ESTIMATION AND INFERENCE FOR VARYING COEFFICIENT MULTIDIMENSIONAL FIXED-EFFECTS PANEL DATA MODELS

Abstract. This paper presents a general estimation method for a varying coefficient multidimensional panel data regression model and offers an array of hypothesis testing avenues. We derive the asymptotic distribution of our estimator, and to construct valid tests, we develop the necessary central limit theory to conduct inference. The presence of multiple effects over differing dimensions requires nontrivial changes to the central limit theory for U-statistics. The types of inference we can conduct offer a diverse array of hypotheses for applied work and we explicitly present test statistics for some of the most important hypothesis tests. A detailed set of simulations supports our estimator's asymptotic developments and reveals that our testing infrastructure possess correct asymptotic size and high power.

1. Introduction

The big data revolution has motivated the rapid emergence of three, four and even higher dimensional panel data sets. This allows researchers to consider complex interconnectedness of big data sets due to, for example, the presence of externalities, spillovers or common shocks. A multidimensional approach is an essential tool for empirical analysis in many scientific milieu (i.e., economics, finance, criminology, etc.). Suppose we have a panel data set $\{(X_{ij\dots lt}, Z_{ij\dots lt}, Y_{ij\dots lt}) : i = 1,\dots,N_1; j = 1,\dots,N_2; \dots; l = 1,\dots,N_l; t = 1,\dots,T\}$, where the index t uniquely defines the time period and indices $i, j, \dots, l$ typically define the units of observations, measured over time. Without further knowledge of how the explanatory variables are related to the dependent variable, one may begin studying the relationship between these variables using the following nonparametric panel data regression model[1]

$$Y_{ij\dots lt} = \beta(X_{ij\dots lt}, Z_{ij\dots lt}) + \pi_{ij\dots lt} + v_{ij\dots lt}, \tag{1.1}$$

where $Y_{ij\dots lt}$ is the response variable corresponding to the $(i, j, \dots, l)$th observation within the groups $(i, j, \dots, l)$ at time t, $X_{ij\dots lt}$ and $Z_{ij\dots lt}$ are $d \times 1$ and $q \times 1$ vectors of covariates, respectively, $\beta(\cdot)$ is a smooth unknown function to be estimated, $\pi_{ij\dots lt}$ captures the unobserved

Date: February 17, 2022.
Christopher F. Parmeter, Department of Economics, University of Miami; e-mail: cparmeter@bus.miami.edu.
* All errors are ours alone. All R code in this paper is available at `the-smooth-operators.com`. This work is part of the I+D+i project Ref. PID2019-105986GB-C22, funded by MCIN/AEI/ 10.13039/501100011033. In addition, this work was partially complete while professors Henderson and Soberon were visiting the School of Mathematical Sciences at Nankai University in Tianjin, China. The usual disclaimer applies.
[1]See, for example, Ai and Li (2008), Henderson and Parmeter (2015), Parmeter and Racine (2019), Rodriguez-Poo and Soberon (2017) and/or Su and Ullah (2011) for intensive reviews of one-way and two-way nonparametric fixed-effects panel data models.

heterogeneity that enables controlling for all unobserved random and/or interactive effects (i.e., individual and/or time-specific effects, interactive effects among individuals and/or individuals and time, and so on), and $v_{ij\ldots lt}$ is the *i.i.d.* idiosyncratic disturbance term.

While there is relatively rich literature on multidimensional linear panel data models (see Matyas (2017) and the references therein), little work has been done in estimating multidimensional nonparametric models given the difficulty of removing potential correlation between unobserved heterogeneity and the covariates. For a three-dimensional setting, Galvao and Montes-Rojas (2017) consider estimation of a multidimensional quantile regression model, while Sun et al. (2017) develop an estimation procedure to estimate the smooth function in a nonparametric specification. More recently, Henderson et al. (2021) provide estimators for the corresponding gradients of the unknown conditional mean.

While it is well-known that nonparametric models are useful to guard against various forms of model misspecification, they exhibit several weaknesses that can lead to misleading inference in empirical studies. First, in many settings, nonparametric methods are unable to incorporate prior information coming from economic theory