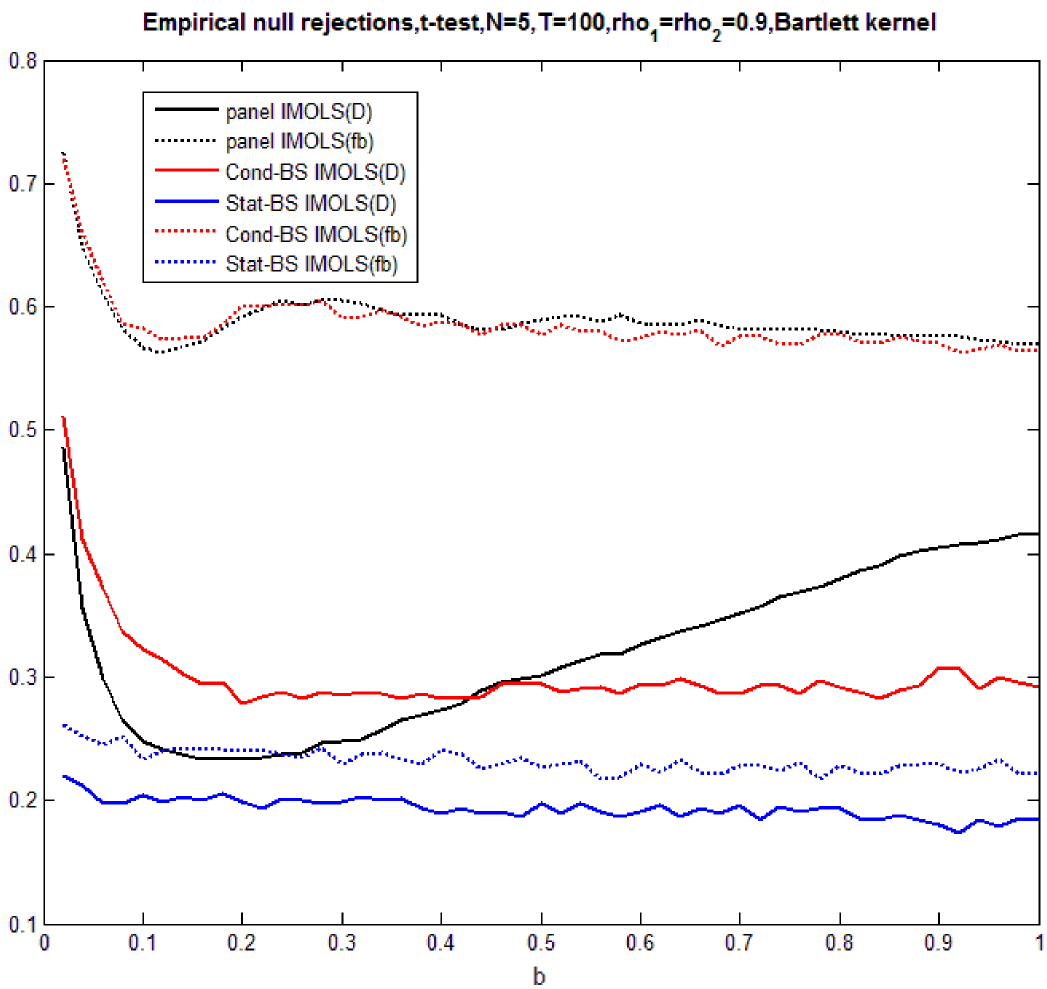


Figure 2.7: Empirical null rejections, t-test, $N = 5$, $T = 100$, $\rho_1 = \rho_2 = 0.9$, Bartlett kernel

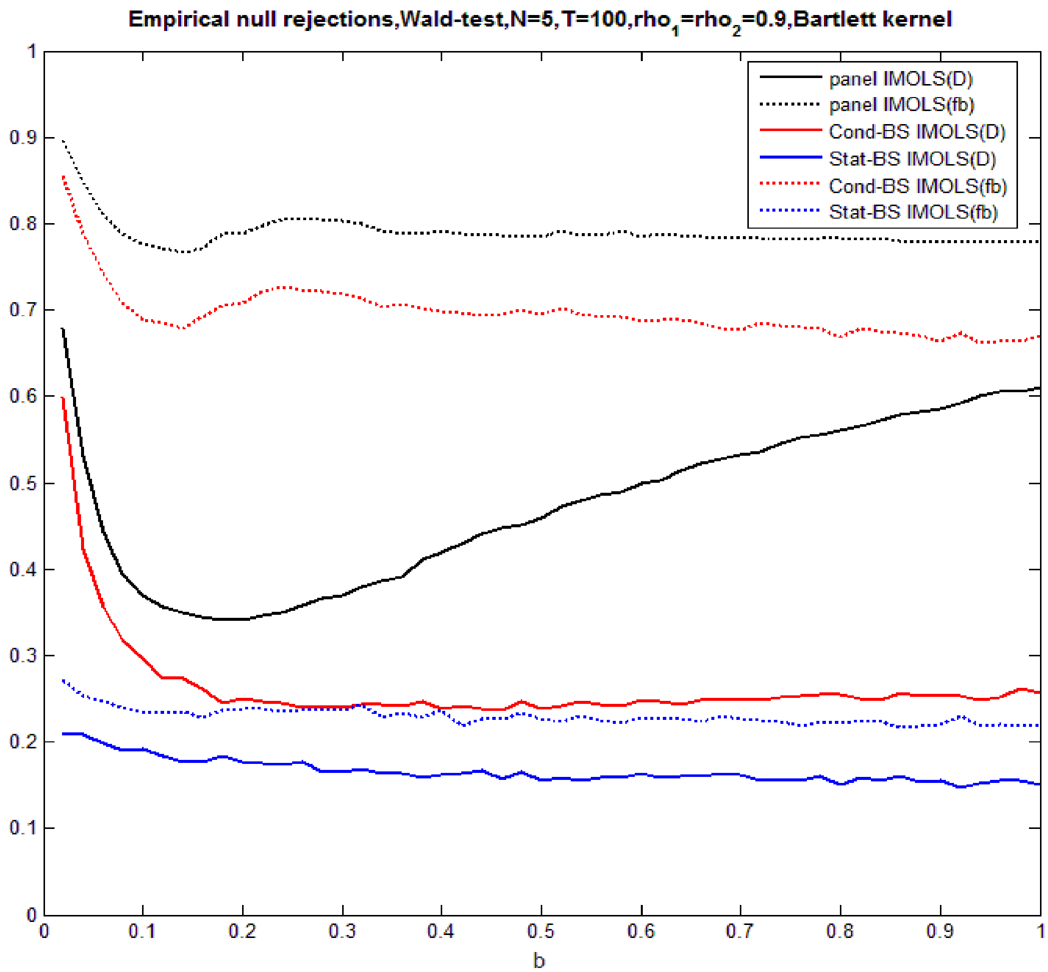


Figure 2.8: Empirical null rejections, Wald test, $N = 5$, $T = 100$, $\rho_1 = \rho_2 = 0.9$, Bartlett kernel

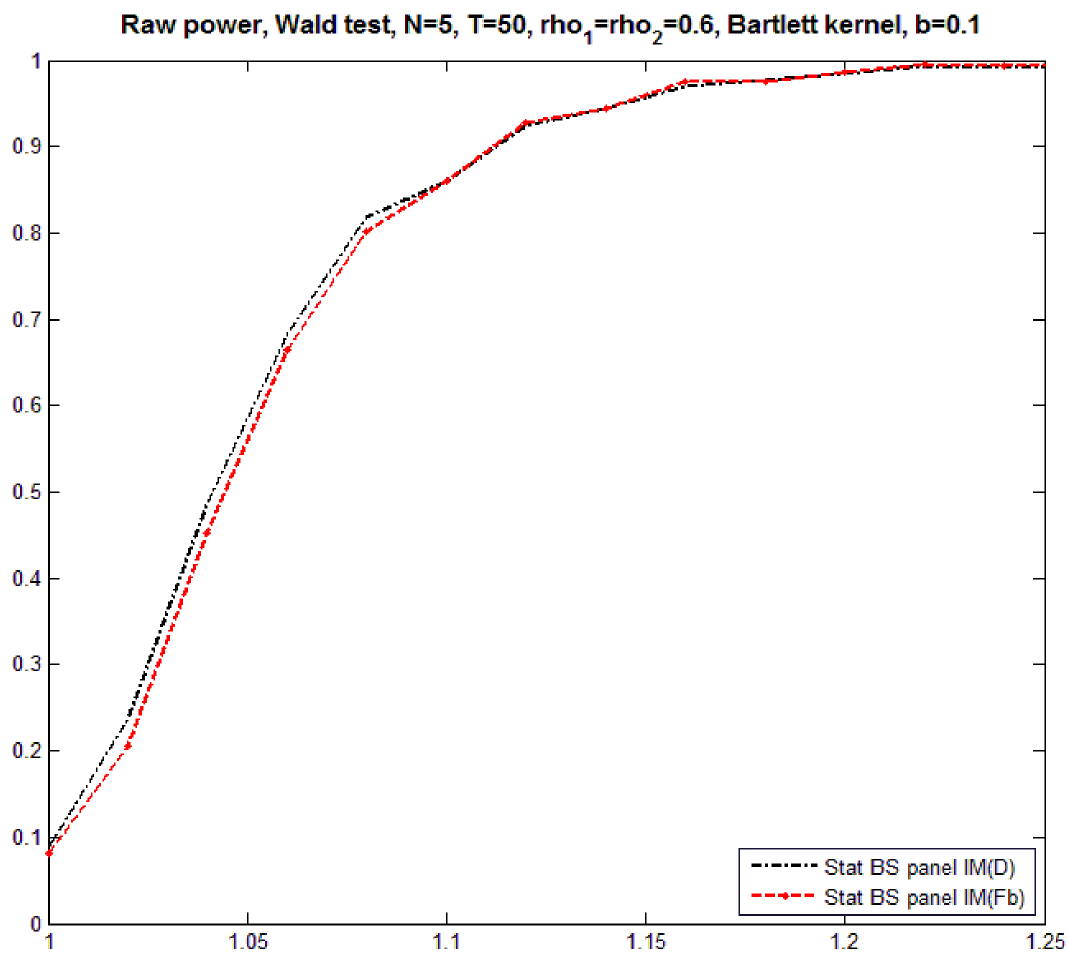


Figure 2.9: Raw power, Wald test, $N = 5$, $T = 50$, $\rho_1 = \rho_2 = 0.6$, $b = 0.1$, Bartlett kernel

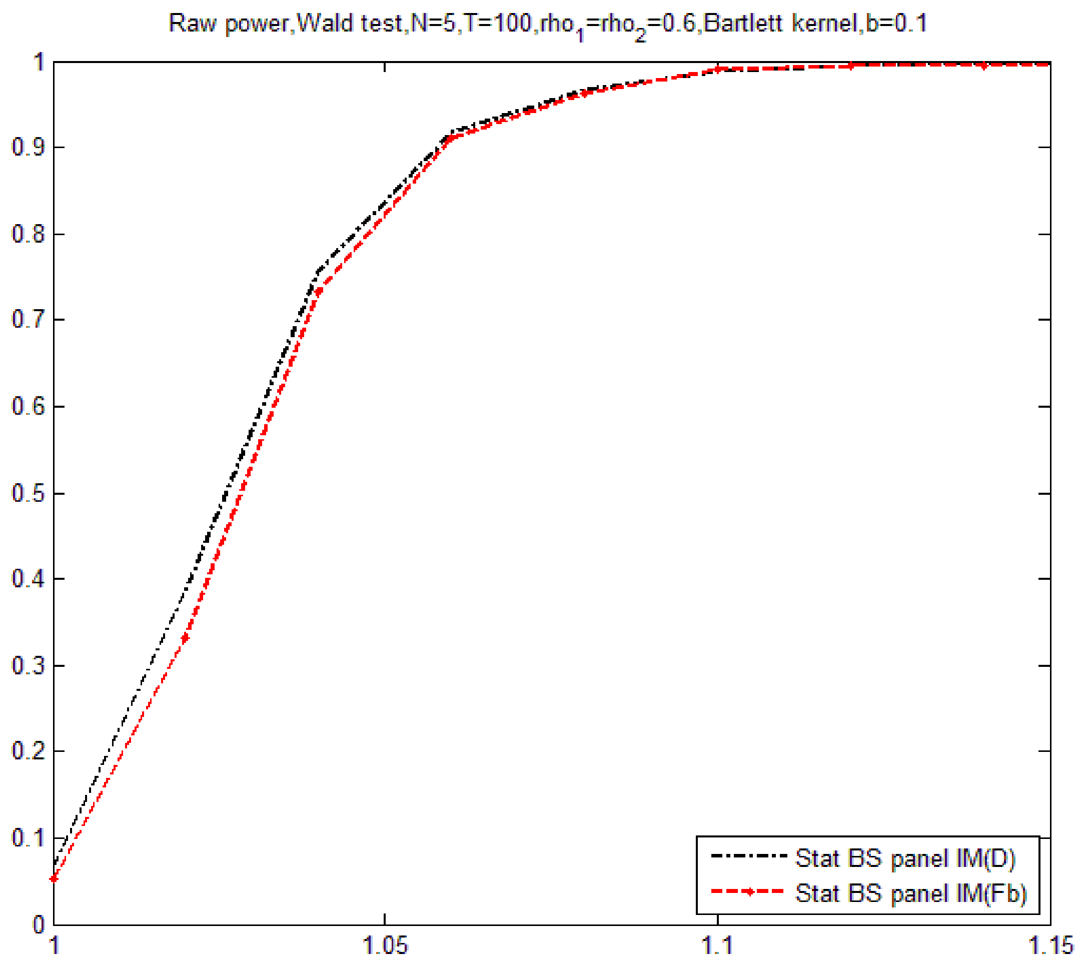


Figure 2.10: Raw power, Wald test, $N = 5$, $T = 100$, $\rho_1 = \rho_2 = 0.6$, $b = 0.1$, Bartlett kernel

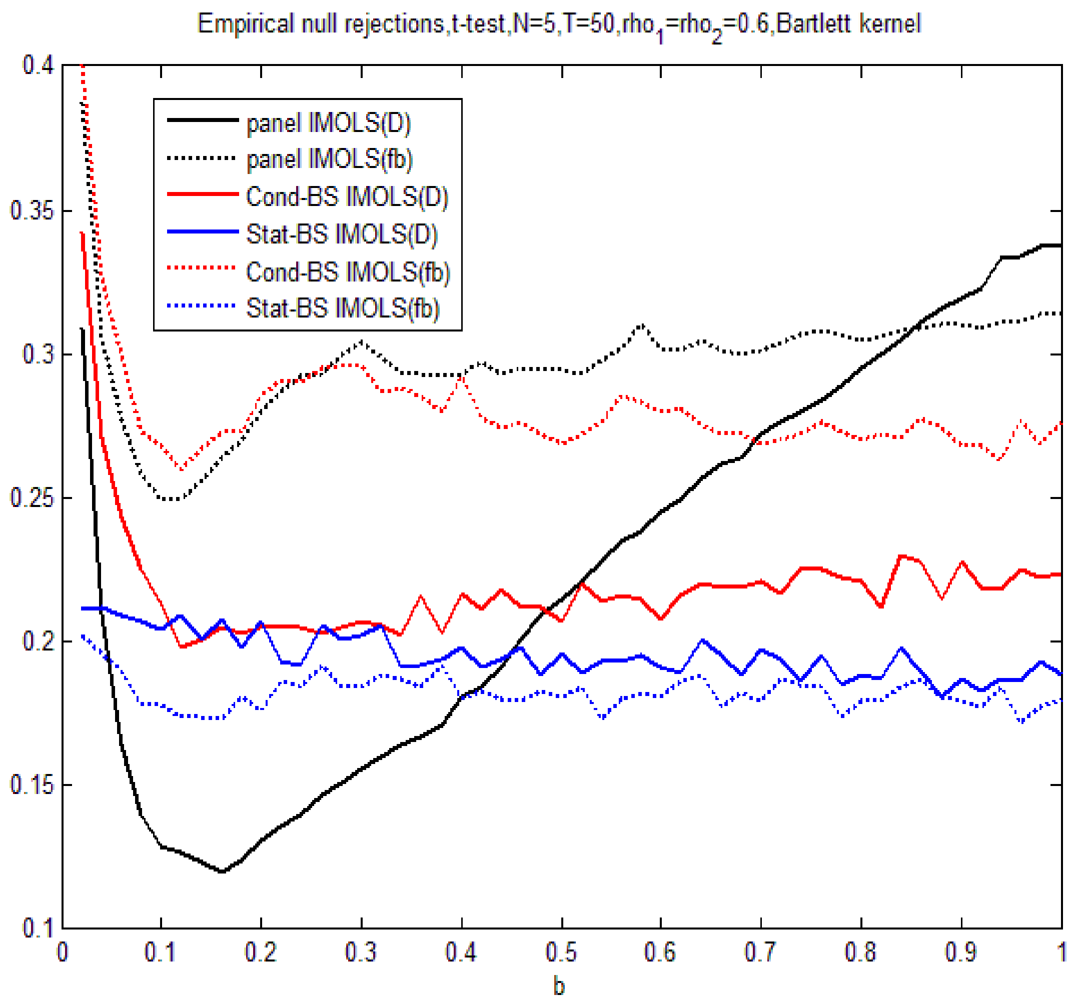


Figure 2.11: Empirical null rejections, t-test, $N = 5$, $T = 50$, $\rho_1 = \rho_2 = 0.6$, Bartlett kernel, residuals from non-augmented partial sum regression

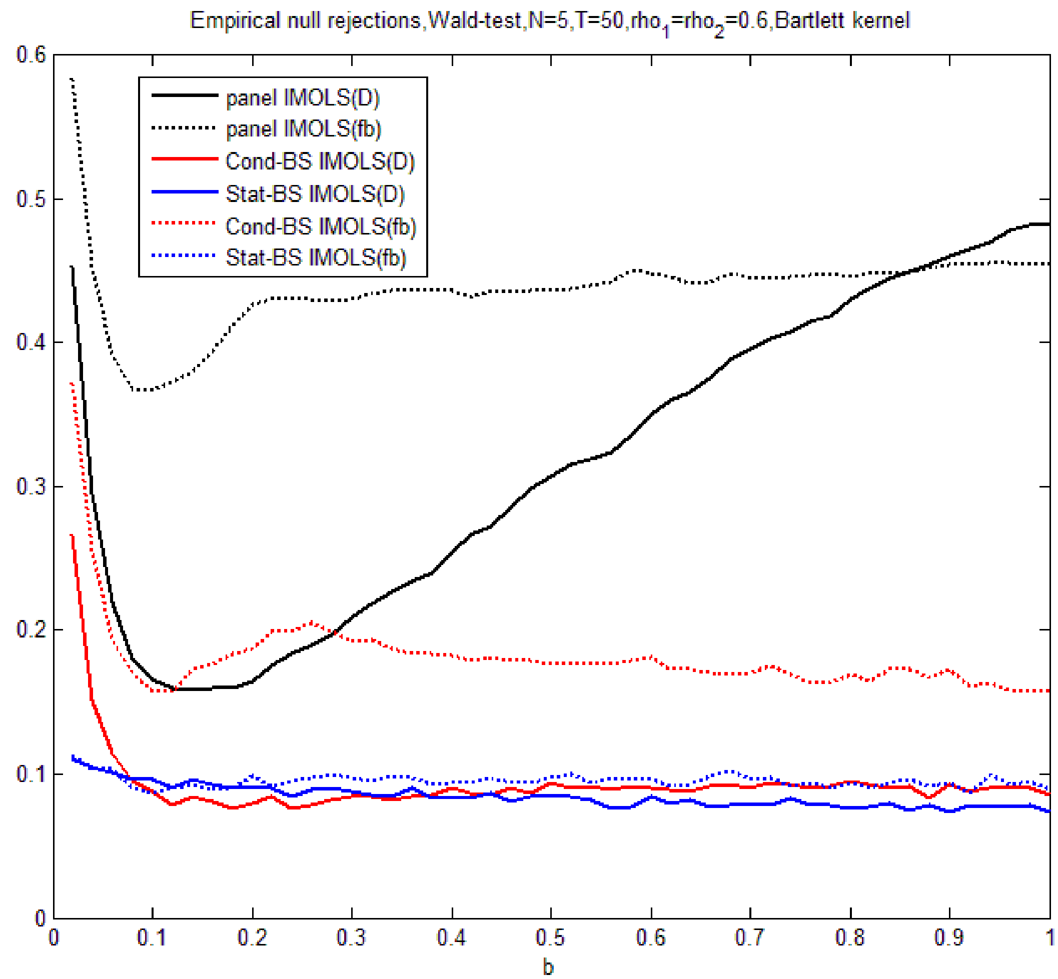


Figure 2.12: Empirical null rejections, Wald test, $N = 5$, $T = 50$, $\rho_1 = \rho_2 = 0.6$, Bartlett kernel, residuals from non-augmented partial sum regression

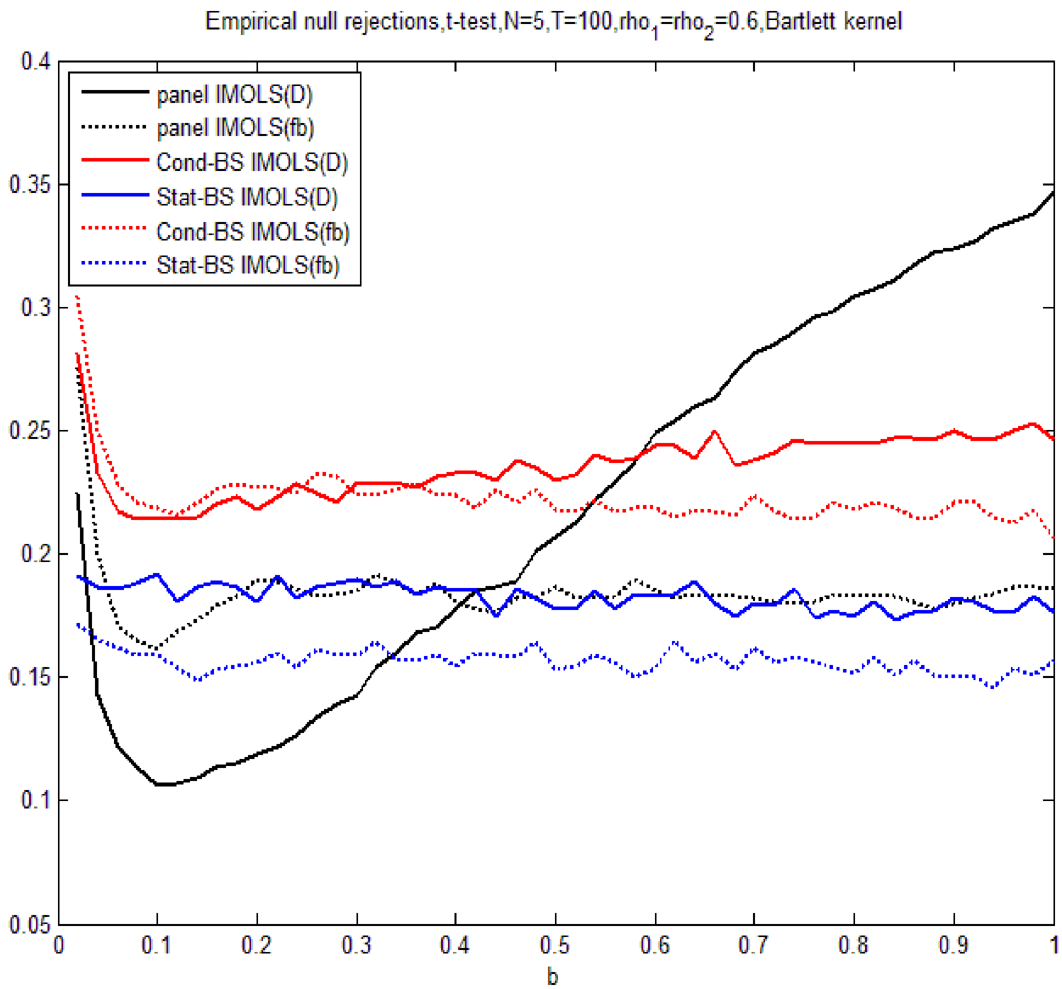


Figure 2.13: Empirical null rejections, t-test, $N = 5$, $T = 100$, $\rho_1 = \rho_2 = 0.6$, Bartlett kernel, residuals from non-augmented partial sum regression

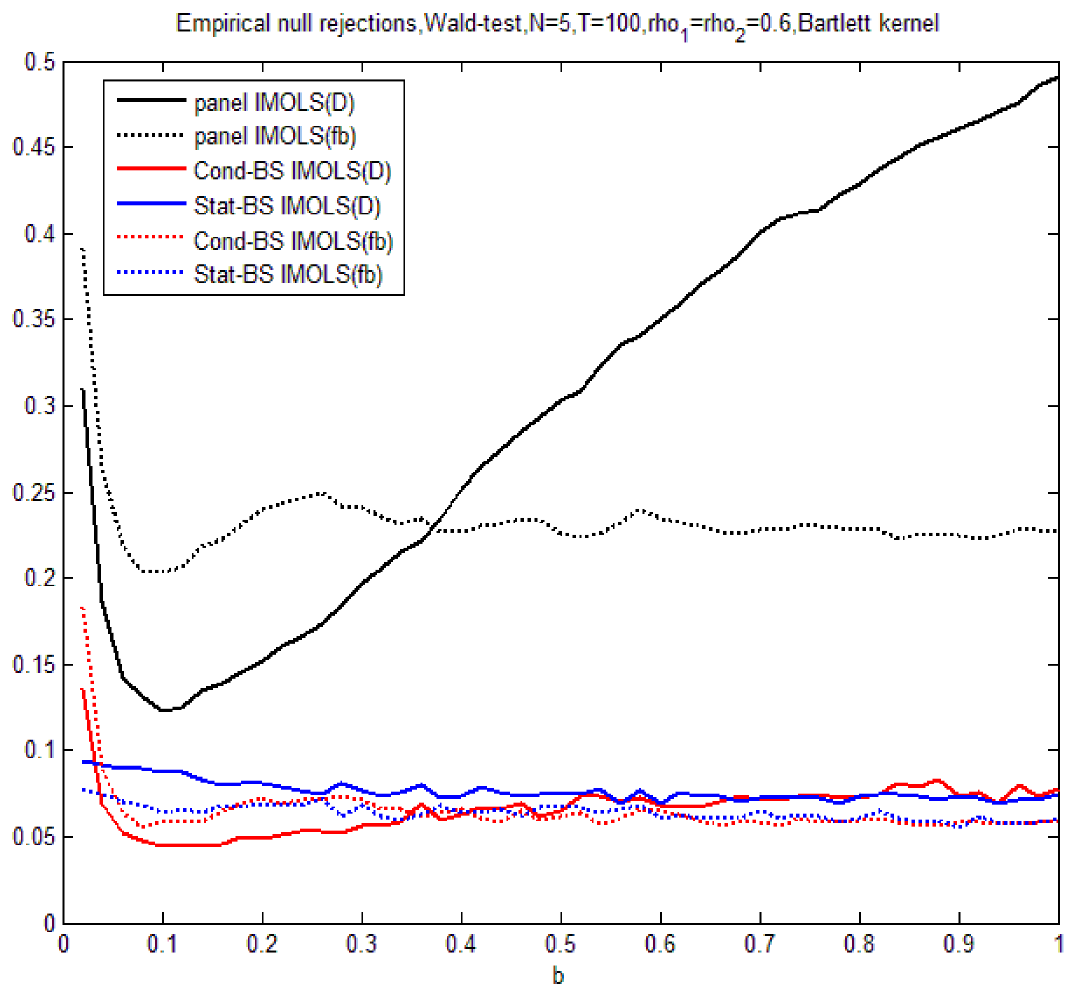


Figure 2.14: Empirical null rejections, Wald test, $N = 5$, $T = 100$, $\rho_1 = \rho_2 = 0.6$, Bartlett kernel, residuals from non-augmented partial sum regression

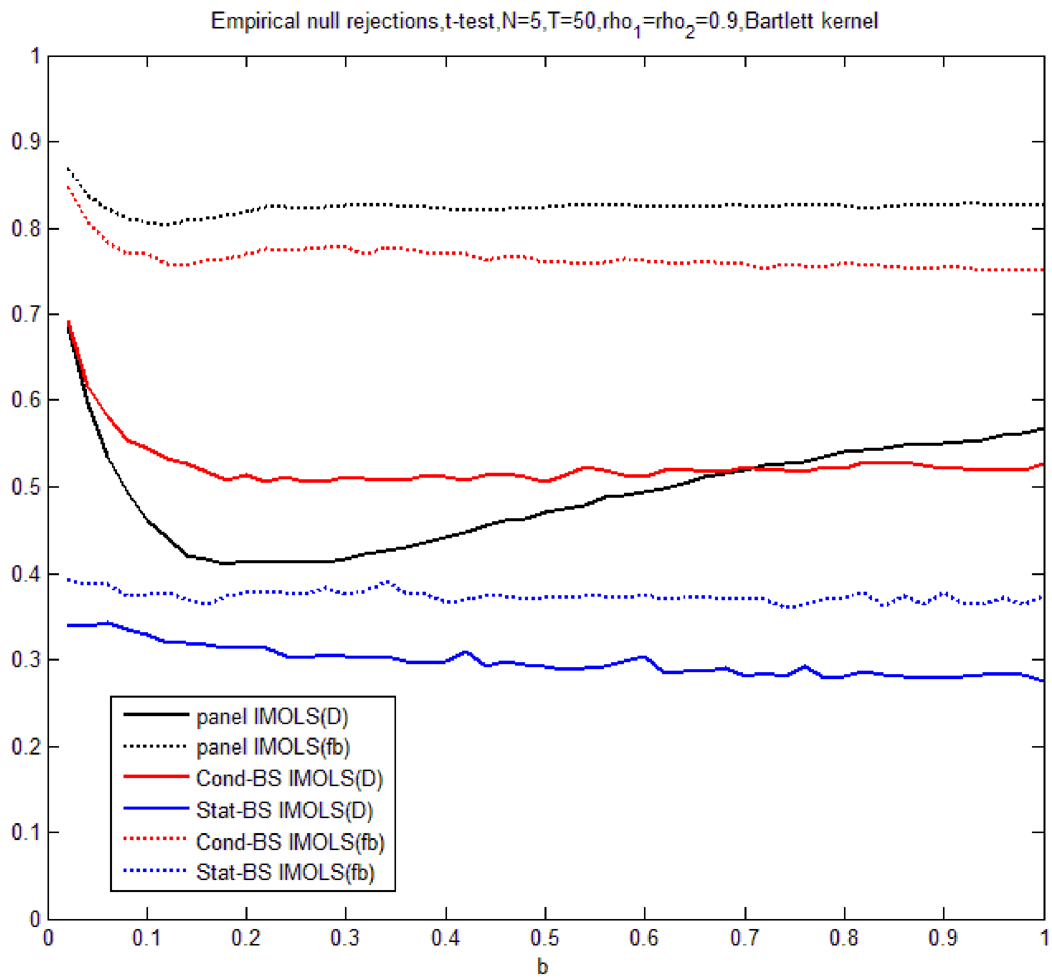


Figure 2.15: Empirical null rejections, t-test, $N = 5$, $T = 50$, $\rho_1 = \rho_2 = 0.9$, Bartlett kernel, residuals from non-augmented partial sum regression

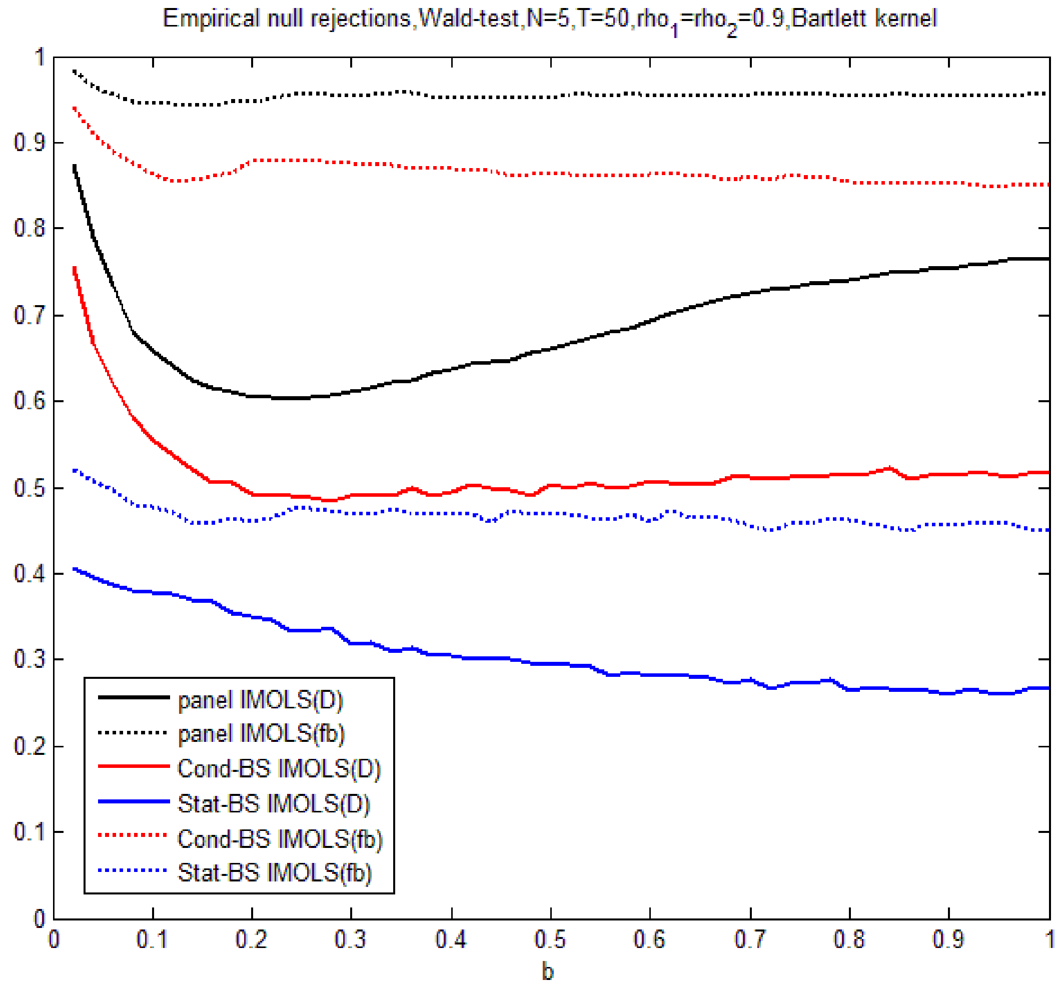


Figure 2.16: Empirical null rejections, Wald test, $N = 5$, $T = 50$, $\rho_1 = \rho_2 = 0.9$, Bartlett kernel, residuals from non-augmented partial sum regression

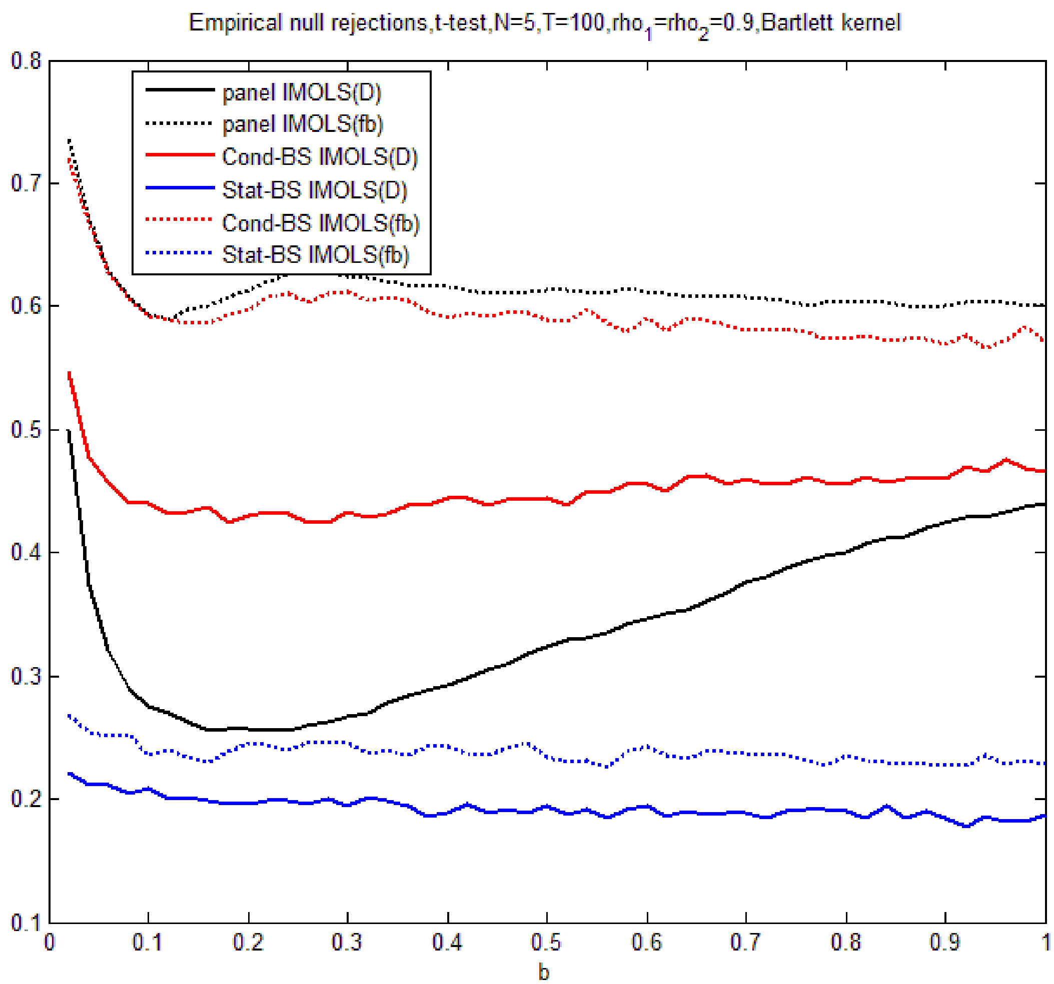


Figure 2.17: Empirical null rejections, t-test, $N = 5$, $T = 100$, $\rho_1 = \rho_2 = 0.9$, Bartlett kernel, residuals from non-augmented partial sum regression

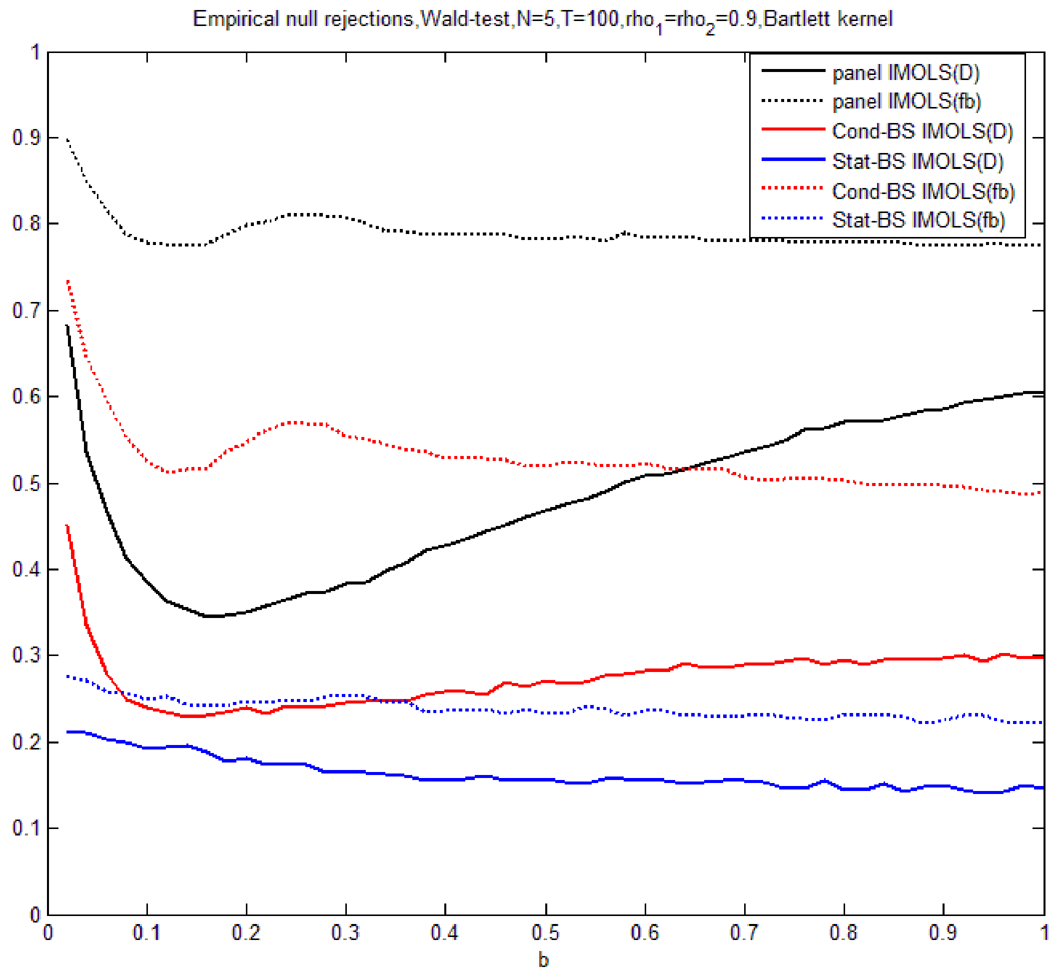


Figure 2.18: Empirical null rejections, Wald test, $N = 5$, $T = 100$, $\rho_1 = \rho_2 = 0.9$, Bartlett kernel, residuals from non-augmented partial sum regression

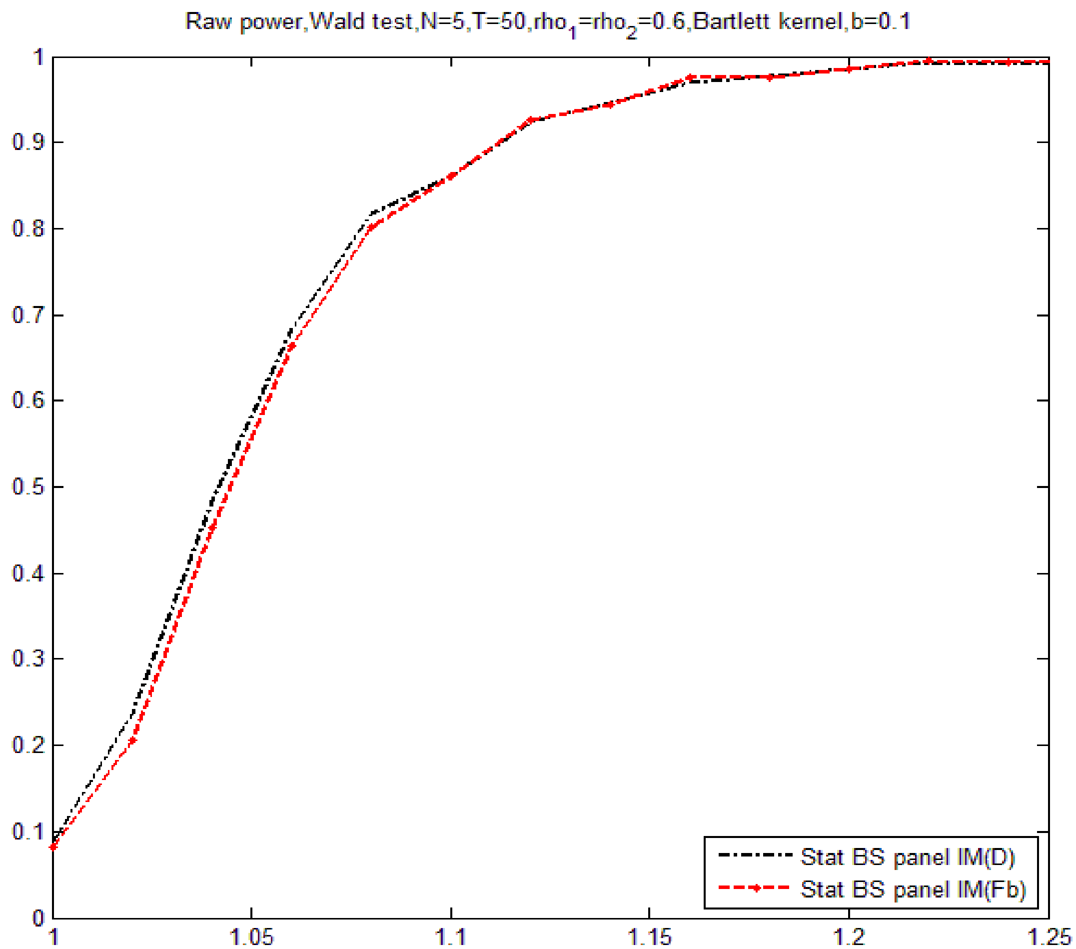


Figure 2.19: Raw power, Wald test, $N = 5$, $T = 50$, $\rho_1 = \rho_2 = 0.6$, $b = 0.1$, Bartlett kernel, residuals from non-augmented partial sum regression

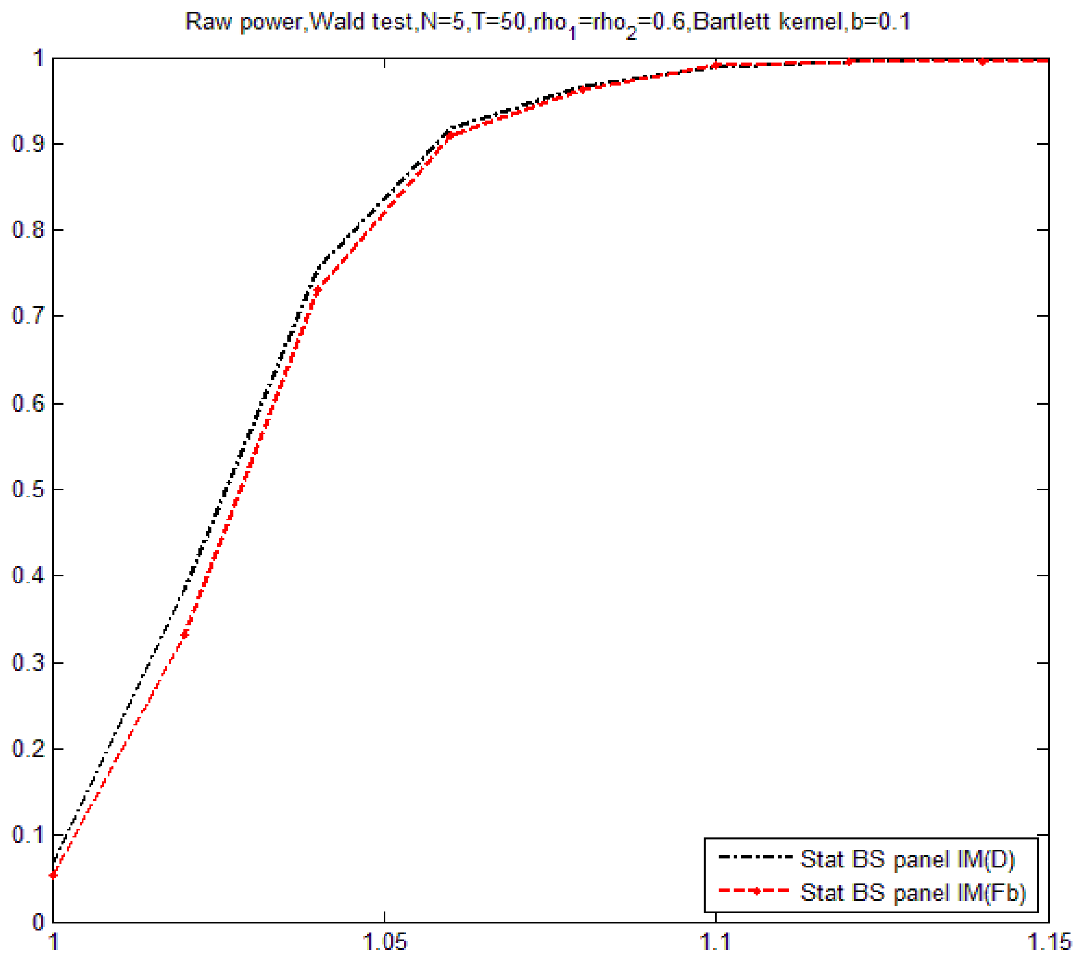


Figure 2.20: Raw power, Wald test, $N = 5$, $T = 100$, $\rho_1 = \rho_2 = 0.6$, $b = 0.1$, Bartlett kernel, residuals from non-augmented partial sum regression

# Chapter 3

# Estimation and Inference for Heterogeneous Cointegrated Panels with Limited Cross Sectional Dependence

This paper is concerned with parameter estimation and inference in a panel cointegrating regression with endogenous regressors and heterogeneous long run variances in the cross section. In addition, the model allows a limited degree of cross-sectional dependence due to a common time effect. The estimator is labeled as panel integrated modified ordinary least squares (panel IM-OLS). Similar to panel fully modified OLS (panel FM-OLS) and panel dynamic OLS (panel DOLS), the panel IM-OLS estimator has a zero mean Gaussian mixture limiting distribution. However, standard asymptotic inference is infeasible due the existence of nuisance parameters. Inference based on panel IM-OLS relies on the stationary bootstrap. The properties of panel IM-OLS are analyzed using the stationary bootstrap in finite sample simulations.

## 3.1 Introduction

In the past decade, panel cointegration methods have drawn much attention in empirical research. The attractive feature of panel cointegration methods is that they permit investigation of the long-run relationship among nonstationary variables more efficiently than using time series data alone. However, panel cointegration is more complicated than single time-series cointegration when cross-sectional dependence and heterogeneity exist. If the cross-sectional dependence and heterogeneity were ignored, it might lead to poor inference and inconsistent estimators. It is well known that the application of the first generation panel unit root tests, which generally assume cross-sectional independence, to the series with cross-sectional correlation leads to size distortion and low power. This might also be the case for the panel cointegration estimation and testing. For example, Westerlund and Edgerton [2008] claim that the tests of McCoskey and Kao [1998], Pedroni [1999], [2004] and Westerlund [2005] all require independence among the cross-sectional units, and their size properties become suspect when this assumption does not hold. The homogeneity assumption is often not well supported by the data. Therefore a framework that allows potential heterogeneity is necessary.

In the panel cointegrated regression literature, the panel fully modified OLS (panel FM-OLS) and the panel dynamic OLS (panel DOLS) methods are the most popular methods (see Kao and Chiang [2000], Pedroni [2000], Bai et al. [2009] and Mark and Sul [2003]). They are the extensions of the single time series fully modified OLS (FM-OLS) and dynamic OLS (DOLS). Integrated modified OLS (IM-OLS), proposed by Vogelsang and Wagner [2014], provides a fully parametric and computationally convenient alternative to the FM-OLS and the DOLS estimators. Vogelsang et al. [2016] extend IM-OLS to panel data models with individual dummies and homogeneous second moment structure. The present paper considers an extension of Vogelsang et al. [2016] by allowing time dummies and heterogeneous variance structure in the model. The benefit of adding time dummies is twofold. First, time dummies can handle deterministic components and common factor shocks, and second, time dummies make the model robust to limited degrees of cross-sectional dependence. Allowing heterogeneous, rather than homogeneous, variance structure makes the framework discussed in this paper more applicable in empirical research. Bai et al. [2009] and Mark and Sul [2003] consider similar problems using the panel FM-OLS and the panel DOLS estimators.

The limit theory considered here is obtained for a fixed number of cross-sectional units $N$, letting the number of the time periods, $T$, go to infinity. The setting of $N$ fixed and $T \to \infty$ is widely used in empirical macroeconomics, empirical energy economics and empirical finance problems (see Christopoulos and Tsionas [2004], Lee [2005], Apergis and Payne [2009], Narayan and Smyth [2008] and Canzoneri et al. [1999]). Under this scenario, even though the panel IM-OLS estimator converges to a zero mean Gaussian mixture distribution, asymptotic inference is complicated by the presence of nuisance parameters. One way to implement valid hypothesis tests is using bootstrap methods. Although bootstrap methods are widely employed for analyzing nonstationary time series data, e.g. bootstrap unit root tests and bootstrap cointegration tests (see Chang [2004], Paparoditis and Politis [2003], Parker et al. [2006], Westerlund and Edgerton [2007]), surprisingly few papers are devoted to bootstrap inference in cointegrated regressions. Psaradakis [2001] and Chang et al. [2006] employ the sieve bootstrap procedure to cointegrated regressions. Li and Maddala [1997] and Shin and Hwang [2013] apply the stationary bootstrap to cointegrated regression.

In the literature, the sieve bootstrap can be applied in fairly general models and performs well in pure time series setting, but it requires fitting a finite order AR or VAR model to the errors. Smeekes and Urbain [2014] questioned the validity of the use of VAR sieve bootstrap in panels with a moderate cross-sectional dimension and showed that the AR sieve bootstrap might be misleading when cross-sectional dependence is present. On the contrary, the stationary bootstrap requires no parametric structure for drawing bootstrap samples. In addition, a working paper by Li [2016] shows that the stationary bootstrap performs well in panel cointegrated regressions with fixed effects and homogeneous variance structure when cross sectional units are uncorrelated. The bootstrap method used in the present paper is the stationary bootstrap.

The rest of the paper is organized as follows. Section 3.2 introduces the model, assumptions and the panel IM-OLS estimator. In Section 3.3 asymptotic inference and stationary bootstrap inference are presented. Section 3.4 provides a Monte Carlo simulation to investigate the finite sample properties of the proposed bootstrap test. Section 3.5 summarizes the results and concludes the paper. All proofs are collected in Appendices C - F.

## 3.2 Model set up and estimation

### 3.2.1 The model and assumptions

Consider the following panel data model

$$
\begin{aligned}
y_{it} &= \alpha_i + x_{it}'\beta + e_{it} && (3.1) \\
x_{it} &= x_{it-1} + v_{it} && (3.2)
\end{aligned}
$$

where $i = 1, 2, \cdots, N$ and $t = 1, 2, \cdots, T$ index the cross-sectional and time series units, respectively; $y_{it}$, $\alpha_i$ and $e_{it}$ are scalars; $x_{it}$, $\beta$ and $v_{it}$ are $k \times 1$ vectors. The regressor, $x_{it}$, is potentially endogenous for each individual $i$.

**Assumption 6.** *Assume that the error term, $e_{it}$, follows a special case of a factor model*

$$e_{it} = F_t'\lambda + u_{it}. \tag{3.3}$$

In Assumption 6, $F_t$ is the common factor and $u_{it}$ is idiosyncratic component. Assumption 6 is a special case of factor model, as the factor loading $\lambda$ is constant across $i$. Under the above assumption, $e_{it}$ and $e_{jt}$ are correlated due to the common factor $F_t$, therefore the panel data model is cross-sectional dependent. Because the regressor considered here is endogenous, $v_{it}$ is assumed to be correlated with $u_{it}$. In addition, there is no restriction on the correlation between $v_{it}$ and $F_t$.

Define the error vector as $\eta_{it} = \begin{bmatrix} u_{it} & v_{it}' \end{bmatrix}'$ and suppose that it is a $(k+1)$ dimensional stationary vector for each $i$. In addition, assume that $F_t$ is a $I(0)$ process. This implies that the model introduced in (3.1) describes a system of panel cointegrated regressions, i.e. $y_{it}$ is cointegrated with $x_{it}$. It might also be interesting to consider the case that $F_t$ is a $I(1)$ process. Bai et al. [2009] consider the CupBC (continuously-updated and bias-corrected) and the CupFM (continuously-updated and fully-modified) estimators for panel cointegration models with cross-sectional dependence generated by unobserved global stochastic trends, where $F_t$ is non-stationary. In this paper, the interest is in estimation and inference about $\beta$ based on the panel IM-OLS estimator when $F_t$ is stationary. In order to derive the panel IM-OLS estimator's limiting distribution, a second assumption is sufficient.

**Assumption 7.** *Assume that $\eta_{it}$ is independent across $i$, and satisfies the Functional CLT*

$$T^{-\frac{1}{2}} \sum_{t=1}^{[rT]} \eta_{it} \Rightarrow B_i(r) = \begin{bmatrix} B_{u,i}(r) \\ B_{v,i}(r) \end{bmatrix} = \Omega_i^{1/2} W_i(r),$$

*where $r \in (0,1]$, and $[rT]$ denotes the largest integer value of $rT$.*

In Assumption 7, $\Omega_i^{1/2}$ is a $(k+1) \times (k+1)$ matrix that satisfies $\Omega_i = \Omega_i^{1/2} \left(\Omega_i^{1/2}\right)'$ where

$$\Omega_i = \sum_{j=-\infty}^{\infty} \mathbb{E}\left(\eta_{it}\eta_{it-j}'\right) = \begin{bmatrix} \Omega_{uu,i} & \Omega_{uv,i} \\ \Omega_{vu,i} & \Omega_{vv,i} \end{bmatrix} > 0,$$

where it is obvious that $\Omega_{uv,i} = \Omega_{vu,i}'$. Assume that $\Omega_{vv,i}$ is non-singular, which implies that $\{x_{it}\}$ are not cointegrated among themselves. Partition $B_i(r)$ as $B_i(r) = \begin{bmatrix} B_{u,i}(r) & B_{v,i}'(r) \end{bmatrix}'$, and likewise partition $W_i(r)$ as $W_i(r) = \begin{bmatrix} w_{u,i}(r) & W_{v,i}'(r) \end{bmatrix}'$, where $w_{u,i}(r)$ and $W_{v,i}(r)$ are a scalar and a $k$-dimensional standard Brownian motion, respectively. Using the Cholesky form of $\Omega_i^{1/2}$,

$$\Omega_i^{1/2} = \begin{bmatrix} \sigma_{u \cdot v,i} & \lambda_{uv,i} \\ 0_{k \times 1} & \Omega_{vv,i}^{1/2} \end{bmatrix},$$

it can be shown that $\sigma_{u \cdot v,i}^2 = \Omega_{uu,i} - \Omega_{uv,i}\Omega_{vv,i}^{-1}\Omega_{vu,i}$ and $\lambda_{uv,i} = \Omega_{uv,i}\left(\Omega_{vv,i}^{-1/2}\right)'$. In addition, it follows that

$$B_i(r) = \begin{bmatrix} B_{u,i}(r) \\ B_{v,i}(r) \end{bmatrix} = \begin{bmatrix} \sigma_{u \cdot v,i} w_{u,i}(r) + \lambda_{uv,i} W_{v,i}(r) \\ \Omega_{vv,i}^{1/2} W_{v,i}(r) \end{bmatrix}.$$

Note that $\lambda_{uv,i} \neq 0$ for all $i$ because the regressors are allowed to be endogenous. Notice that the 2nd order moment structure is heterogeneous across $i$.

### 3.2.2 Panel IM-OLS estimator

From regression (3.1) and Assumption 1, the system can be rewritten as

$$y_{it} = \alpha_i + x_{it}'\beta + F_t'\lambda + u_{it},$$

and its cross-sectional mean is given by

$$\bar{y}_t = \bar{\alpha} + \bar{x}_t'\beta + F_t'\lambda + \bar{u}_t,$$

where

$$\begin{aligned}
\bar{y}_t &= \frac{1}{N}\sum_{i=1}^{N} y_{it} \\
\bar{\alpha} &= \frac{1}{N}\sum_{i=1}^{N} \alpha_i \\
\bar{x}_t &= \frac{1}{N}\sum_{i=1}^{N} x_{it} \\
\bar{u}_t &= \frac{1}{N}\sum_{i=1}^{N} u_{it}.
\end{aligned}$$

Cross-sectional demeaning can be used to remove $F_t'\lambda$ and provides an estimation equation that is exactly invariant to $F_t'\lambda$. Note that the cross-sectional demeaning is exactly the same as including time period dummies and projecting them out of the regression. Since $F_t$ could be unobserved time shock, therefore projecting it out before partial summing is crucial. Cross-sectional demeaning gives

$$y_{it} - \bar{y}_t = \alpha_i - \bar{\alpha} + \left(x_{it}' - \bar{x}_t'\right) \beta + u_{it} - \bar{u}_t, \tag{3.4}$$

which is denoted as

$$\ddot{y}_{it} = \mu_i + \ddot{x}_{it}'\beta + \ddot{u}_{it}, \tag{3.5}$$

where

$$\begin{aligned}
\mu_i &= \alpha_i - \bar{\alpha} \\
\ddot{y}_{it} &= y_{it} - \bar{y}_t \\
\ddot{x}_{it} &= x_{it} - \bar{x}_t \\
\ddot{u}_{it} &= u_{it} - \bar{u}_t.
\end{aligned}$$

Following Vogelsang and Wagner [2014], compute the partial sum of regression (3.5) to give

$$S_{it}^{\ddot{y}} = t\mu_i + S_{it}^{\ddot{x}\prime}\beta + S_{it}^{\ddot{u}}, \tag{3.6}$$

where

$$\begin{aligned}
S_{it}^{\ddot{y}} &= \sum_{j=1}^{t} \ddot{y}_{ij} \\
S_{it}^{\ddot{x}} &= \sum_{j=1}^{t} \ddot{x}_{ij} \\
S_{it}^{\ddot{u}} &= \sum_{j=1}^{t} \ddot{u}_{ij}.
\end{aligned}$$

In order to deal with the endogeneity problem generated by the correlation between $u_{it}$ and $v_{it}$, it is sufficient to add additional regressors into regression (3.6). A natural candidate is the demeaned regressor $\ddot{x}_{it}$, however, this does not work due to the heterogeneity in the model. This is formally shown in the Appendix. The endogeneity problem, which is complicated by the heterogeneity in the variance structure, can be solved by adding the decomposed $\ddot{x}_{it}$ into (3.6). The decomposed $\ddot{x}_{it}$ can be expressed as

$$\frac{-1}{N}x_{1t},\ \cdots \frac{-1}{N}x_{i-1,t},\ \frac{N-1}{N}x_{it},\ \frac{-1}{N}x_{i+1,t},\ \cdots\ \frac{-1}{N}x_{Nt}.$$

Adding these regressors separately will overcome the heterogeneous variance problem when dealing with endogeneity. Details are given in the Appendix.

**Remark 4.** *In regression (3.5), if $\beta$ is the only parameter of interest, then it is possible to demean across time to remove $\mu_i$ before partial summing. That is*

$$\ddot{y}_{it}^{+} = \ddot{x}_{it}^{+\prime}\beta + \ddot{u}_{it}^{+},$$

*where*

$$\ddot{y}_{it}^{+} = \ddot{y}_{it} - \frac{1}{T}\sum_{k=1}^{T}\ddot{y}_{ik},\ \ \ddot{x}_{it}^{+} = \ddot{x}_{it} - \frac{1}{T}\sum_{k=1}^{T}\ddot{x}_{ik},\ \ \ddot{u}_{it}^{+} = \ddot{u}_{it} - \frac{1}{T}\sum_{k=1}^{T}\ddot{u}_{ik}.$$

*Then regression (3.6) becomes to*

$$S_{it}^{\ddot{y}^{+}} = S_{it}^{\ddot{x}^{+}\prime}\beta + S_{it}^{\ddot{u}^{+}},$$

*where*

$$S_{it}^{\ddot{y}^{+}} = \sum_{j=1}^{t}\ddot{y}_{ij}^{+},\ \ S_{it}^{\ddot{x}^{+}} = \sum_{j=1}^{t}\ddot{x}_{ij}^{+},\ \ S_{it}^{\ddot{u}^{+}} = \sum_{j=1}^{t}\ddot{u}_{ij}^{+}.$$

*However, in this case, including the components of $\ddot{x}_{it}$ is not sufficient to deal with the endogeneity problem. Finding the additional regressors for this partial sum regression is much more challenging if not impossible, therefore this method is not considered in this paper.*

**Remark 5.** *If the system does have homogeneous 2nd order moment structures, i.e. $\Omega_i = \Omega_j$ for any $i$, $j$ for $\{1, 2, \cdots, N\}$, then adding $\ddot{x}_{it}$ to the partial sum regression will be sufficient for solving the endogeneity problem.*

Including the additional regressors in (3.6) gives

$$S_{it}^{\ddot{y}} = t\mu_i + S_{it}^{\ddot{x}\prime}\beta + \frac{N-1}{N}x_{it}'\gamma_i + \left(\frac{-1}{N}\sum_{j=1,j\neq i}^{N} x_{jt}'\gamma_j\right) + S_{it}^{\tilde{u}} \tag{3.7}$$

where

$$S_{it}^{\tilde{u}} = S_{it}^{\ddot{u}} - \frac{N-1}{N}x_{it}'\gamma_i + \frac{1}{N}\sum_{j=1,j\neq i}^{N} x_{jt}'\gamma_j.$$

Stacking all time periods and all individuals' data together, the matrix form of the system is given by

$$S^{\ddot{y}} = S^{\ddot{x}}\theta + S^{\ddot{u}}, \tag{3.8}$$

where

$$S^{\ddot{y}} = \begin{bmatrix} S^{\ddot{y}}_{11} \\ \vdots \\ S^{\ddot{y}}_{1T} \\ \vdots \\ S^{\ddot{y}}_{N1} \\ \vdots \\ S^{\ddot{y}}_{NT} \end{bmatrix}, \ \theta = \begin{bmatrix} \beta \\ \gamma_1 \\ \vdots \\ \gamma_N \\ \mu_1 \\ \vdots \\ \mu_N \end{bmatrix}, \ S^{\ddot{u}} = \begin{bmatrix} S^{\tilde{u}}_{11} \\ \vdots \\ S^{\tilde{u}}_{1T} \\ \vdots \\ S^{\tilde{u}}_{N1} \\ \vdots \\ S^{\tilde{u}}_{NT} \end{bmatrix},$$

$$S^{\ddot{x}} = \begin{bmatrix} S^{\ddot{x}\prime}_{11} & \frac{N-1}{N}x'_{11} & \frac{-1}{N}x'_{21} & \cdots & \frac{-1}{N}x'_{N1} & 1 & \cdots & 0 \\ S^{\ddot{x}\prime}_{12} & \frac{N-1}{N}x'_{12} & \frac{-1}{N}x'_{22} & \cdots & \frac{-1}{N}x'_{N2} & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ S^{\ddot{x}\prime}_{1T} & \frac{N-1}{N}x'_{1T} & \frac{-1}{N}x'_{2T} & \cdots & \frac{-1}{N}x'_{NT} & T & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ S^{\ddot{x}\prime}_{N1} & \frac{-1}{N}x'_{11} & \frac{-1}{N}x'_{21} & \cdots & \frac{N-1}{N}x'_{N1} & 0 & \cdots & 1 \\ S^{\ddot{x}\prime}_{N2} & \frac{-1}{N}x'_{12} & \frac{-1}{N}x'_{22} & \cdots & \frac{N-1}{N}x'_{N2} & 0 & \cdots & 2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ S^{\ddot{x}\prime}_{NT} & \frac{-1}{N}x'_{1T} & \frac{-1}{N}x'_{2T} & \cdots & \frac{N-1}{N}x'_{NT} & 0 & \cdots & T \end{bmatrix}.$$

The panel IM-OLS estimator is the OLS estimator of regression (3.8), which is given by

$$\hat{\theta} = \left(S^{\ddot{x}\prime}S^{\ddot{x}}\right)^{-1}\left(S^{\ddot{x}\prime}S^{\ddot{y}}\right).$$

It follows that

$$\begin{aligned}\hat{\theta}-\theta &= \left(S^{\ddot{x}\prime}S^{\ddot{x}}\right)^{-1}\left(S^{\ddot{x}\prime}S^{\ddot{u}}\right)\\ &= \left(\sum_{i=1}^{N}\sum_{t=1}^{T} q_{it}q_{it}'\right)^{-1}\left(\sum_{i=1}^{N}\sum_{t=1}^{T} q_{it}S_{it}^{\tilde{u}}\right)\end{aligned}$$

where

$$\begin{aligned}q_{1t} &= \begin{bmatrix} S_{1t}^{\ddot{x}\prime} & \frac{N-1}{N}x_{1t}' & \frac{-1}{N}x_{2t}' & \cdots & \frac{-1}{N}x_{Nt}' & t & 0 & \cdots & 0\end{bmatrix}'\\ q_{2t} &= \begin{bmatrix} S_{2t}^{\ddot{x}\prime} & \frac{-1}{N}x_{1t}' & \frac{N-1}{N}x_{2t}' & \cdots & \frac{-1}{N}x_{Nt}' & 0 & t & \cdots & 0\end{bmatrix}'\\ \vdots &\ \vdots \qquad\qquad \vdots\\ q_{Nt} &= \begin{bmatrix} S_{Nt}^{\ddot{x}\prime} & \frac{-1}{N}x_{1t}' & \frac{-1}{N}x_{2t}' & \cdots & \frac{N-1}{N}x_{Nt}' & 0 & 0 & \cdots & t\end{bmatrix}'.\end{aligned}$$

Define the scaling matrix

$$A_{PIM}^{-1} = \begin{bmatrix} T\cdot I_k & & 0\\ & I_N\otimes I_k & \\ 0 & & I_N\otimes T^{\frac{1}{2}}\end{bmatrix}$$

as a $(k+Nk+N)\times(k+Nk+N)$ diagonal matrix.

The following theorem gives the asymptotic distribution of the panel IM-OLS estimator.

**Theorem 3.** *Assume that the data are generated by (3.1) and (3.2), and that Assumptions 6 and 7 hold. Define $\theta$ by stacking $\beta$, $\gamma_i$ and $\mu_i$. Then for fixed $N$, as $T\to\infty$*

$$
\begin{aligned}
A_{PIM}^{-1}\left(\hat{\theta}-\theta\right) &= \begin{bmatrix} T\left(\hat{\beta}-\beta\right) \\ \left(\hat{\gamma}_1-\gamma_1\right) \\ \vdots \\ \left(\hat{\gamma}_N-\gamma_N\right) \\ \sqrt{T}\left(\hat{\mu}_1-\mu_1\right) \\ \vdots \\ \sqrt{T}\left(\hat{\mu}_N-\mu_N\right) \end{bmatrix} \\
&= \left(T^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}A_{1T}^{-1}q_{it}q_{it}'A_{1T}^{-1}\right)^{-1}\left(T^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}A_{1T}^{-1}q_{it}T^{-\frac{1}{2}}S_{it}^{\tilde{u}}\right) \\
&\Rightarrow \left(\int_0^1\sum_{i=1}^{N}h_i(r)h_i'(r)dr\right)^{-1}\times \\
&\quad \left(\int_0^1\sum_{i=1}^{N}\left\{h_i(r)\left[\sigma_{u\cdot v,i}w_{u,i}(r)-\frac{1}{N}\sum_{j=1}^{N}\sigma_{u\cdot v,j}w_{u,j}(r)\right]\right\}dr\right) \\
&= \Psi
\end{aligned}
$$

*where*

$$h_i(r) = \begin{bmatrix} \int_0^r \left[ \Omega_{vv,i}^{\frac{1}{2}} W_{v,i}(s) - \frac{1}{N} \sum_{j=1}^{N} \Omega_{vv,j}^{\frac{1}{2}} W_{v,j}(s) \right] ds \\ \frac{-1}{N} \Omega_{vv,1}^{\frac{1}{2}} W_{v,1}(r) \\ \vdots \\ \frac{N-1}{N} \Omega_{vv,i}^{\frac{1}{2}} W_{v,i}(r) \\ \vdots \\ \frac{-1}{N} \Omega_{vv,N}^{\frac{1}{2}} W_{v,N}(r) \\ 0 \\ \vdots \\ r \\ \vdots \\ 0 \end{bmatrix}.$$

Conditional on $h_i(r)$ for $i = 1, 2, \cdots, N$, it can be shown that $\Psi \sim N(0, V_{PIM})$, where $V_{PIM}$ is given by

$$\begin{aligned} V_{PIM} &= \left( \int_0^1 \sum_{i=1}^{N} h_i(r) h_i'(r) dr \right)^{-1} \times \\ & \left[ \sum_{i=1}^{N} \sigma_{u \cdot v,i}^2 \int_0^1 \left[ \ddot{H}_i(1) - \ddot{H}_i(r) \right] \left[ \ddot{H}_i(1) - \ddot{H}_i(r) \right]' dr \right] \times \\ & \left( \int_0^1 \sum_{i=1}^{N} h_i(r) h_i'(r) dr \right)^{-1} \end{aligned}$$

and

$$\ddot{H}_i(r) = H_i(r) - \frac{1}{N} \sum_{j=1}^{N} H_j(r).$$

The derivation of this conditional variance is given in the Appendix.

## 3.3 Inference about $\theta$

### 3.3.1 Inference using panel IM-OLS

This section provides a discussion of hypothesis testing using the panel IM-OLS estimator. In particular, the hypothesis being considered is given by

$$H_0: \ R\theta = r$$

where $R \in \mathbb{R}^{q \times (k+Nk+N)}$ with full rank $q$ and $r \in \mathbb{R}^q$. Because the vector $\hat{\theta}$ has elements that converge at different rates, restrictions on $R$ are necessary. Assume that there exists a non-singular $q \times q$ matrix $A_R$ such that

$$\underset{T\to\infty}{lim} A_R^{-1} R A_{PIM} = R^*$$

with $R^*$ has rank $q$.

In order to carry out statistical inference, the asymptotic variance, $V_{PIM}$, needs to be estimated. The outside parts of the sandwich form can be estimated by

$$\left( T^{-2} \sum_{i=1}^{N} \sum_{t=1}^{T} A_{PIM} q_{it} q_{it}' A_{PIM} \right)^{-1} .$$

The tricky part is estimating the middle part of the sandwich form of the variance. Suppose that $\breve{\sigma}^2_{u \cdot v,i}$ is an estimator for $\sigma^2_{u \cdot v,i}$, then an estimator for the middle part of the variance is given by

$$T^{-4}\sum_{t=1}^{T}\sum_{i=1}^{N}\breve{\sigma}_{u\cdot v,i}^{2}\left(A_{PIM}\left[\ddot{S}_{iT}^{q}-\ddot{S}_{i,t-1}^{q}\right]\right)\left(A_{PIM}\left[\ddot{S}_{iT}^{q}-\ddot{S}_{i,t-1}^{q}\right]\right)',$$

with $\ddot{S}_{it}^{q}=S_{it}^{q}-\frac{1}{N}\sum_{j=1}^{N}S_{jt}^{q}$ and $S_{it}^{q}=\sum_{k=1}^{t}q_{ik}$. Therefore, the estimator of $V_{PIM}$ takes the form

$$\begin{aligned}\breve{V}_{PIM} \;=\;& \left(T^{-2}\sum_{i=1}^{N}\sum_{t=1}^{T}A_{PIM}q_{it}q_{it}'A_{PIM}\right)^{-1}\times\\ & \left[T^{-4}\sum_{t=1}^{T}\sum_{i=1}^{N}\breve{\sigma}_{u\cdot v,i}^{2}\left(A_{PIM}\left[\ddot{S}_{iT}^{q}-\ddot{S}_{i,t-1}^{q}\right]\right)\left(A_{PIM}\left[\ddot{S}_{iT}^{q}-\ddot{S}_{i,t-1}^{q}\right]\right)'\right]\times\\ & \left(T^{-2}\sum_{i=1}^{N}\sum_{t=1}^{T}A_{PIM}q_{it}q_{it}'A_{PIM}\right)^{-1}.\end{aligned}$$

Here, two potential candidates for $\breve{\sigma}_{u\cdot v,i}^{2}$ are considered.

1. The first candidate, $\hat{\sigma}_{u\cdot v,i}^{2}$, is based on the residuals of regression (3.7), i.e.

$$\hat{S}_{it}^{\tilde{u}}=S_{it}^{\ddot{y}}-t\hat{\mu}_i-S_{it}^{\ddot{x}\prime}\hat{\beta}-\frac{N-1}{N}x_{it}'\hat{\gamma}_i+\frac{1}{N}\sum_{j=1,j\neq i}^{N}x_{jt}'\hat{\gamma}_j$$

where $\hat{\mu}_i$, $\hat{\beta}$ and $\hat{\gamma}_i$ $(i=1,2,\cdots,N)$ are the panel IM-OLS estimators. Define a HAC estimator using the first difference of $\hat{S}_{it}^{\tilde{u}}$:

$$\hat{\sigma}_{u\cdot v,i}^{2}=T^{-1}\sum_{j=2}^{T}\sum_{h=2}^{T}k\left(\frac{|j-h|}{M}\right)\triangle\hat{S}_{ij}^{\tilde{u}}\triangle\hat{S}_{ih}^{\tilde{u}}.$$

2. The second candidate, $\tilde{\sigma}_{u\cdot v,i}^{2}$, is based on the residuals of a further augmented regres-

sion of the partial sum regression (3.6), i.e.

$$\tilde{S}_{it}^{\tilde{u}} = S_{it}^{\ddot{y}} - t\tilde{\mu}_i - S_{it}^{\ddot{x}\prime}\tilde{\beta}_i - \frac{N-1}{N}x_{it}'\tilde{\gamma}_{i,i} + \frac{1}{N}\sum_{j=1,j\neq i}^{N} x_{jt}'\tilde{\gamma}_{i,j} - z_{it}'\tilde{\lambda}_i$$

where

$$z_{it} = t\sum_{j=1}^{T} D_{ij} - \sum_{j=1}^{t-1}\sum_{s=1}^{j} D_{is},$$

$$D_{it} = \begin{bmatrix} S_{it}^{\ddot{x}\prime} & t & \frac{-1}{N}x_{1t}' & \cdots & \frac{N-1}{N}x_{it}' & \cdots & \frac{-1}{N}x_{Nt}' \end{bmatrix}',$$

and $\tilde{\mu}_i$, $\tilde{\beta}_i$, $\tilde{\gamma}_{i,i}$ and $\tilde{\gamma}_{i,j}$ are OLS from the further augmented regression given by

$$S_{it}^{\ddot{y}} = t\mu_i + S_{it}^{\ddot{x}\prime}\beta_i + \frac{N-1}{N}x_{it}'\gamma_{i,i} - \frac{1}{N}\sum_{j=1,j\neq i}^{N} x_{jt}'\gamma_{i,j} + z_{it}'\lambda_i. \tag{3.9}$$

Note that for given $i$, the estimators $\gamma_{i,i}$ is the parameter associate with $i^{th}$ individual's $x_{it}$ regressor, and $\gamma_{i,j}$ are the parameters associate with $i^{th}$ individual's all $x_{jt}$ regressor for $j = 1, 2, \cdots, N$ and $j \neq i$. They are allowed to be different across different individual because the further augmented regressions are being done *individual by individual.* Therefore, the HAC estimator using $\tilde{S}_{it}^{\tilde{u}}$ is defined as

$$\tilde{\sigma}_{u\cdot v,i}^2 = T^{-1}\sum_{j=2}^{T}\sum_{h=2}^{T} k\left(\frac{|j-h|}{M}\right)\triangle\tilde{S}_{ij}^{\tilde{u}}\triangle\tilde{S}_{ih}^{\tilde{u}}.$$

**Remark 6.** *The reason for considering the second variance estimator, $\tilde{\sigma}_{u\cdot v,i}^2$, is that it delivers an asymptotic pivotal limit in the following two cases: (i) $N = 1$ (See Vogelsang and Wagner [2014]); (ii) $N > 1$ with homogeneous variance structure (See Vogelsang et al. [2016]). However, when $N > 1$ and heterogeneous variance structure exists, $\tilde{\sigma}_{u\cdot v,i}^2$ no* *longer leads to an asymptotic pivotal limit. In practice, this estimator should be considered for $N = 1$ case or $N > 1$ with homogeneous variance structure case.*

Let $\hat{t}$ and $\hat{W}$ denote statistics defined using $\hat{\sigma}^2_{u\cdot v,i}$ to construct $\breve{V}_{PIM}$, and likewise $\tilde{t}$ and $\tilde{W}$ denote statistics defined using $\tilde{\sigma}^2_{u\cdot v,i}$ to construct $\breve{V}_{PIM}$. Letting $\breve{t}$ and $\breve{W}$ denote either $\hat{t}$ and $\hat{W}$ or $\tilde{t}$ and $\tilde{W}$, define the $t$ and $Wald$ statistics as:

$$
\begin{aligned}
\breve{t} &= \frac{\left(R\hat{\theta} - r\right)}{\sqrt{RA_{PIM}\breve{V}_{PIM}A_{PIM}R'}} \\
\breve{W} &= \left(R\hat{\theta} - r\right)' \left[RA_{PIM}\breve{V}_{PIM}A_{PIM}R'\right]^{-1} \left(R\hat{\theta} - r\right).
\end{aligned}
$$

**Theorem 4.** *Assume that the data are generated by (3.1) and (3.2), and that Assumptions 6 and 7 hold. Under traditional bandwidth and kernel assumptions, with $N$ fixed as $T \to \infty$*

-
$$\hat{W} \Rightarrow \frac{\chi^2_q}{1 + d'_\gamma d_\gamma}$$

  *and when $q = 1$,*

$$\hat{t} \Rightarrow \frac{Z}{\sqrt{(1 + d'_\gamma d_\gamma)}}$$

  *where $\chi^2_q$ is a chi-square random variable with $q$ degrees of freedom, $Z$ is a standard* *normal random variable,*

$$
\begin{aligned}
d'_{\gamma}d_{\gamma} =& V_{PIM}^{-1}\left(\int_0^1 \sum_{i=1}^{N} h_i(r)h'_i(r)dr\right)^{-1} \times \\
& \left[\sum_{i=1}^{N} \sigma^2_{u\cdot v,i}\left(d'_{\gamma_i}d_{\gamma_i}\right)\int_0^1 \left[\ddot{H}_i(1)-\ddot{H}_i(r)\right]\left[\ddot{H}_i(1)-\ddot{H}_i(r)\right]' dr\right] \times \\
& \left(\int_0^1 \sum_{i=1}^{N} h_i(r)h'_i(r)dr\right)^{-1},
\end{aligned}
$$

*and*

$$
d'_{\gamma_i}d_{\gamma_i} = \sigma^{-2}_{u\cdot v,i}d'_{\Psi_i}\Omega_{vv,i}d_{\Psi_i},
$$

*where* $d_{\Psi_i}$ *is the* $(i+1)^{th}$ $k \times 1$ *block of the distribution* $\Psi$*.*

- *Under fixed-b asymptotics where* $M = bT$*,* $b \in (0,1]$ *is held fixed as* $T \to \infty$*, then the fixed-b limits of* $\hat{W}$ *and* $\hat{t}$ *are given by*

$$
\hat{W} \Rightarrow \frac{\chi^2_q}{\hat{Q}(b)}
$$

*and when* $q = 1$*,*

$$
\hat{t} \Rightarrow \frac{Z}{\sqrt{\hat{Q}(b)}}
$$

*where*

$$
\begin{aligned}
\hat{Q}(b) =& V_{PIM}^{-1}\left(\int_0^1 \sum_{i=1}^{N} h_i(r)h_i'(r)dr\right)^{-1} \times \\
& \left[\sum_{i=1}^{N} \sigma_{u\cdot v,i}^2 Q_b\left(\hat{P}_i(r)\right)\int_0^1 \left[\ddot{H}_i(1)-\ddot{H}_i(r)\right]\left[\ddot{H}_i(1)-\ddot{H}_i(r)\right]' dr\right] \times \\
& \left(\int_0^1 \sum_{i=1}^{N} h_i(r)h_i'(r)dr\right)^{-1}
\end{aligned}
$$

*is a stochastic process that depends on the kernel function, bandwidth and* $W_i(r)$*.*

- *Under fixed-b asymptotics where* $M = bT$*,* $b \in (0,1]$ *is held fixed as* $T \to \infty$*, then*

$$\tilde{W} \Rightarrow \frac{\chi_q^2}{\tilde{Q}(b)}$$

*and when* $q = 1$*,*

$$\tilde{t} \Rightarrow \frac{Z}{\sqrt{\tilde{Q}(b)}}$$

*where*

$$
\begin{aligned}
\tilde{Q}(b) =& V_{PIM}^{-1}\left(\int_0^1 \sum_{i=1}^{N} h_i(r)h_i'(r)dr\right)^{-1} \times \\
& \left[\sum_{i=1}^{N} \sigma_{u\cdot v,i}^2 Q_b\left(\tilde{P}_i(r)\right)\int_0^1 \left[\ddot{H}_i(1)-\ddot{H}_i(r)\right]\left[\ddot{H}_i(1)-\ddot{H}_i(r)\right]' dr\right] \times \\
& \left(\int_0^1 \sum_{i=1}^{N} h_i(r)h_i'(r)dr\right)^{-1}
\end{aligned}
$$

*is a stochastic process that depends on the kernel function, bandwidth and* $W_i(r)$*.*

### 3.3.2 Inference using the stationary bootstrap

Unfortunately, the statistics $\breve{t}$ and $\breve{W}$ are not asymptotic pivotal, which makes asymptotic inference infeasible. One possible solution is applying the stationary bootstrap to mimic the non-pivotal asymptotic distribution of those statistics. The stationary bootstrap, proposed by Politis and Romano [1994], is a special type of block bootstrap where the block size follows a geometric distribution instead of a fixed number. For a geometric distribution with parameter $p_T$, the expected block size of the stationary bootstrap is $1/p_T$. The stationary bootstrap has been used in the literature of unit root tests, cointegration tests and cointegrated regression inference; see Swensen [2003], Paparoditis and Politis [2005] , Parker et al. [2006], Shin [2015], and Shin and Hwang [2013]. It can capture the serial correlation structure in the original sample by block resampling, and it produces stationary bootstrap samples. A formal description of the stationary bootstrap inference procedure is given below.

1. Calculate the residuals based on regression (3.7) as

$$\hat{S}_{it}^{\tilde{u}} = S_{it}^{\ddot{y}} - t\hat{\mu}_i - S_{it}^{\ddot{x}\prime}\hat{\beta} - \frac{N-1}{N}x_{it}'\hat{\gamma}_i + \frac{1}{N}\sum_{j=1,j\neq i}^{N} x_{jt}'\hat{\gamma}_j$$

   where $\hat{\mu}_i$, $\hat{\beta}$, $\hat{\gamma}_i$ and $\hat{\gamma}_j$ for $j = 1, \cdots, N$ and $j \neq i$ are the panel IM-OLS estimators.

2. Define $\triangle\hat{S}_{it}^{\tilde{u}}$ as a proxy for $\ddot{u}_{it}$, and $\triangle\ddot{x}_{it}$ as a proxy for $\ddot{v}_{it}$, which is the cross-sectional demeaned $v_{it}$. Based on those proxies, define $\hat{\ddot{\eta}}_{it} = \begin{bmatrix}\triangle\hat{S}_{it}^{\tilde{u}} & \triangle\ddot{x}_{it}'\end{bmatrix}' = \begin{bmatrix}\hat{\ddot{u}}_{it} & \hat{\ddot{v}}_{it}'\end{bmatrix}'$ as a proxy for $\ddot{\eta}_{it} = \begin{bmatrix}\ddot{u}_{it} & \ddot{v}_{it}'\end{bmatrix}'$ for $t = 1, 2, \cdots, T$, and set $\hat{S}_{i0}^{\tilde{u}} = 0$ and $\ddot{x}_{i0}$ be zero vector for all $i$.

3. Re-sample the series $\left\{\hat{\ddot{\eta}}_{it}\right\}$ via the stationary bootstrap, obtaining $\left\{\hat{\ddot{\eta}}^{\star}\right\}$, which can be partitioned the same as $\hat{\ddot{\eta}}_{it}$ into $\hat{\ddot{\eta}}^{\star} = \left[\begin{matrix}\hat{\ddot{u}}_{it}^{\star} & \hat{\ddot{v}}_{it}^{\star\prime}\end{matrix}\right]^{\prime}$.

4. Obtain the bootstrap samples $\{\ddot{x}_{it}^{\star}\}$ by

$$\ddot{x}_{it}^{\star} = \sum_{j=1}^{t} \hat{\ddot{v}}_{ij}^{\star},$$

and generate the bootstrap samples $\{\ddot{y}_{it}^{\star}\}$ from [1]

$$\ddot{y}_{it}^{\star} = \hat{\mu}_i + \ddot{x}_{it}^{\star\prime}\hat{\beta} + \hat{\ddot{u}}_{it}^{\star}.$$

5. After obtaining the bootstrap demeaned variables $\ddot{x}_{it}^{\star}$ and $\ddot{y}_{it}^{\star}$, follow the same procedure as discussed before to estimate $\theta$, denoted by $\hat{\theta}^{\star}$, and compute the bootstrap estimator of the limiting variance $V_{PIM}$, say $\breve{V}_{PIM}^{\star}$. Define the bootstrap statistics as follows

$$\begin{aligned}
\breve{t}^{\star} &= \frac{\left(R\hat{\theta}^{\star} - R\hat{\theta}\right)}{\sqrt{RA_{PIM}\breve{V}_{PIM}^{\star}A_{PIM}R^{\prime}}} \\
\breve{W}^{\star} &= \left(R\hat{\theta}^{\star} - R\hat{\theta}\right)^{\prime}\left[RA_{PIM}\breve{V}_{PIM}^{\star}A_{PIM}R^{\prime}\right]^{-1}\left(R\hat{\theta}^{\star} - R\hat{\theta}\right).
\end{aligned}$$

6. Repeat steps 3-5 independently $B$ times to obtain samples $\left\{\breve{t}_j^{\star}\right\}_{j=1}^{B}$ and $\left\{\breve{W}_j^{\star}\right\}_{j=1}^{B}$.

[1] Note that there is another method to obtain $\ddot{x}_{it}^{\star}$. One can resample directly from the original regressor $x_{it}$ to obtain $x_{it}^{\star}$ and then apply cross-sectional demeaning to $x_{it}^{\star}$ to obtain $\ddot{x}_{it}^{\star}$ . However, the size of the tests based on this method is higher than that of the method introduced in procedure 1 to 4. Therefore, this method is not included in this paper.

7. Compute the equal tail bootstrap p-value as

$$
\begin{aligned}
p^{\star}\left(\breve{t}\right) &= 2\min\left(\frac{1}{B}\sum_{j=1}^{B} I\left(\breve{t}_j^{\star} \leq \breve{t}\right), \frac{1}{B}\sum_{j=1}^{B} I\left(\breve{t}_j^{\star} > \breve{t}\right)\right) \\
p^{\star}\left(\breve{W}\right) &= \frac{1}{B}\sum_{j=1}^{B} I\left(\breve{W}_j^{\star} > \breve{W}\right),
\end{aligned}
$$

where $I(\cdot)$ is the indicator function. Reject the null hypothesis if the equal tail bootstrap p-value is less than 5%.

## 3.4 Finite sample simulation

This section investigates finite sample size and power of the bootstrap tests based on the panel IM-OLS estimators. The data generating process is given by

$$
\begin{aligned}
y_{it} &= x1_{it}\beta_1 + x2_{it}\beta_2 + u_{it} \\
x1_{it} &= x1_{i,t-1} + v1_{it} \\
x2_{it} &= x2_{i,t-1} + v2_{it}
\end{aligned}
$$

where for all $i = 1, 2, \cdots, N$, $u_{i0} = 0$, $x1_{i0}$ and $x2_{i0}$ are zero vectors, and

$$
\begin{aligned}
u_{it} &= \rho_1 u_{i,t-1} + \rho_2(e1_{it} + e2_{it}) + \varepsilon_{it} \\
v1_{it} &= e1_{it} + 0.5e1_{i,t-1} \\
v2_{it} &= e2_{it} + 0.5e2_{i,t-1}
\end{aligned}
$$

where for $i^{th}$ individual, $\varepsilon_{it}$, $e1_{it}$ and $e2_{it}$ are $i.i.d.$ $N\left(0, i^2\right)$ random variables. There is no individual effect term and common time effect term included in the data generating process because the focus here is on $\beta_1$ and $\beta_2$, and the estimates of $\beta_1$ and $\beta_2$ are exactly invariant to those terms. The parameter values are $\beta_1 = \beta_2 = 1$. In addition, $\rho_1$ and $\rho_2$ are chosen from $\{0.6, 0.9\}$. The parameter $\rho_1$ controls serial correlation in the regression error, and $\rho_2$ determines the endogeneity of the regressors. The kernel function used in this simulation study is the Bartlett kernel, and the bandwidths are given by $M = bT$ with $b \in \{0.1, 0.5, 1\}$. For the block length parameter $p_T$ in the stationary bootstrap, two different settings are presented. One is $p_T = 0.01(4 - j)(T/50)^{-1/3}$ with $j \in \{1, 2, 3\}$, the other is $p_T = 0.04(4 - j)(T/50)^{-1/3}$ with $j \in \{1, 2, 3\}$. The sample sizes are $N = 5$, $T \in \{50, 500\}$. The number of bootstrap replications is $B = 399$, and the number of simulation replications is 1000.

Results only for cases where $\rho_1 = \rho_2$ are reported. The results include t-statistics for testing the null hypothesis $H_0 : \beta_1 = 1$ and Wald statistics for testing the joint null hypothesis $H_0 : \beta_1 = \beta_2 = 1$. The bootstrap panel IM-OLS statistics were implemented in two ways. The first one uses the stationary bootstrap procedures with the bootstrap version of $\hat{\sigma}^2_{u\cdot v,i}$ and is labeled Stat-BS IM-OLS(D). The second one uses the stationary bootstrap procedures with the bootstrap version of $\tilde{\sigma}^2_{u\cdot v,i}$ and is labeled Stat-BS IM-OLS(fb). Rejections for the bootstrap statistics are carried out by comparing the bootstrap p-value with the nominal level, which is 5% in this simulation.

Tables 3.1 to 3.4 report empirical null rejection probabilities of the t and Wald tests. In each table Panel A corresponds to $T = 50$ and Panel B to $T = 500$. Some common findings about the t and Wald tests can be summarized as follows. For both t and Wald tests, Stat-BS IM-OLS(D) statistics tend to have smaller null rejection probabilities than those of Stat-BS

IM-OLS(fb) statistics. When the bandwidth parameter $b$ varies, the rejection probabilities are relatively stable for both t and Wald tests, which shows that the bootstrap method can successfully capture the impact of the bandwidth on the test statistics. In addition, when the sample size $T$ increases from 50 to 500, rejection probabilities approach 0.05 as expected.

As the values of $\rho_1$, $\rho_2$ increase from 0.6 to 0.9, there exists strong serial correlation and endogeneity. It can be seen from Tables 3.1-3.4 that the rejection probabilities in all cases generally increase, but those increases depend on the sample size and the test statistics. If the time sample is small ($T = 50$), the rejection probabilities increase quite a lot for all tests. In contrast, if the time sample size is large ($T = 500$), the Stat-BS IM-OLS(D) statistics have similar rejection probabilities as $\rho_1 = \rho_2 = 0.6$, whereas the rejection probabilities increase quite a bit for the Stat-BS IMOLS(fb) statistics. This implies that when the time sample size is large enough, the Stat-BS IM-OLS(D) statistics can effectively handle strong serial correlation and endogeneity.

Another important pattern in Tables 3.1-3.4 is that the size of the tests depends heavily on the tuning parameter $p_T$. It is not a surprise because the stationary bootstrap is a moving block bootstrap with changing block lengths. Theoretically, there is no rule of thumb for choosing the value of $p_T$ to ensure the hypothesis test has correct size. In a given sample, Politis and White [2004] and Patton et al. [2009] propose a method to obtain an optimal block length parameter for the stationary bootstrap. However, that optimal block length parameter is based on minimizing the MSE of the stationary bootstrap sample mean, which doesn't necessarily guarantee the correct size of the tests. Therefore, several different values for $p_T$ were used in this simulation study. In Tables 3.1-3.4, for both t and Wald tests, when $p_T$ is small, corresponding to large average block length, the tests tend to have over rejection problems. As $p_T$ increases, the over rejection problem becomes less severe and under-rejection problems appear in some of the cases. To obtain the correct size, the t tests require $p_T$ to be larger than that of the Wald tests.

Next consider the power properties of the tests. When the alternative is true, some bootstrap methods fail to simulate critical values that are valid under the null in which case the tests have no power. Therefore, the analysis of the power properties of bootstrap tests is important. Here, only results for the case $\rho_1 = \rho_2 \in \{0.6, 0.9\}$ for the Wald test for $N \in \{5, 15\}$, $T \in \{50, 500\}$ with the Bartlett kernel are provided. If the power of the test is not an issue for small sample size, like $T = 50$, then it will not be a concern when the sample size is large, like $T = 500$. Starting from the null values of $\beta_1$ and $\beta_2$ equal to 1, the alternative values being considered are $\beta_1 = \beta_2 = \beta \in (1, 1.4]$, which are total of 21 values on a grid with mesh 0.02 including the null value. Power and size-adjusted power are reported. Note that size-adjusted power is not feasible in practice, but it allows us to see the theoretical power differences across tests while holding null rejection probabilities constant at 0.05.

Figures 3.1-3.4 show that using the bootstrap method, the Stat-BS IM-OLS(D) and Stat-BS IM-OLS(fb) Wald tests do have power. Figure 3.1 shows the power comparison of the Stat-BS IM-OLS(D) Wald test for small ($T = 50$) and large ($T = 500$) sample sizes and using respective block size parameter values, $p_T \in \{0.08, 0.00464\}$, give null rejections close to 5%. It can be seen that the power of the tests with the larger sample size ($T = 500$) and smaller $p_T$ ($p_T = 0.00464$) grows dramatically fast. This implies that if the sample size is large enough and the resampling block size parameter $p_T$ can be wisely chosen, the Stat-BS IM-OLS(D) Wald test tends to have very high power. And even if the sample size is relatively small ($T = 50$), the power of the test is still acceptable if $p_T$ is carefully chosen.

Next consider the impact of the serial correlation and endogeneity on the power of the

Stat-BS IM-OLS(D) Wald test. Figure 3.2 displays the power comparison of the Stat-BS IM-OLS(D) Wald test for small ($\rho_1 = \rho_2 = 0.6$) and large ($\rho_1 = \rho_2 = 0.9$) serial correlation and endogeneity with respective block size parameter values, $p_T \in \{0.00464, 0.00696\}$, give null rejections close to 5%. The power of the test with smaller serial correlation and endogeneity ($\rho_1 = \rho_2 = 0.6$) is higher than that of the test with larger serial correlation and endogeneity ($\rho_1 = \rho_2 = 0.9$). If the sample size is small, the power of the tests is lower as expected.

Figures 3.3 and 3.4 provide size-corrected power comparisons between the Stat-BS IM-OLS(D) and Stat-BS IM-OLS(fb) Wald tests for the same values of $T$, $\rho_1$, $\rho_2$, $b$ but using different sample size $N$. In Figure 3.3, the sample size $N$ is 5, while in Figure 3.4, the sample size $N$ is 15. These two figures allow us to see power differences across tests while holding null rejection probabilities constant at 0.05. It can be seen that when the cross sectional sample size is small, the Stat-BS IM-OLS(fb) test has slightly higher power than that of the Stat-BS IM-OLS(D) test. However, when the cross sectional sample size increases, the power of Stat-BS IM-OLS(D) test is much higher. This implies that we should not consider using the Stat-BS IM-OLS(fb) test when $N$ is large, because it has large size distortions and lower power in this scenario.

## 3.5 Summary and conclusions

This paper considers the estimation and inference of a homogeneous cointegrated vector in a panel data model with individual heterogeneity and heterogeneous variance structure. In addition, the model allows a limited degree of cross-sectional dependence due to a common time effect. The estimator is labeled as panel IM-OLS. It is a fully parametric estimator that is based on a partial sum transformed regression augmented by the decomposed demeaned original regressor. The advantage is that it leads to a zero mean mixed Gaussian limiting distribution without requiring the choice of tuning parameters (like bandwidth, kernel function, numbers of leads and lags). Asymptotic inference is infeasible due to the presence of nuisance parameters, and the stationary bootstrap is used for hypothesis testing. Monte Carlo simulations show that the bootstrap method can deliver good size and power for t and Wald tests, depending on the sample size, serial correlation, endogeneity and the stationary bootstrap block length resampling parameter. When there is strong serial correlation and endogeneity, for moderate time sample sizes, the size of the tests are close to nominal level for certain values of $p_T$.

Unlike in Vogelsang et al. [2016], the further augmented regression residuals do not lead to an asymptotic pivotal test, and the bootstrap hypothesis test based on it has more size distortion. When the cross sectional sample size $N$ is small, the power of the test based on the further augmented regression residuals is a little bit higher than that of the test based on augmented regression residuals. However, when the cross sectional sample size $N$ increases, the power of the test based on the further augmented regression residuals is much lower than that of the test based on augmented regression residuals. This power loss as $N$ increases is because the further augmented regression requires adding many additional regressors to compute the residuals. Therefore, in practice, when $N$ is large and the panel has cross sectional dependence and heterogeneous variance structure, inference based on the further augmented regression residuals is not recommended.

One limitation of the present paper is that the cross-sectional dependence is only coming from a common time effect with a constant factor loading. This might be restrictive in some applications. Therefore, a model with more general cross-sectional dependence may be worth considering in the future. In that more general scenario, the theory of the inference based on the panel IM-OLS type estimators will rely on more general bootstrap procedures. If the stationary bootstrap can mimic the non-pivotal limit of the original statistics, then formally proving the asymptotic equivalence between the stationary bootstrap statistics and the original test statistics may be a viable research topic in the future.

APPENDIX

## Tables and Figures

Table 3.1: Empirical null rejection probabilities, 5% level, t-tests for $H_0 : \beta_1 = 1$, $N = 5$, $\rho = 0.6$, Bartlett kernel

| $p_T$ | Stat-BS(D) | | | Stat-BS(fb) | | |
|---|---|---|---|---|---|---|
| | b=0.1 | b=0.5 | b=1 | b=0.1 | b=0.5 | b=1 |
| Panel A: $T = 50$ | | | | | | |
| 0.01 | 0.246 | 0.214 | 0.226 | 0.346 | 0.335 | 0.336 |
| 0.02 | 0.219 | 0.191 | 0.198 | 0.31 | 0.309 | 0.31 |
| 0.03 | 0.199 | 0.17 | 0.173 | 0.3 | 0.308 | 0.303 |
| 0.04 | 0.184 | 0.162 | 0.157 | 0.285 | 0.286 | 0.287 |
| 0.08 | 0.127 | 0.114 | 0.115 | 0.231 | 0.239 | 0.239 |
| 0.12 | 0.102 | 0.094 | 0.095 | 0.213 | 0.209 | 0.207 |
| Panel B: $T = 500$ | | | | | | |
| 0.00464 | 0.152 | 0.145 | 0.138 | 0.174 | 0.176 | 0.162 |
| 0.00928 | 0.099 | 0.101 | 0.102 | 0.123 | 0.122 | 0.106 |
| 0.01393 | 0.074 | 0.075 | 0.081 | 0.086 | 0.074 | 0.085 |
| 0.01857 | 0.053 | 0.062 | 0.065 | 0.066 | 0.065 | 0.068 |
| 0.03713 | 0.032 | 0.044 | 0.033 | 0.037 | 0.039 | 0.033 |
| 0.05570 | 0.023 | 0.028 | 0.024 | 0.03 | 0.02 | 0.025 |

Table 3.2: Empirical null rejection probabilities, 5% level, t-tests for $H_0 : \beta_1 = 1$, $N = 5$, $\rho = 0.9$, Bartlett kernel

| $p_T$ | Stat-BS(D) | | | Stat-BS(fb) | | |
|---|---|---|---|---|---|---|
| | b=0.1 | b=0.5 | b=1 | b=0.1 | b=0.5 | b=1 |
| Panel A: $T = 50$ | | | | | | |
| 0.01 | 0.427 | 0.375 | 0.371 | 0.74 | 0.745 | 0.75 |
| 0.02 | 0.405 | 0.349 | 0.347 | 0.747 | 0.743 | 0.75 |
| 0.03 | 0.395 | 0.328 | 0.328 | 0.738 | 0.733 | 0.734 |
| 0.04 | 0.356 | 0.307 | 0.291 | 0.738 | 0.74 | 0.73 |
| 0.08 | 0.333 | 0.279 | 0.267 | 0.724 | 0.724 | 0.72 |
| 0.12 | 0.307 | 0.247 | 0.235 | 0.715 | 0.716 | 0.715 |
| Panel B: $T = 500$ | | | | | | |
| 0.00464 | 0.144 | 0.138 | 0.137 | 0.243 | 0.23 | 0.23 |
| 0.00928 | 0.095 | 0.096 | 0.098 | 0.175 | 0.172 | 0.165 |
| 0.01393 | 0.07 | 0.07 | 0.074 | 0.144 | 0.125 | 0.13 |
| 0.01857 | 0.053 | 0.061 | 0.06 | 0.115 | 0.113 | 0.114 |
| 0.03713 | 0.031 | 0.035 | 0.033 | 0.093 | 0.08 | 0.077 |
| 0.05570 | 0.022 | 0.027 | 0.024 | 0.073 | 0.077 | 0.072 |

Table 3.3: Empirical null rejection probabilities, 5% level, Wald-tests for $H_0 : \beta_1 = 1, \beta_2 = 1$, $N = 5$, $\rho = 0.6$, Bartlett kernel

| $p_T$ | Stat-BS(D) | | | Stat-BS(fb) | | |
|---|---|---|---|---|---|---|
| | b=0.1 | b=0.5 | b=1 | b=0.1 | b=0.5 | b=1 |
| Panel A: $T = 50$ | | | | | | |
| 0.01 | 0.159 | 0.142 | 0.153 | 0.296 | 0.299 | 0.302 |
| 0.02 | 0.125 | 0.122 | 0.122 | 0.266 | 0.268 | 0.269 |
| 0.03 | 0.107 | 0.109 | 0.109 | 0.239 | 0.236 | 0.236 |
| 0.04 | 0.092 | 0.101 | 0.086 | 0.216 | 0.22 | 0.214 |
| 0.08 | 0.049 | 0.047 | 0.047 | 0.155 | 0.151 | 0.157 |
| 0.12 | 0.029 | 0.031 | 0.027 | 0.1 | 0.113 | 0.119 |
| Panel B: $T = 500$ | | | | | | |
| 0.00464 | 0.053 | 0.053 | 0.051 | 0.081 | 0.073 | 0.074 |
| 0.00928 | 0.02 | 0.023 | 0.022 | 0.037 | 0.038 | 0.039 |
| 0.01393 | 0.01 | 0.011 | 0.019 | 0.019 | 0.019 | 0.021 |
| 0.01857 | 0.002 | 0.01 | 0.007 | 0.008 | 0.01 | 0.016 |
| 0.03713 | 0.002 | 0.003 | 0.004 | 0.002 | 0.002 | 0.005 |
| 0.05570 | 0 | 0.003 | 0.003 | 0 | 0.002 | 0.003 |

Table 3.4: Empirical null rejection probabilities, 5% level, Wald-tests for $H_0 : \beta_1 = 1, \beta_2 = 1$, $N = 5$, $\rho = 0.9$, Bartlett kernel

| $p_T$ | Stat-BS(D) | | | Stat-BS(fb) | | |
|---|---|---|---|---|---|---|
| | b=0.1 | b=0.5 | b=1 | b=0.1 | b=0.5 | b=1 |
| Panel A: $T = 50$ | | | | | | |
| 0.01 | 0.474 | 0.402 | 0.393 | 0.876 | 0.876 | 0.867 |
| 0.02 | 0.448 | 0.374 | 0.368 | 0.874 | 0.873 | 0.864 |
| 0.03 | 0.43 | 0.346 | 0.348 | 0.874 | 0.87 | 0.865 |
| 0.04 | 0.402 | 0.337 | 0.316 | 0.869 | 0.867 | 0.865 |
| 0.08 | 0.325 | 0.261 | 0.25 | 0.859 | 0.861 | 0.846 |
| 0.12 | 0.268 | 0.212 | 0.21 | 0.848 | 0.854 | 0.841 |
| Panel B: $T = 500$ | | | | | | |
| 0.00464 | 0.07 | 0.062 | 0.069 | 0.16 | 0.159 | 0.166 |
| 0.00928 | 0.031 | 0.04 | 0.041 | 0.107 | 0.096 | 0.097 |
| 0.01393 | 0.014 | 0.016 | 0.021 | 0.069 | 0.063 | 0.069 |
| 0.01857 | 0.011 | 0.018 | 0.013 | 0.057 | 0.043 | 0.056 |
| 0.03713 | 0.004 | 0.006 | 0.008 | 0.028 | 0.031 | 0.029 |
| 0.05570 | 0.003 | 0.005 | 0.004 | 0.019 | 0.019 | 0.021 |

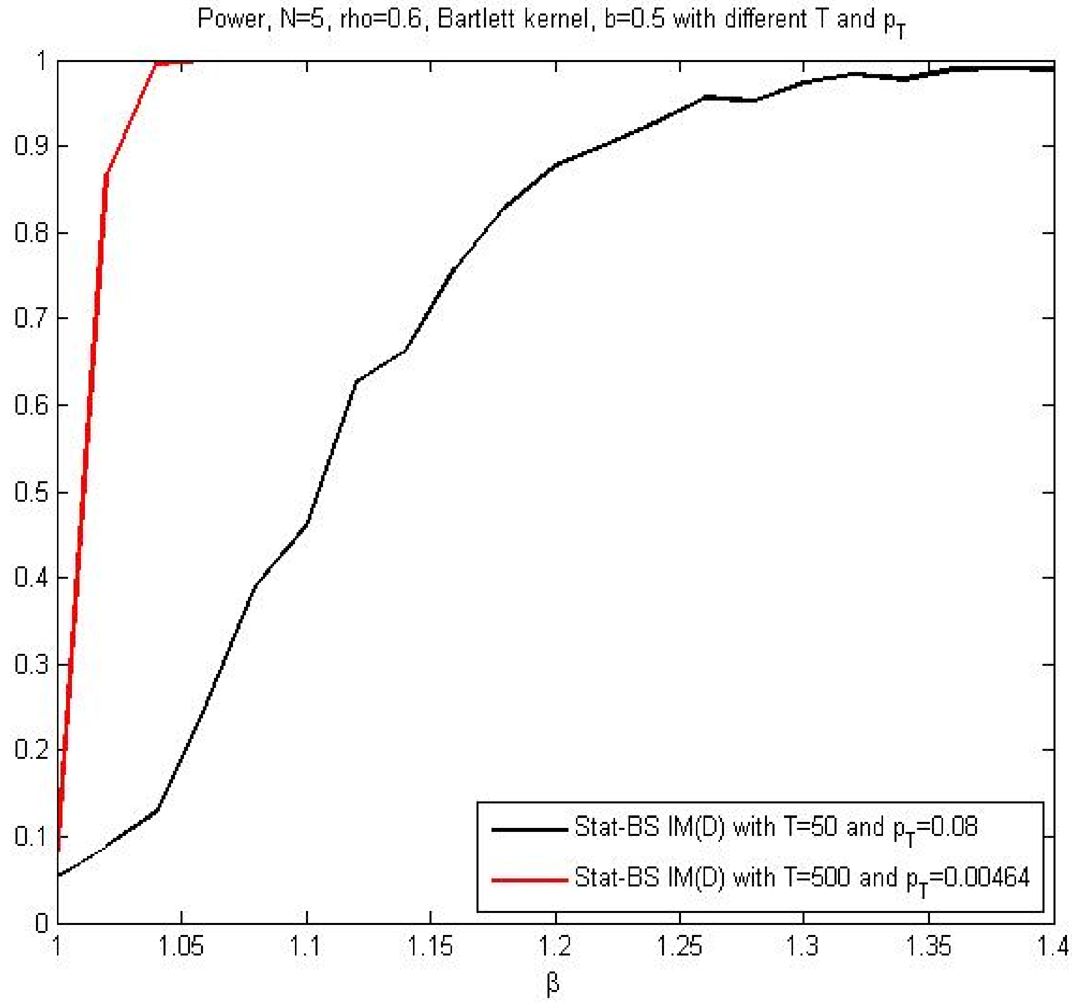


Figure 3.1: Power of bootstrap Stat-BS IM (D), Wald test, $N = 5$, $\rho_1 = \rho_2 = 0.6$, $b = 0.5$, Bartlett kernel with different $T$ and $p_T$

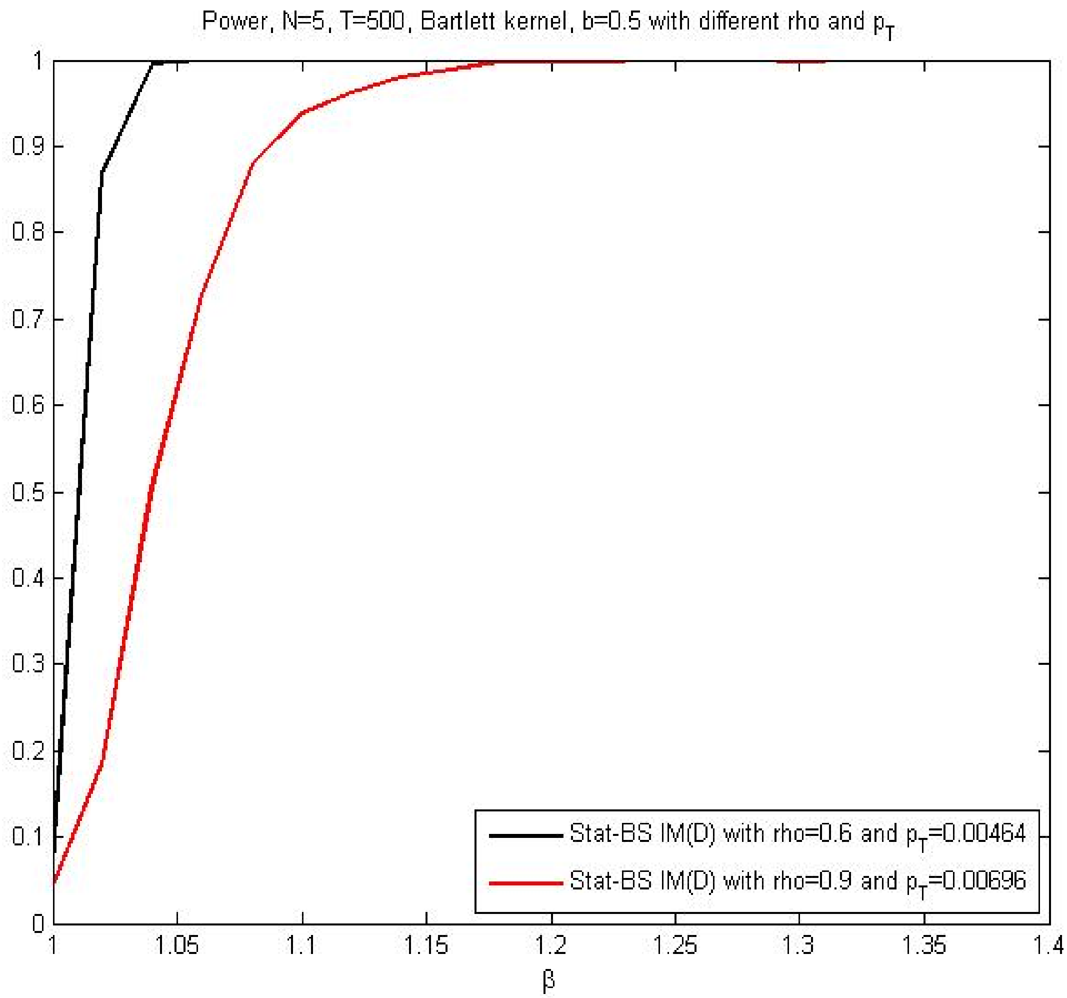


Figure 3.2: Power of bootstrap Stat-BS IM (D), Wald test, $N = 5$, $T = 500$, $b = 0.5$, Bartlett kernel with different $\rho$ and $p_T$

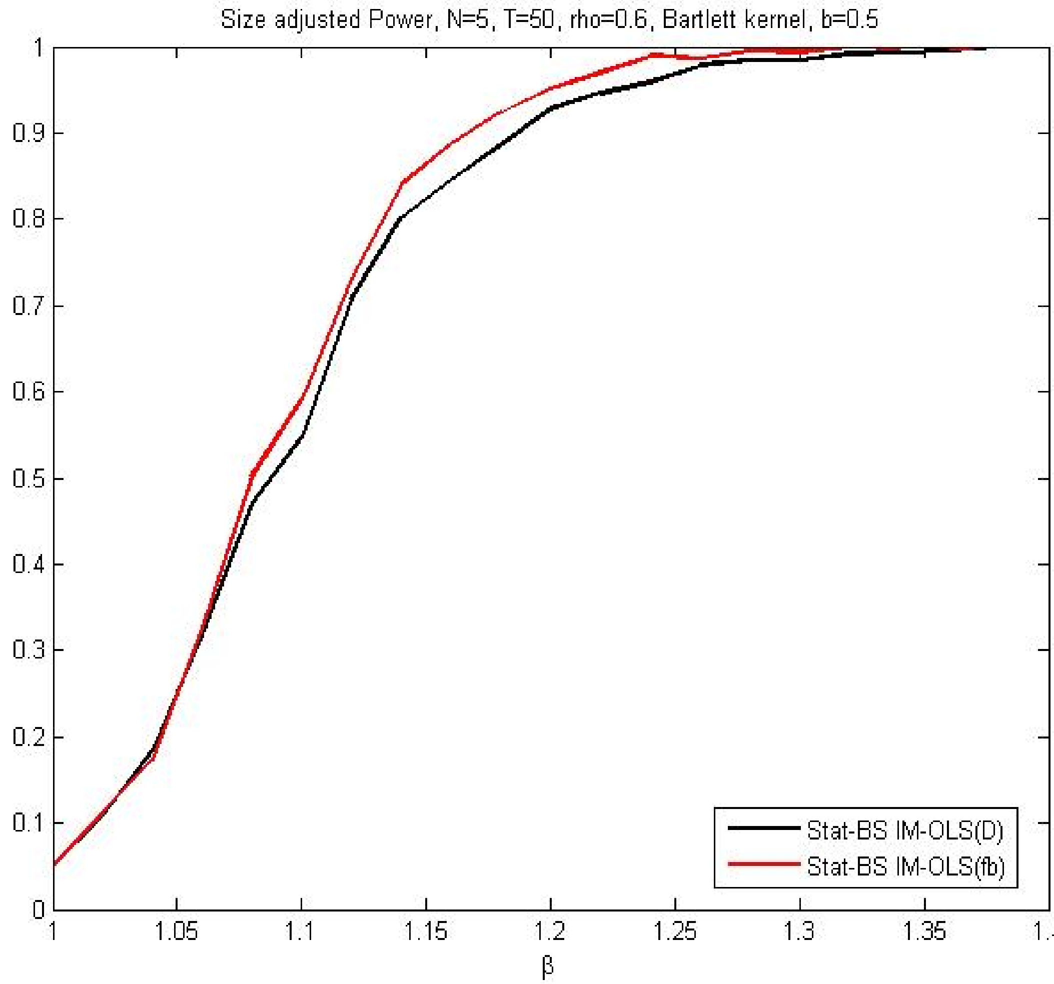


Figure 3.3: Size adjusted power, Wald-tests, $N = 5$, $T = 50$, $\rho_1 = \rho_2 = 0.6$, $b = 0.5$, Bartlett kernel

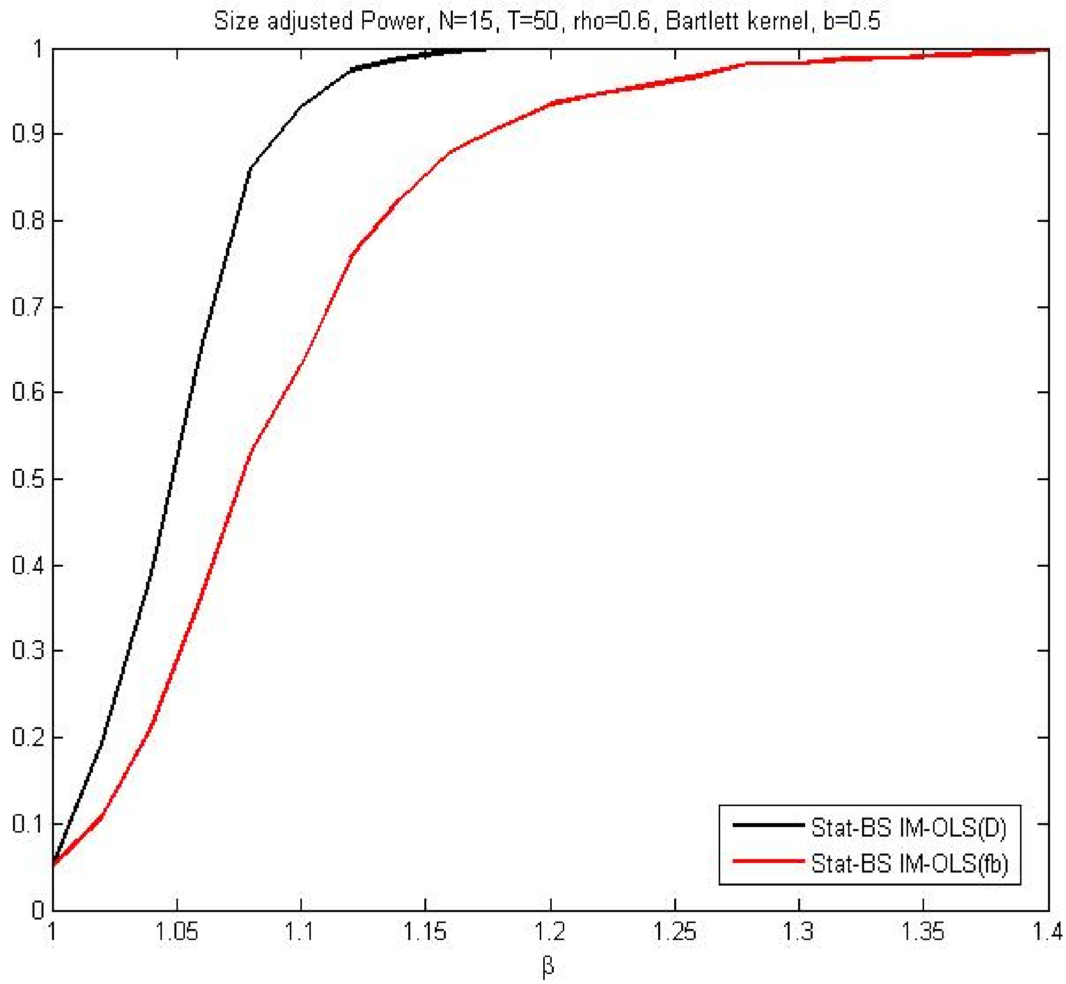


Figure 3.4: Size adjusted power, Wald-tests, $N = 15$, $T = 50$, $\rho_1 = \rho_2 = 0.6$, $b = 0.5$, Bartlett kernel

## Proof of failure of using $\ddot{x}_{it}$ to solve endogeneity problem

This is the proof showing that directly adding $\ddot{x}_{it}$ to regression (3.6) cannot fully deal with the endogeneity problem in the model considered in this paper. Suppose, we add $\ddot{x}_{it}$ into the partial sum model, which gives

$$S_{it}^{\ddot{y}} = t\mu_i + S_{it}^{\ddot{x}\prime}\beta + \ddot{x}_{it}'\gamma_i + S_{it}^{\ddot{u}} - \ddot{x}_{it}'\gamma_i$$

.Consider the behavior of $T^{-\frac{1}{2}}\left(S_{it}^{\ddot{u}} - \ddot{x}_{it}'\gamma_i\right)$ as $T \to \infty$,

$$\begin{aligned}
T^{-\frac{1}{2}}\left(S_{it}^{\ddot{u}} - \ddot{x}_{it}'\gamma_i\right) &= T^{-\frac{1}{2}}\left(S_{it}^{u} - \frac{1}{N}\sum_{j=1}^{N} S_{jt}^{u} - x_{it}'\gamma_i + \frac{1}{N}\sum_{j=1}^{N} x_{jt}'\gamma_i\right) \\
&= T^{-\frac{1}{2}}\left(S_{it}^{u} - x_{it}'\gamma_i\right) - \frac{1}{N}\sum_{j=1}^{N} T^{-\frac{1}{2}}\left(S_{jt}^{u} - x_{jt}'\gamma_i\right) \\
&\Rightarrow \left[B_{u,i}(r) - B_{v,i}'(r)\gamma_i\right] - \frac{1}{N}\sum_{j=1}^{N}\left[B_{u,j}(r) - B_{v,j}'(r)\gamma_i\right] \\
&= \sigma_{u\cdot v,i}w_{u,i}(r) + \lambda_{uv,i}W_{v,i}(r) - \left[\Omega_{vv,i}^{\frac{1}{2}}W_{v,i}(r)\right]'\gamma_i \\
&- \frac{1}{N}\sum_{j=1}^{N}\left[\sigma_{u\cdot v,j}w_{u,j}(r) + \lambda_{uv,j}W_{v,j}(r) - \left(\Omega_{vv,j}^{\frac{1}{2}}W_{v,j}(r)\right)'\gamma_i\right] \\
&= \left[\sigma_{u\cdot v,i}w_{u,i}(r) - \frac{1}{N}\sum_{j=1}^{N}\sigma_{u,v,j}w_{u,j}(r)\right] \\
&- W_{v,i}'(r)\Omega_{vv,i}^{\frac{1}{2}\prime}\left(\gamma_i - \Omega_{vv,i}^{-\frac{1}{2}\prime}\lambda_{uv,i}'\right) \\
&+ \frac{1}{N}\sum_{j=1}^{N}\left[W_{v,j}'(r)\Omega_{vv,j}^{\frac{1}{2}\prime}\left(\gamma_i - \Omega_{vv,j}^{-\frac{1}{2}\prime}\lambda_{uv,j}'\right)\right] \\
&\neq \sigma_{u\cdot v,i}w_{u,i}(r) - \frac{1}{N}\sum_{j=1}^{N}\sigma_{u,v,j}w_{u,j}(r)
\end{aligned}$$

Note that the last inequality holds because when there is heterogeneity in the 2nd moment structure, it is almost impossible that

$$\gamma_i = \Omega_{vv,i}^{-\frac{1}{2}\prime}\lambda_{uv,i}' = \Omega_{vv,j}^{-\frac{1}{2}\prime}\lambda_{uv,j}'$$

for all $j = 1, 2, \cdots, N$. Therefore, just adding $\ddot{x}_{it}$ to regression (3.6) cannot fully deal with the endogeneity problem. Note that, if the 2nd moment structure is homogeneous, then only adding $\ddot{x}_{it}$ to regression (3.6) will work, because

$$\gamma_i = \Omega_{vv,i}^{-\frac{1}{2}\prime}\lambda_{uv,i}' = \Omega_{vv,j}^{-\frac{1}{2}\prime}\lambda_{uv,j}' = \gamma_j = \Omega_{vv}^{-\frac{1}{2}\prime}\lambda_{uv}'$$

for all $i$, $j$.

## Proof of Theorem 3

In order to derive the asymptotic distribution of the panel IM-OLS estimator, we start with regression (3.7). First consider the limit of $T^{-1/2}S_{it}^{\tilde{u}}$ with $N$ fixed and $T \to \infty$.

$$
\begin{aligned}
T^{-1/2}S_{it}^{\tilde{u}} &= T^{-1/2}\left(S_{it}^{\ddot{u}} - \frac{N-1}{N}x_{it}'\gamma_i + \frac{1}{N}\sum_{j=1,j\neq i}^{N} x_{jt}'\gamma_j\right) \\
&= T^{-1/2}\left(S_{it}^{u} - \frac{1}{N}\sum_{j=1}^{N} S_{jt}^{u} - x_{it}'\gamma_i + \frac{1}{N}\sum_{j=1}^{N} x_{jt}'\gamma_j\right) \\
&= T^{-\frac{1}{2}}\left(S_{it}^{u} - x_{it}'\gamma_i\right) - \frac{1}{N}\sum_{j=1}^{N} T^{-\frac{1}{2}}\left(S_{jt}^{u} - x_{jt}'\gamma_j\right) \\
&\Rightarrow \left[B_{u,i}(r) - B_{v,i}'(r)\gamma_i\right] - \frac{1}{N}\sum_{j=1}^{N}\left[B_{u,j}(r) - B_{v,j}'(r)\gamma_j\right] \\
&= \sigma_{u\cdot v,i}w_{u,i}(r) + \lambda_{uv,i}W_{v,i}(r) - W_{v,i}'(r)\Omega_{vv,i}^{\frac{1}{2}\prime}\gamma_i \\
&\quad -\frac{1}{N}\sum_{j=1}^{N}\left[\sigma_{u\cdot v,j}w_{u,j}(r) + \lambda_{uv,j}W_{v,i}(r) - W_{v,j}'(r)\Omega_{vv,j}^{\frac{1}{2}\prime}\gamma_j\right] \\
&= \left[\sigma_{u\cdot v,i}w_{u,i}(r) - \frac{1}{N}\sum_{j=1}^{N}\sigma_{u\cdot v,j}w_{u,j}(r)\right] - W_{v,i}'(r)\Omega_{vv,i}^{\frac{1}{2}\prime}\left[\gamma_i - \Omega_{vv,i}^{\frac{-1}{2}\prime}\lambda_{uv,i}'\right] \\
&\quad +\frac{1}{N}\sum_{j=1}^{N}W_{v,j}'(r)\Omega_{vv,j}^{\frac{1}{2}\prime}\left[\gamma_j - \Omega_{vv,j}^{\frac{-1}{2}\prime}\lambda_{uv,j}'\right]
\end{aligned}
$$

Therefore, when $\gamma_i = \Omega_{vv,i}^{-\frac{1}{2}\prime}\lambda_{uv,i}' = \Omega_{vv,i}^{-1}\Omega_{vu,i}$, it follows that

$$
T^{-1/2}S_{it}^{\tilde{u}} \Longrightarrow \sigma_{u\cdot v,i}w_{u,i}(r) - \frac{1}{N}\sum_{j=1}^{N}\sigma_{u\cdot v,j}w_{u,j}(r).
$$

Define $A_{1T}^{-1} = T^{-\frac{1}{2}}A_{PIM}$. The next step of the proof is to obtain the limit of $A_{1T}^{-1}q_{it}$ for

$N$ fixed and $T \to \infty$.

$$
\begin{aligned}
A_{1T}^{-1} q_{it} &= \begin{bmatrix} T^{-\frac{3}{2}} S_{it}^{\ddot{x}} \\ T^{-\frac{1}{2}} \left(\frac{-1}{N} x_{1t}\right) \\ \vdots \\ T^{-\frac{1}{2}} \left(\frac{N-1}{N} x_{it}\right) \\ \vdots \\ T^{-\frac{1}{2}} \left(\frac{-1}{N} x_{Nt}\right) \\ 0 \\ \vdots \\ \frac{t}{T} \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} T^{-\frac{3}{2}} S_{it}^{x} - \frac{1}{N} \sum_{j=1}^{N} T^{-\frac{3}{2}} S_{jt}^{x} \\ \frac{-1}{N} T^{-\frac{1}{2}} x_{1t} \\ \vdots \\ \frac{N-1}{N} T^{-\frac{1}{2}} x_{it} \\ \vdots \\ \frac{-1}{N} T^{-\frac{1}{2}} x_{Nt} \\ 0 \\ \vdots \\ \frac{t}{T} \\ \vdots \\ 0 \end{bmatrix} \\
&\Rightarrow \begin{bmatrix} \int_0^r B_{v,i}(s)ds - \frac{1}{N} \sum_{j=1}^{N} \int_0^r B_{v,j}(s)ds \\ \frac{-1}{N} B_{v,1}(r) \\ \vdots \\ \frac{N-1}{N} B_{v,i}(r) \\ \vdots \\ \frac{-1}{N} B_{v,N}(r) \\ 0 \\ \vdots \\ r \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} \int_0^r \left[ \Omega_{vv,i}^{\frac{1}{2}} W_{v,i}(s) - \frac{1}{N} \sum_{j=1}^{N} \Omega_{vv,j}^{\frac{1}{2}} W_{v,j}(s) \right] ds \\ \frac{-1}{N} \Omega_{vv,1}^{\frac{1}{2}} W_{v,1}(r) \\ \vdots \\ \frac{N-1}{N} \Omega_{vv,i}^{\frac{1}{2}} W_{v,i}(r) \\ \vdots \\ \frac{-1}{N} \Omega_{vv,N}^{\frac{1}{2}} W_{v,N}(r) \\ 0 \\ \vdots \\ r \\ \vdots \\ 0 \end{bmatrix} \\
&= h_i(r).
\end{aligned}
$$

Therefore, as $T \to \infty$,

$$A_{1T}^{-1} q_{it} \Rightarrow h_i(r).$$

For fixed $N$, as $T \to \infty$,

$$
\begin{aligned}
A_{PIM}^{-1}\left(\hat{\theta}-\theta\right) &= T^{-\frac{1}{2}} A_{1T}\left(\hat{\theta}-\theta\right) = \begin{bmatrix} T\left(\hat{\beta}-\beta\right) \\ (\hat{\gamma}_1-\gamma_1) \\ \vdots \\ (\hat{\gamma}_N-\gamma_N) \\ \sqrt{T}\,(\hat{\mu}_1-\mu_1) \\ \vdots \\ \sqrt{T}\,(\hat{\mu}_N-\mu_N) \end{bmatrix} \\
&= \left(T^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T} A_{1T}^{-1} q_{it} q_{it}' A_{1T}^{-1}\right)^{-1} \left(T^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T} A_{1T}^{-1} q_{it} T^{-\frac{1}{2}} S_{it}^{\tilde{u}}\right) \\
&\Rightarrow \left(\int_0^1 \sum_{i=1}^{N} h_i(r) h_i'(r) dr\right)^{-1} \times \\
&\quad \left(\int_0^1 \sum_{i=1}^{N} \left\{ h_i(r) \left[\sigma_{u\cdot v,i} w_{u,i}(r) - \frac{1}{N}\sum_{j=1}^{N} \sigma_{u\cdot v,j} w_{u,j}(r)\right]\right\} dr\right) \\
&= \Psi
\end{aligned}
$$

## Proof of the derivation of the form of the asymptotic variance $\Psi$

We start from rewriting $\int_0^1 \sum\limits_{i=1}^{N} \left\{ h_i(r) \left[ \sigma_{u\cdot v,i} w_{u,i}(r) - \frac{1}{N} \sum_{j=1}^{N} \sigma_{u\cdot v,j} w_{u,j}(r) \right] \right\} dr$ as

$$
\begin{aligned}
&\int_0^1 \sum_{i=1}^{N} \left\{ h_i(r) \left[ \sigma_{u\cdot v,i} w_{u,i}(r) - \frac{1}{N} \sum_{j=1}^{N} \sigma_{u\cdot v,j} w_{u,j}(r) \right] \right\} dr \\
=&\int_0^1 \sum_{i=1}^{N} \left[ h_i(r) - \frac{1}{N} \sum_{j=1}^{N} h_j(r) \right] \sigma_{u\cdot v,i} w_{u,i}(r) dr \\
=&\sum_{i=1}^{N} \sigma_{u\cdot v,i} \int_0^1 \left[ h_i(r) - \frac{1}{N} \sum_{j=1}^{N} h_j(r) \right] w_{u,i}(r) dr \\
=&\sum_{i=1}^{N} \sigma_{u\cdot v,i} \int_0^1 w_{u,i}(r) d \left[ H_i(r) - \frac{1}{N} \sum_{j=1}^{N} H_j(r) \right] \\
=&\sum_{i=1}^{N} \sigma_{u\cdot v,i} \left( w_{u,i}(r) \left[ H_i(r) - \frac{1}{N} \sum_{j=1}^{N} H_j(r) \right] |_0^1 - \int_0^1 \left[ H_i(r) - \frac{1}{N} \sum_{j=1}^{N} H_j(r) \right] dw_{u,i}(r) \right) \\
=&\sum_{i=1}^{N} \sigma_{u\cdot v,i} \left( w_{u,i}(1) \left[ H_i(1) - \frac{1}{N} \sum_{j=1}^{N} H_j(1) \right] - \int_0^1 \left[ H_i(r) - \frac{1}{N} \sum_{j=1}^{N} H_j(r) \right] dw_{u,i}(r) \right) \\
=&\sum_{i=1}^{N} \sigma_{u\cdot v,i} \left( \int_0^1 \left[ H_i(1) - \frac{1}{N} \sum_{j=1}^{N} H_j(1) \right] dw_{u,i}(r) - \int_0^1 \left[ H_i(r) - \frac{1}{N} \sum_{j=1}^{N} H_j(r) \right] dw_{u,i}(r) \right) \\
=&\sum_{i=1}^{N} \sigma_{u\cdot v,i} \int_0^1 [\ddot{H}_i(1) - \ddot{H}_i(r)] dw_{u,i}(r).
\end{aligned}
$$

Therefore, the variance of

$$
\int_0^1 \sum_{i=1}^{N} \left\{ h_i(r) \left[ \sigma_{u\cdot v,i} w_{u,i}(r) - \frac{1}{N} \sum_{j=1}^{N} \sigma_{u\cdot v,j} w_{u,j}(r) \right] \right\} dr
$$

will be same as the variance of

$$
\sum_{i=1}^{N} \sigma_{u\cdot v,i} \int_0^1 [\ddot{H}_i(1) - \ddot{H}_i(r)] dw_{u,i}(r)
$$

which is

$$\sum_{i=1}^{N} \sigma_{u\cdot v,i}^2 \int_0^1 [\ddot{H}_i(1) - \ddot{H}_i(r)][\ddot{H}_i(1) - \ddot{H}_i(r)]' dr.$$

Then, the variance of $\Psi$ is

$$\begin{aligned} V_{PIM} = & \left( \int_0^1 \sum_{i=1}^{N} h_i(r) h_i'(r) dr \right)^{-1} \times \\ & \left[ \sum_{i=1}^{N} \sigma_{u\cdot v,i}^2 \int_0^1 [\ddot{H}_i(1) - \ddot{H}_i(r)][\ddot{H}_i(1) - \ddot{H}_i(r)]' dr \right] \times \\ & \left( \int_0^1 \sum_{i=1}^{N} h_i(r) h_i'(r) dr \right)^{-1}. \end{aligned}$$

## Proof of Theorem 4

This is the proof of the null limiting distribution of the test statistics in Theorem 4. First, consider the behavior of $\hat{\sigma}^2_{u\cdot v,i}$. It is based on $\triangle\hat{S}^{\tilde{u}}_{it}$, where

$$\begin{aligned}
\triangle\hat{S}^{\tilde{u}}_{it} &= \triangle S^{\ddot{y}}_{it} - \triangle t\hat{\mu}_i - \triangle S^{\ddot{x}\prime}_{it}\hat{\beta} - \frac{N-1}{N}\triangle x'_{it}\hat{\gamma}_i + \frac{1}{N}\sum_{j=1,j\neq i}^{N}\triangle x'_{jt}\hat{\gamma}_j \\
&= \ddot{y}_{it} - \hat{\mu}_i - \ddot{x}'_{it}\hat{\beta} - \frac{N-1}{N}v'_{it}\hat{\gamma}_i + \frac{1}{N}\sum_{j=1,j\neq i}^{N} v'_{jt}\hat{\gamma}_j \\
&= \mu_i + \ddot{x}'_{it}\beta + \ddot{u}_{it} - \hat{\mu}_i - \ddot{x}'_{it}\hat{\beta} - \frac{N-1}{N}v'_{it}\hat{\gamma}_i + \frac{1}{N}\sum_{j=1,j\neq i}^{N} v'_{jt}\hat{\gamma}_j \\
&= \mu_i + \ddot{x}'_{it}\beta + u_{it} - \frac{1}{N}\sum_{j=1}^{N} u_{jt} - \hat{\mu}_i - \ddot{x}'_{it}\hat{\beta} - v'_{it}\hat{\gamma}_i + \frac{1}{N}\sum_{j=1}^{N} v'_{jt}\hat{\gamma}_j \\
&= \left(u_{it} - v'_{it}\gamma_i\right) - \frac{1}{N}\sum_{j=1}^{N}\left(u_{jt} - v'_{jt}\gamma_j\right) - v'_{it}\left(\hat{\gamma}_i - \gamma_i\right) \\
&+ \frac{1}{N}\sum_{j=1}^{N} v'_{jt}\left(\hat{\gamma}_j - \gamma_j\right) - \left(\hat{\mu}_i - \mu_i\right) - \ddot{x}'_{it}\left(\hat{\beta} - \beta\right) \\
&= \left[u^+_{it} - v'_{it}\left(\hat{\gamma}_i - \gamma_i\right)\right] - \frac{1}{N}\sum_{j=1}^{N}\left[u^+_{jt} - v'_{jt}\left(\hat{\gamma}_j - \gamma_j\right)\right] - \left(\hat{\mu}_i - \mu_i\right) - \ddot{x}'_{it}\left(\hat{\beta} - \beta\right)
\end{aligned}$$

where $u^+_{it} = u_{it} - v'_{it}\gamma_i$. It can be shown that the last three parts of the formula can be neglected for long run variance estimation of $\triangle\hat{S}^{\tilde{u}}_{it}$. Thus, the long run variance estimator based on $\triangle\hat{S}^{\tilde{u}}_{it}$, asymptotically coincides with long run variance estimator based on $u^+_{it} - v'_{it}\left(\hat{\gamma}_i - \gamma_i\right)$.

Define $\eta^+_{it} = \left[u^+_{it},\ \ v'_{it}\right]'$, and then its long run variance is $\Omega^+_i = \begin{bmatrix}\sigma^2_{u\cdot v,i} & 0 \\ 0 & \Omega_{vv,i}\end{bmatrix}$. Using unobserved $\eta^+_{it}$, an infeasible long run variance estimator, $\widehat{\Omega^+_i}$, is consistent. That is $\widehat{\Omega^+_i} \xrightarrow{p} \Omega^+_i$.

Note that: $u^+_{it} - v'_{it}\left(\hat{\gamma}_i - \gamma_i\right) = \eta^{+\prime}_{it}\begin{bmatrix}1 \\ -\left(\hat{\gamma}_i - \gamma_i\right)\end{bmatrix}$, then HAC estimator, $\widetilde{\Omega^+_i}$, for $u^+_{it} -$

$v_{it}' \left(\hat{\gamma}_i - \gamma_i\right)$ can be written as

$$\begin{bmatrix} 1 & -\left(\hat{\gamma}_i - \gamma_i\right)' \end{bmatrix} \widehat{\Omega_i^+} \begin{bmatrix} 1 \\ -\left(\hat{\gamma}_i - \gamma_i\right) \end{bmatrix}$$

with

$$\begin{aligned}
\left(\hat{\gamma}_i - \gamma_i\right) \Rightarrow & \begin{bmatrix} 0_{k\times k} & 0_{k\times k} & \cdots & I_k & \cdots & 0_{k\times k} & 0 & \cdots & 0 \end{bmatrix} \times \\
& \left( \int_0^1 \sum_{i=1}^N h_i(r) h_i'(r) dr \right)^{-1} \left\{ \sum_{i=1}^N \sigma_{u\cdot v,i} \int_0^1 \left[ \ddot{H}_i(1) - \ddot{H}_i(r) \right] dw_{u,i}(r) \right\} \\
= & \, d_{\Psi_i}
\end{aligned}$$

where $d_{\Psi_i}$ represents the $(i+1)^{th}$ $k \times 1$ block of the distribution $\Psi$.

Combining the above results shows that $\widetilde{\Omega_i^+}$ converges to

$$\begin{aligned}
\begin{bmatrix} 1 & -d_{\Psi_i}' \end{bmatrix} \begin{bmatrix} \sigma_{u\cdot v,i}^2 & 0 \\ 0 & \Omega_{vv,i} \end{bmatrix} \begin{bmatrix} 1 \\ -d_{\Psi_i} \end{bmatrix} &= \sigma_{u\cdot v,i}^2 + d_{\Psi_i}' \Omega_{vv,i} d_{\Psi_i} \\
&= \sigma_{u\cdot v,i}^2 \left( 1 + \sigma_{u\cdot v,i}^{-2} d_{\Psi_i}' \Omega_{vv,i} d_{\Psi_i} \right) \\
&= \sigma_{u\cdot v,i}^2 \left( 1 + d_{\gamma_i}' d_{\gamma_i} \right),
\end{aligned}$$

which leads to $\hat{\sigma}_{u\cdot v,i}^2 \Rightarrow \sigma_{u\cdot v,i}^2 \left( 1 + d_{\gamma_i}' d_{\gamma_i} \right)$. This implies that $\hat{V}_{PIM}$, using $\hat{\sigma}_{u\cdot v,i}^2$, converges to

$$\begin{aligned}
\hat{V}_{PIM} \Rightarrow & \left( \int_0^1 \sum_{i=1}^N h_i(r) h_i'(r) dr \right)^{-1} \times \\
& \left\{ \sum_{i=1}^N \sigma_{u\cdot v,i}^2 \left(1 + d_{\gamma_i}' d_{\gamma_i}\right) \int_0^1 \left[ \ddot{H}_i(1) - \ddot{H}_i(r) \right] \left[ \ddot{H}_i(1) - \ddot{H}_i(r) \right]' dr \right\} \times \\
& \left( \int_0^1 \sum_{i=1}^N h_i(r) h_i'(r) dr \right)^{-1} \\
= & V_{PIM} \left(1 + d_\gamma' d_r\right)
\end{aligned}$$

where

$$
\begin{aligned}
d_\gamma' d_\gamma &= V_{PIM}^{-1} \left( \int_0^1 \sum_{i=1}^N h_i(r) h_i'(r) dr \right)^{-1} \times \\
&\left\{ \sum_{i=1}^N \sigma_{u \cdot v,i}^2 \left( d_{\gamma_i}' d_{\gamma_i} \right) \int_0^1 \left[ \ddot{H}_i(1) - \ddot{H}_i(r) \right] \left[ \ddot{H}_i(1) - \ddot{H}_i(r) \right]' dr \right\} \times \\
&\left( \int_0^1 \sum_{i=1}^N h_i(r) h_i'(r) dr \right)^{-1} .
\end{aligned}
$$

The null limiting distribution of Wald and t statistics can be computed as follows.

$$
\begin{aligned}
\hat{W} &= \left(R\hat{\theta} - r\right)' \left[ R A_{PIM} \breve{V}_{PIM} A_{PIM} R' \right]^{-1} \left(R\hat{\theta} - r\right) \\
&= \left[ R \left( \hat{\theta} - \theta \right) \right]' \left[ R A_{PIM} \breve{V}_{PIM} A_{PIM} R' \right]^{-1} \left[ R \left( \hat{\theta} - \theta \right) \right] \\
&= \left[ A_R^{-1} R A_{PIM} A_{PIM}^{-1} \left( \hat{\theta} - \theta \right) \right]' \left[ A_R^{-1} R A_{PIM} \breve{V}_{PIM} A_{PIM} R' \left( A_R^{-1} \right)' \right]^{-1} \times \\
&\quad \left[ A_R^{-1} R A_{PIM} A_{PIM}^{-1} \left( \hat{\theta} - \theta \right) \right] \\
&\Rightarrow [R^* \Psi]' \left[ R^* V_{PIM} \left( 1 + d_\gamma' d_\gamma \right) (R^*)' \right]^{-1} [R^* \Psi] \\
&= \frac{\chi_q^2}{\left( 1 + d_\gamma' d_\gamma \right)}
\end{aligned}
$$

and for $q = 1$,

$$
\begin{aligned}
\hat{t} &= \frac{\left(R\hat{\theta} - r\right)}{\sqrt{R A_{PIM} \breve{V}_{PIM} A_{PIM} R'}} \\
&= \frac{\left[ R \left( \hat{\theta} - \theta \right) \right]}{\sqrt{R A_{PIM} \breve{V}_{PIM} A_{PIM} R'}} \\
&= \frac{A_R^{-1} R A_{PIM} A_{PIM}^{-1} \left( \hat{\theta} - \theta \right)}{\sqrt{A_R^{-1} R A_{PIM} \breve{V}_{PIM} A_{PIM} R' A_R^{-1}}} \\
&\Rightarrow \frac{R^* \Psi}{\sqrt{R^* V_{PIM} \left( 1 + d_\gamma' d_\gamma \right) (R^*)'}} \\
&= \frac{Z}{\sqrt{\left( 1 + d_\gamma' d_\gamma \right)}}.
\end{aligned}
$$

Second, consider the fixed-b limit of the $\hat{t}$ and $\hat{W}$. Recall that

$$\begin{aligned}
\hat{S}_{it}^{\tilde{u}} &= S_{it}^{\ddot{y}} - t\hat{\mu}_i - S_{it}^{\ddot{x}\prime}\hat{\beta} - \frac{N-1}{N}x_{it}'\hat{\gamma}_i + \frac{1}{N}\sum_{j=1,j\neq i}^{N} x_{jt}'\hat{\gamma}_j \\
&= t\mu_i + S_{it}^{\ddot{x}\prime}\beta + \frac{N-1}{N}x_{it}'\gamma_i - \frac{1}{N}\sum_{j=1,j\neq i}^{N} x_{jt}'\gamma_j + S_{it}^{\ddot{u}} \\
&\quad - t\hat{\mu}_i - S_{it}^{\ddot{x}\prime}\hat{\beta} - \frac{N-1}{N}x_{it}'\hat{\gamma}_i + \frac{1}{N}\sum_{j=1,j\neq i}^{N} x_{jt}'\hat{\gamma}_j \\
&\quad - \frac{N-1}{N}x_{it}'\Omega_{vv,i}^{-1}\Omega_{vu,i} + \frac{1}{N}\sum_{j=1,j\neq i}^{N} x_{jt}'\Omega_{vv,j}^{-1}\Omega_{vu,j} \\
&= S_{it}^{\ddot{u}} - \frac{N-1}{N}x_{it}'\Omega_{vv,i}^{-1}\Omega_{vu,i} + \frac{1}{N}\sum_{j=1,j\neq i}^{N} x_{jt}'\Omega_{vv,j}^{-1}\Omega_{vu,j} - q_{it}'\left(\hat{\theta}-\theta\right) \\
&= S_{it}^{u} - \frac{1}{N}\sum_{j=1}^{N} S_{jt}^{u} - x_{it}'\Omega_{vv,i}^{-1}\Omega_{vu,i} + \frac{1}{N}\sum_{j=1}^{N} x_{jt}'\Omega_{vv,j}^{-1}\Omega_{vu,j} - q_{it}'\left(\hat{\theta}-\theta\right)
\end{aligned}$$

where $q_{it}$ is defined above.

Then, the first difference of $\hat{S}_{it}^{\tilde{u}}$ can be written as

$$\begin{aligned}
\triangle\hat{S}_{it}^{\tilde{u}} &= \triangle S_{it}^{u} - \frac{1}{N}\sum_{j=1}^{N} \triangle S_{jt}^{u} - \triangle x_{it}'\Omega_{vv,i}^{-1}\Omega_{vu,i} + \frac{1}{N}\sum_{j=1}^{N} \triangle x_{jt}'\Omega_{vv,j}^{-1}\Omega_{vu,j} - \triangle q_{it}'\left(\hat{\theta}-\theta\right) \\
&= u_{it} - \frac{1}{N}\sum_{j=1}^{N} u_{jt} - v_{it}'\Omega_{vv,i}^{-1}\Omega_{vu,i} + \frac{1}{N}\sum_{j=1}^{N} v_{jt}'\Omega_{vv,j}^{-1}\Omega_{vu,j} - \triangle q_{it}'\left(\hat{\theta}-\theta\right).
\end{aligned}$$

Consequently,

$$\begin{aligned}
T^{-\frac{1}{2}}\sum_{t=1}^{[rT]}\triangle\hat{S}_{it}^{\tilde{u}} &= T^{-\frac{1}{2}}\sum_{t=1}^{[rT]}u_{it}-\frac{1}{N}\sum_{j=1}^{N}\left(T^{-\frac{1}{2}}\sum_{t=1}^{[rT]}u_{jt}\right)-T^{-\frac{1}{2}}x'_{i[rT]}\Omega_{vv,i}^{-1}\Omega_{vu,i} \\
&+\frac{1}{N}\sum_{j=1}^{N}T^{-\frac{1}{2}}x'_{j[rT]}\Omega_{vv,j}^{-1}\Omega_{vu,j}-T^{-\frac{1}{2}}q'_{i[rT]}\left(\hat{\theta}-\theta\right) \\
&= T^{-\frac{1}{2}}S_{i[rT]}^{u}-\frac{1}{N}\sum_{j=1}^{N}T^{-\frac{1}{2}}S_{j[rT]}^{u}-T^{-\frac{1}{2}}x'_{i[rT]}\Omega_{vv,i}^{-1}\Omega_{vu,i} \\
&+\frac{1}{N}\sum_{j=1}^{N}T^{-\frac{1}{2}}x'_{j[rT]}\Omega_{vv,j}^{-1}\Omega_{vu,j}-T^{-\frac{1}{2}}q'_{i[rT]}A_{PIM}A_{PIM}^{-1}\left(\hat{\theta}-\theta\right) \\
&\Rightarrow B_{u,i}(r)-\frac{1}{N}\sum_{j=1}^{N}B_{u,j}(r)-B'_{v,i}(r)\Omega_{vv,i}^{-1}\Omega_{vu,i} \\
&+\frac{1}{N}\sum_{j=1}^{N}B'_{v,j}(r)\Omega_{vv,j}^{-1}\Omega_{vu,j}-h'_{i}(r)\Psi \\
&= \sigma_{u\cdot v,i}w_{u,i}(r)+\lambda_{uv,i}W_{v,i}(r)-\frac{1}{N}\sum_{j=1}^{N}\left[\sigma_{u\cdot v,j}w_{u,j}(r)+\lambda_{uv,j}W_{v,j}(r)\right] \\
&-W'_{v,i}(r)\Omega_{vv,i}^{\frac{1}{2}\prime}\Omega_{vv,i}^{-1}\Omega_{vu,i}+\frac{1}{N}\sum_{j=1}^{N}\left[W'_{v,j}(r)\Omega_{vv,j}^{\frac{1}{2}\prime}\Omega_{vv,j}^{-1}\Omega_{vu,j}\right]-h'_{i}(r)\Psi \\
&= \sigma_{u\cdot v,i}w_{u,i}(r)-\frac{1}{N}\sum_{j=1}^{N}\left[\sigma_{u\cdot v,j}w_{u,j}(r)\right]-h'_{i}(r)\Psi \\
&= \sigma_{u\cdot v,i}\left[w_{u,i}(r)-\frac{1}{N}\sum_{j=1}^{N}\frac{\sigma_{u\cdot v,j}w_{u,j}(r)}{\sigma_{u\cdot v,i}}-\sigma_{u\cdot v,i}^{-1}h'_{i}(r)\Psi\right] \\
&= \sigma_{u\cdot v,i}\cdot\hat{P}_{i}(r)
\end{aligned}$$

where

$$
\begin{aligned}
\hat{P}_i(r) &= w_{u,i}(r) - \frac{1}{N}\sum_{j=1}^{N}\frac{\sigma_{u\cdot v,j}w_{u,j}(r)}{\sigma_{u\cdot v,i}} - \sigma_{u\cdot v,i}^{-1}h_i'(r)\Psi \\
&= w_{u,i}(r) - \frac{1}{N}\sum_{j=1}^{N}\frac{\sigma_{u\cdot v,j}w_{u,j}(r)}{\sigma_{u\cdot v,i}} \\
&- \sigma_{u\cdot v,i}^{-1}h_i'(r)\left(\int_0^1\sum_{i=1}^{N}h_i(r)h_i'(r)dr\right)^{-1}\times \\
&\left(\int_0^1\sum_{i=1}^{N}h_i(r)\left[\sigma_{u\cdot v,i}w_{u,i}(r) - \frac{1}{N}\sum_{j=1}^{N}\sigma_{u\cdot v,j}w_{u,j}(r)\right]dr\right).
\end{aligned}
$$

In sum,

$$
T^{-\frac{1}{2}}\sum_{t=1}^{[rT]}\triangle\hat{S}_{it}^{\tilde{u}} = T^{-\frac{1}{2}}\hat{S}_{i[rT]}^{\tilde{u}} \Longrightarrow \sigma_{u\cdot v,i}\hat{P}_i(r). \tag{3.10}
$$

Next, write $\hat{\sigma}_{u\cdot v,i}^2$ in terms of $T^{-\frac{1}{2}}\hat{S}_{i[rT]}^{\tilde{u}}$. The kernel function used here is the Bartlett kernel. Define

$$
\begin{aligned}
K_{ts} &= k\left(\frac{|t-s|}{M}\right) \\
\triangle^2 K_{ts} &= (K_{ts} - K_{t,s+1}) - (K_{t+1,s} - K_{t+1,s+1}).
\end{aligned}
$$

Simple algebra gives

$$
\begin{aligned}
\hat{\sigma}_{u\cdot v,i}^2 &= T^{-1}\sum_{j=2}^{T}\sum_{h=2}^{T}k\left(\frac{|j-h|}{M}\right)\triangle\hat{S}_{ij}^{\tilde{u}}\triangle\hat{S}_{ih}^{\tilde{u}} \\
&= T^{-1}\sum_{j=2}^{T}\left(\triangle\hat{S}_{ij}^{\tilde{u}}\sum_{h=2}^{T}K_{jh}\triangle\hat{S}_{ih}^{\tilde{u}}\right) = T^{-1}\sum_{j=2}^{T}a_jb_j
\end{aligned}
$$

where

$$
a_j = \triangle\hat{S}_{ij}^{\tilde{u}}, \quad b_j = \sum_{h=2}^{T}K_{jh}\triangle\hat{S}_{ih}^{\tilde{u}}.
$$

Using summation by parts we can write

$$\sum_{t=2}^{T} a_t b_t = \left(\sum_{s=1}^{T} a_s\right) b_T - a_1 b_2 + \sum_{t=2}^{T-1}\left[\left(\sum_{s=1}^{t} a_s\right)(b_t - b_{t+1})\right] \tag{3.11}$$

which gives

$$\begin{aligned}\hat{\sigma}^2_{u\cdot v,i} &= T^{-1}\hat{S}^{\tilde{u}}_{iT}\sum_{h=2}^{T} K_{Th}\triangle\hat{S}^{\tilde{u}}_{ih} - T^{-1}\hat{S}^{\tilde{u}}_{i1}\sum_{h=2}^{T} K_{2h}\triangle\hat{S}^{\tilde{u}}_{ih}\\ &+ T^{-1}\sum_{j=2}^{T-1}\left[\hat{S}^{\tilde{u}}_{ij}\left(\sum_{h=2}^{T} K_{jh}\triangle\hat{S}^{\tilde{u}}_{ih} - \sum_{h=2}^{T} K_{j+1,h}\triangle\hat{S}^{\tilde{u}}_{ih}\right)\right]\end{aligned}$$

We need to apply (3.11) to the sums over $h$:

$$(1)\ \sum_{h=2}^{T} K_{Th}\triangle\hat{S}^{\tilde{u}}_{ih} = \sum_{h=2}^{T}\triangle\hat{S}^{\tilde{u}}_{ih}K_{Th} = \hat{S}^{\tilde{u}}_{iT}K_{TT} - \hat{S}^{\tilde{u}}_{i1}K_{T2} + \sum_{h=2}^{T-1}\hat{S}^{\tilde{u}}_{ih}\left(K_{Th} - K_{T,h+1}\right)$$

$$(2)\quad \sum_{h=2}^{T} K_{2h}\triangle\hat{S}^{\tilde{u}}_{ih} = \sum_{h=2}^{T}\triangle\hat{S}^{\tilde{u}}_{ih}K_{2h} = \hat{S}^{\tilde{u}}_{iT}K_{2T} - \hat{S}^{\tilde{u}}_{i1}K_{22} + \sum_{h=2}^{T-1}\hat{S}^{\tilde{u}}_{ih}\left(K_{2h} - K_{2,h+1}\right)$$

$$(3)\quad \sum_{h=2}^{T} K_{jh}\triangle\hat{S}^{\tilde{u}}_{ih} = \sum_{h=2}^{T}\triangle\hat{S}^{\tilde{u}}_{ih}K_{jh} = \hat{S}^{\tilde{u}}_{iT}K_{jT} - \hat{S}^{\tilde{u}}_{i1}K_{j2} + \sum_{h=2}^{T-1}\hat{S}^{\tilde{u}}_{ih}\left(K_{jh} - K_{j,h+1}\right)$$

$$\begin{aligned}(4)\quad \sum_{h=2}^{T} K_{j+1,h}\triangle\hat{S}^{\tilde{u}}_{ih} &= \sum_{h=2}^{T}\triangle\hat{S}^{\tilde{u}}_{ih}K_{j+1,h}\\ &= \hat{S}^{\tilde{u}}_{iT}K_{j+1,T} - \hat{S}^{\tilde{u}}_{i1}K_{j+1,2} + \sum_{h=2}^{T-1}\hat{S}^{\tilde{u}}_{ih}\left(K_{j+1,h} - K_{j+1,h+1}\right).\end{aligned}$$

Plugging in these expressions to $\hat{\sigma}^2_{u\cdot v,i}$ gives

$$\begin{aligned}
\hat{\sigma}^2_{u\cdot v,i} &= T^{-1}\hat{S}^{\tilde{u}}_{iT}\left[\hat{S}^{\tilde{u}}_{iT}K_{TT} - \hat{S}^{\tilde{u}}_{i1}K_{T2} + \sum_{h=2}^{T-1}\hat{S}^{\tilde{u}}_{ih}\left(K_{Th} - K_{T,h+1}\right)\right] \\
&\quad -T^{-1}\hat{S}^{\tilde{u}}_{i1}\left[\hat{S}^{\tilde{u}}_{iT}K_{2T} - \hat{S}^{\tilde{u}}_{i1}K_{22} + \sum_{h=2}^{T-1}\hat{S}^{\tilde{u}}_{ih}\left(K_{2h} - K_{2,h+1}\right)\right] \\
&\quad +T^{-1}\sum_{j=2}^{T-1}\hat{S}^{\tilde{u}}_{ij}\left[\hat{S}^{\tilde{u}}_{iT}K_{jT} - \hat{S}^{\tilde{u}}_{i1}K_{j2} + \sum_{h=2}^{T-1}\hat{S}^{\tilde{u}}_{ih}\left(K_{jh} - K_{j,h+1}\right)\right] \\
&\quad -T^{-1}\sum_{j=2}^{T-1}\hat{S}^{\tilde{u}}_{ij}\left[\hat{S}^{\tilde{u}}_{iT}K_{j+1,T} - \hat{S}^{\tilde{u}}_{i1}K_{j+1,2} + \sum_{h=2}^{T-1}\hat{S}^{\tilde{u}}_{ih}\left(K_{j+1,h} - K_{j+1,h+1}\right)\right] \\
&= T^{-1}\hat{S}^{\tilde{u}}_{iT}K_{TT}\hat{S}^{\tilde{u}}_{iT} + T^{-1}\sum_{h=2}^{T-1}\hat{S}^{\tilde{u}}_{iT}(K_{Th} - K_{T,h+1})\hat{S}^{\tilde{u}}_{ih} \\
&\quad +T^{-1}\sum_{j=2}^{T-1}\hat{S}^{\tilde{u}}_{ij}(K_{jT} - K_{j+1,T})\hat{S}^{\tilde{u}}_{iT} + \text{terms related to}\hat{S}^{\tilde{u}}_{i1} \\
&\quad +T^{-1}\sum_{j=2}^{T-1}\sum_{h=2}^{T-1}\hat{S}^{\tilde{u}}_{ij}\left[(K_{jh} - K_{j,h+1}) - (K_{j+1,h} - K_{j+1,h+1})\right]\hat{S}^{\tilde{u}}_{ih} \\
&= T^{-1}\sum_{j=2}^{T-1}\sum_{h=2}^{T-1}\hat{S}^{\tilde{u}}_{ij}\triangle^2 K_{jh}\hat{S}^{\tilde{u}}_{ih} + T^{-1}\sum_{j=2}^{T-1}\hat{S}^{\tilde{u}}_{ij}(K_{jT} - K_{j+1,T})\hat{S}^{\tilde{u}}_{iT} \\
&\quad +T^{-1}\sum_{h=2}^{T-1}\hat{S}^{\tilde{u}}_{iT}(K_{Th} - K_{T,h+1})\hat{S}^{\tilde{u}}_{ih} + T^{-1}\hat{S}^{\tilde{u}}_{iT}K_{TT}\hat{S}^{\tilde{u}}_{iT} + \text{terms related to}\hat{S}^{\tilde{u}}_{i1}
\end{aligned}$$

Note that the terms related to $\hat{S}^{\tilde{u}}_{i1}$ vanish as $T \to \infty$, because $T^{-\frac{1}{2}}\hat{S}^{\tilde{u}}_{i1}$ converges to $\sigma_{u\cdot v,i}\hat{P}_i(0)$, which equals zero. For the Bartlett kernel we have

$$K_{ts} = k\left(\frac{|t-s|}{M}\right) = \begin{cases} 1 - \frac{|t-s|}{M}, & |t-s| \leqslant M \\ 0 & |t-s| > M \end{cases}$$

Then it follows that

$$K_{ts} - K_{t,s+1} = \begin{cases} 0, & t \leqslant s - M \\ \frac{1}{M}, & s+1-M \leqslant t \leqslant s \\ -\frac{1}{M}, & s+1 \leqslant t \leqslant s+M \\ 0, & t \geqslant s+M+1 \end{cases}$$

$$K_{t+1,s}-K_{t+1,s+1}=\begin{cases}0, & t\leqslant s-M-1\\ \frac{1}{M}, & s-M\leqslant t\leqslant s-1\\ -\frac{1}{M}, & s\leqslant t\leqslant s-1+M\\ 0, & t\geqslant s+M\end{cases}$$

and

$$\triangle^2 K_{ts}=\begin{cases}\frac{2}{M}, & t=s\\ -\frac{1}{M}, & t=s\pm M\\ 0, & \text{otherwise}\end{cases}$$

Using these result we have

$$\begin{aligned}\hat{\sigma}^2_{u\cdot v,i} &= T^{-1}\left[\frac{2}{M}\sum_{j=2}^{T-1}\hat{S}^{\tilde{u}}_{ij}\hat{S}^{\tilde{u}}_{ij}-\frac{1}{M}\sum_{j=2}^{T-M-1}\left(\hat{S}^{\tilde{u}}_{i,j+M}\hat{S}^{\tilde{u}}_{ij}+\hat{S}^{\tilde{u}}_{ij}\hat{S}^{\tilde{u}}_{i,j+M}\right)\right]\\ &+T^{-1}\left[-\frac{1}{M}\sum_{j=T-M}^{T-1}\hat{S}^{\tilde{u}}_{ij}\hat{S}^{\tilde{u}}_{iT}-\frac{1}{M}\sum_{h=T-M}^{T-1}\hat{S}^{\tilde{u}}_{iT}\hat{S}^{\tilde{u}}_{ih}\right]\\ &+T^{-1}\hat{S}^{\tilde{u}}_{iT}\hat{S}^{\tilde{u}}_{iT}+\text{terms related to }\hat{S}^{\tilde{u}}_{i1}\\ &=\frac{2}{MT}\sum_{j=2}^{T-1}\hat{S}^{\tilde{u}}_{ij}\hat{S}^{\tilde{u}}_{ij}-\frac{2}{MT}\sum_{j=2}^{T-M-1}\hat{S}^{\tilde{u}}_{ij}\hat{S}^{\tilde{u}}_{i,j+M}-\frac{2}{MT}\sum_{j=T-M}^{T-1}\hat{S}^{\tilde{u}}_{ij}\hat{S}^{\tilde{u}}_{iT}\\ &+T^{-1}\hat{S}^{\tilde{u}}_{iT}\hat{S}^{\tilde{u}}_{iT}+\text{terms related to }\hat{S}^{\tilde{u}}_{i1},\end{aligned}$$

where the last term follows from the fact that $K_{TT}=1$.

Under fixed-b asymptotics we set $M=bT$ where $b\in(0,1]$ is held fixed as $T\to\infty$. Plugging in $bT$ for $M$ into $\hat{\sigma}^2_{u\cdot v,i}$ gives

$$\begin{aligned}\hat{\sigma}^2_{u\cdot v,i} &= \frac{2}{bT}\sum_{j=2}^{T-1}T^{-\frac{1}{2}}\hat{S}^{\tilde{u}}_{ij}T^{-\frac{1}{2}}\hat{S}^{\tilde{u}}_{ij}-\frac{2}{bT}\sum_{j=2}^{T-bT-1}T^{-\frac{1}{2}}\hat{S}^{\tilde{u}}_{ij}T^{-\frac{1}{2}}\hat{S}^{\tilde{u}}_{i,j+M}\\ & -\frac{2}{bT}\sum_{j=T-bT}^{T-1}T^{-\frac{1}{2}}\hat{S}^{\tilde{u}}_{ij}T^{-\frac{1}{2}}\hat{S}^{\tilde{u}}_{iT}+T^{-\frac{1}{2}}\hat{S}^{\tilde{u}}_{iT}T^{-\frac{1}{2}}\hat{S}^{\tilde{u}}_{iT}\\ & +\text{terms related to }T^{-\frac{1}{2}}\hat{S}^{\tilde{u}}_{i1}.\end{aligned}$$

Using (3.10) and the continuous mapping theorem gives

$$
\begin{aligned}
\hat{\sigma}^2_{u\cdot v,i} &\Rightarrow \frac{2}{b}\int_0^1 \left[\sigma_{u\cdot v,i}\hat{P}_i(r)\right]^2 dr - \frac{2}{b}\int_0^{1-b} \sigma_{u\cdot v,i}\hat{P}_i(r)\sigma_{u\cdot v,i}\hat{P}_i(r+b)dr \\
&\quad -\frac{2}{b}\int_{1-b}^1 \sigma_{u\cdot v,i}\hat{P}_i(r)\sigma_{u\cdot v,i}\hat{P}_i(1)dr + \left[\sigma_{u\cdot v,i}\hat{P}_i(1)\right]^2 \\
&= \sigma^2_{u\cdot v,i}\left[\frac{2}{b}\int_0^1 \hat{P}_i^2(r)dr - \frac{2}{b}\int_0^{1-b} \hat{P}_i(r)\hat{P}_i(r+b)dr - \frac{2}{b}\int_{1-b}^1 \hat{P}_i(r)\hat{P}_i(1)dr + \hat{P}_i^2(1)\right] \\
&= \sigma^2_{u\cdot v,i}Q_b\left(\hat{P}_i(r)\right)
\end{aligned}
$$

where

$$
Q_b\left(\hat{P}_i(r)\right) = \frac{2}{b}\int_0^1 \hat{P}_i^2(r)dr - \frac{2}{b}\int_0^{1-b} \hat{P}_i(r)\hat{P}_i(r+b)dr - \frac{2}{b}\int_{1-b}^1 \hat{P}_i(r)\hat{P}_i(1)dr + \hat{P}_i^2(1).
$$

Therefore, based on $\hat{\sigma}^2_{u\cdot v,i}$, the fixed-b limit of the covariance matrix is given by

$$
\begin{aligned}
\hat{V}_{PIM} \Rightarrow &\left(\int_0^1 \sum_{i=1}^N h_i(r)h_i'(r)dr\right)^{-1} \times \\
&\left\{\sum_{i=1}^N \sigma^2_{u\cdot v,i}Q_b\left(\hat{P}_i(r)\right)\int_0^1 \left[\ddot{H}_i(1)-\ddot{H}_i(r)\right]\left[\ddot{H}_i(1)-\ddot{H}_i(r)\right]' dr\right\} \times \\
&\left(\int_0^1 \sum_{i=1}^N h_i(r)h_i'(r)dr\right)^{-1} \\
=&V_{PIM}\cdot\hat{Q}(b)
\end{aligned}
$$

where

$$
\begin{aligned}
\hat{Q}(b) = &V_{PIM}^{-1}\left(\int_0^1 \sum_{i=1}^N h_i(r)h_i'(r)dr\right)^{-1} \times \\
&\left\{\sum_{i=1}^N \sigma^2_{u\cdot v,i}Q_b\left(\hat{P}_i(r)\right)\int_0^1 \left[\ddot{H}_i(1)-\ddot{H}_i(r)\right]\left[\ddot{H}_i(1)-\ddot{H}_i(r)\right]' dr\right\} \times \\
&\left(\int_0^1 \sum_{i=1}^N h_i(r)h_i'(r)dr\right)^{-1}.
\end{aligned}
$$

This implies that

$$\hat{W} \Rightarrow \frac{\chi_q^2}{\hat{Q}(b)},$$

and for $q = 1$,

$$\hat{t} \Rightarrow \frac{Z}{\sqrt{\hat{Q}(b)}}.$$

Lastly, consider the result for the fixed-b test statistics. Similar as above and Vogelsang and Wagner [2014], $\tilde{\sigma}^2_{u\cdot v,i} \Rightarrow \sigma^2_{u\cdot v,i} Q_b\left(\tilde{P}_i(r)\right)$ where $Q_b(\cdot)$ is the same as above, and $\tilde{P}_i(r)$ is similar as $\hat{P}_i(r)$ but its component is from the further augment regression (3.9). Therefore, the fixed-b limit of the covariance matrix is such that

$$\begin{aligned}
\tilde{V}_{PIM} \Rightarrow & \left(\int_0^1 \sum_{i=1}^N h_i(r) h_i'(r) dr\right)^{-1} \times \\
& \left\{\sum_{i=1}^N \sigma^2_{u\cdot v,i} Q_b\left(\tilde{P}_i(r)\right) \int_0^1 \left[\ddot{H}_i(1) - \ddot{H}_i(r)\right]\left[\ddot{H}_i(1) - \ddot{H}_i(r)\right]' dr\right\} \times \\
& \left(\int_0^1 \sum_{i=1}^N h_i(r) h_i'(r) dr\right)^{-1} \\
= & V_{PIM} \cdot \tilde{Q}(b)
\end{aligned}$$

where

$$\begin{aligned}
\tilde{Q}(b) = & V_{PIM}^{-1} \left(\int_0^1 \sum_{i=1}^N h_i(r) h_i'(r) dr\right)^{-1} \times \\
& \left\{\sum_{i=1}^N \sigma^2_{u\cdot v,i} Q_b\left(\tilde{P}_i(r)\right) \int_0^1 \left[\ddot{H}_i(1) - \ddot{H}_i(r)\right]\left[\ddot{H}_i(1) - \ddot{H}_i(r)\right]' dr\right\} \times \\
& \left(\int_0^1 \sum_{i=1}^N h_i(r) h_i'(r) dr\right)^{-1}.
\end{aligned}$$

This implies that

$$\tilde{W} \Rightarrow \frac{\chi_q^2}{\tilde{Q}(b)},$$

and for $q = 1$,

$$\tilde{t} \Rightarrow \frac{Z}{\sqrt{\tilde{Q}(b)}}.$$

# BIBLIOGRAPHY

# BIBLIOGRAPHY


Donald W. K. Andrews. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, 59(3):817–858, May 1991.

Nicholas Apergis and James E. Payne. Energy consumption and economic growth in central america: Evidence from a panel cointegration and error correction model. *Energy Economics*, 31(2):211–216, 2009.

Jushan Bai and Chihwa Kao. Chapter 1 on the estimation and inference of a panel cointegration model with cross-sectional dependence. In Badi H. Baltagi, editor, *Panel Data EconometricsTheoretical Contributions and Empirical Applications*, volume 274 of *Contributions to Economic Analysis*, pages 3–30. Elsevier, 2006.

Jushan Bai and Serena Ng. A panic attack on unit roots and cointegration. *Econometrica*, 72(4):1127–1177, July 2004.

Jushan Bai, Chihwa Kao, and Serena Ng. Panel cointegration with global stochastic trends. *Journal of Econometrics*, 149(1):82–99, 2009.

Matthew B Canzoneri, Robert E Cumby, and Behzad Diba. Relative labor productivity and the real exchange rate in the long run: evidence for a panel of oecd countries. *Journal of International Economics*, 47(2):245–266, 1999.

Yoosoon Chang. Bootstrap unit root tests in panels with cross-sectional dependency. *Journal of Econometrics*, 120(2):263–293, 2004.

Yoosoon Chang, Joon Y. Park, and Kevin Song. Bootstrapping cointegrating regressions. *Journal of Econometrics*, 133(2):703–739, 2006.

In Choi and Eiji Kurozumi. Model selection criteria for the leads-and-lags cointegrating regression. *Journal of Econometrics*, 169(2):224–238, August 2012.

Dimitris K Christopoulos and Efthymios G Tsionas. Financial development and economic growth: evidence from panel unit root and cointegration tests. *Journal of Development Economics*, 73(1):55–74, 2004.

Lung fei Lee and Jihai Yu. Estimation of spatial autoregressive panel data models with fixed effects. *Journal of Econometrics*, 154(2):165–185, February 2010.

Chihwa Kao and Min-Hsien Chiang. On the estimation and inference of a cointegrated regression in panel data. In *Nonstationary Panels, Panel Cointegration, and Dynamic Panels*, volume 1st, chapter 7, pages 179–222. Emerald Group Publishing Limited, 2000.

G. Kapetanios, M. Hashem Pesaran, and T. Yamagata. Panels with non-stationary multifactor error structures. *Journal of Econometrics*, 160(2):326–348, February 2011.

Mudit Kapoor, Harry H. Kelejian, and Ingmar R. Prucha. Panel data models with spatially correlated error components. *Journal of Econometrics*, 140(1):97–130, September 2007. Analysis of spatially dependent data.

Mohitosh Kejriwal and Pierre Perron. Data dependent rules for selection of the number of leads and lags in the dynamic ols cointegrating regression. *Econometric Theory*, 24(5): 1425–1441, October 2008.

Chien-Chiang Lee. Energy consumption and gdp in developing countries: A cointegrated panel analysis. *Energy Economics*, 27(3):415–427, 2005.

Hongyi Li and G.S. Maddala. Bootstrapping cointegrating regressions. *Journal of Econometrics*, 80(2):297–318, 1997.

Yi Li. Hypothesis testing in cointegrated panels: Asymptotic and bootstrap method. 2016.

Nelson C. Mark and Donggyu Sul. Cointegration vector estimation by panel dols and long-run money demand. *Oxford Bulletin of Economics & Statistics*, 65(5):655–680, 2003.

Suzanne McCoskey and Chihwa Kao. A residual-based test of the null of cointegration in panel data. *Econometric Reviews*, 17(1):57–84, 1998.

Paresh Kumar Narayan and Russell Smyth. Energy consumption and real gdp in g7 countries: New evidence from panel cointegration with structural breaks. *Energy Economics*, 30(5): 2331–2341, 2008.

Efstathios Paparoditis and Dimitris N. Politis. Residual-based block bootstrap for unit root testing. *Econometrica*, 71(3):813–855, 2003.

Efstathios Paparoditis and Dimitris N Politis. Bootstrapping unit root tests for autoregressive time series. *Journal of the American Statistical Association*, 100(470):545–553, 2005.

Cameron Parker, Efstathios Paparoditis, and Dimitris N. Politis. Unit root testing via the stationary bootstrap. *Journal of Econometrics*, 133(2):601–638, 2006.

Andrew Patton, Dimitris N. Politis, and Halbert White. Correction to "automatic block-length selection for the dependent bootstrap" by d. politis and h. white. *Econometric Reviews*, 28(4):372–375, 2009.

Peter Pedroni. Critical values for cointegration tests in heterogeneous panels with multiple regressors. *Oxford Bulletin of Economics & Statistics*, 61(4):653, 1999.

Peter Pedroni. Fully modified ols for heterogeneous cointegrated panels. In *Nonstationary Panels, Panel Cointegration, and Dynamic Panels.*, volume 1st, chapter 4, pages 93–130. Emerald Group Publishing Limited, 2000.

Peter Pedroni. Panel cointegration: Asymptotic and finite sample properties of pooled time series tests with an application to the ppp hypothesis. *Econometric Theory*, 20(3):597–625, 2004.

M. Hashem Pesaran. Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica*, 74(4):967–1012, July 2006.

Peter C. B. Phillips and Bruce E. Hansen. Statistical inference in instrumental variables regression with i(1) processes. *The Review of Economic Studies*, 57(1):99–125, January 1990.

Dimitris N. Politis and Joseph P. Romano. The stationary bootstrap. *Journal of the American Statistical Association*, 89(428):1303–1313, 1994.

Dimitris N. Politis and Halbert White. Automatic block-length selection for the dependent bootstrap. *Econometric Reviews*, 23(1):53–70, 2004.

Zacharias Psaradakis. On bootstrap inference in cointegrating regressions. *Economics Letters*, 72(1):1–10, 2001.

Dong Wan Shin. Stationary bootstrapping for panel cointegration tests under cross-sectional dependence. *Statistics*, 49(1):209–223, 2015.

Dong Wan Shin and Eunju Hwang. Stationary bootstrapping for cointegrating regressions. *Statistics & Probability Letters*, 83(2):474–480, 2013.

Stephan Smeekes and Jean-Pierre Urbain. On the applicability of the sieve bootstrap in time series panels. *Oxford Bulletin of Economics and Statistics*, 76(1):139–151, 2014.

Anders Rygh Swensen. Bootstrapping unit root tests for integrated processes. *Journal of Time Series Analysis*, 24(1):99–126, 2003.

Timothy J. Vogelsang and Martin Wagner. Integrated modified ols estimation and fixed-inference for cointegrating regressions. *Journal of Econometrics*, 178(2):741–760, 2014.

Timothy J. Vogelsang, Martin Wagner, and Yi Li. Integrated modified ols estimation and fixed-b inference for homogeneous cointegrated panels. 2016.

Joakim Westerlund. New simple tests for panel cointegration. *Econometric Reviews*, 24(3): 297–316, 2005.

Joakim Westerlund and David L. Edgerton. A panel bootstrap cointegration test. *Economics Letters*, 97(3):185–190, 2007.

Joakim Westerlund and David L. Edgerton. A simple test for cointegration in dependent panels with structural breaks. *Oxford Bulletin of Economics and Statistics*, 70(5):665–704, 2008.

Jihai Yu, Robert de Jong, and Lung fei Lee. Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both n and t are large. *Journal of Econometrics*, 146(1):118–134, September 2008.

Jihai Yu, Robert de Jong, and Lung fei Lee. Estimation for spatial dynamic panel data with fixed effects: The case of spatial cointegration. *Journal of Econometrics*, 167(1):16–37, March 2012.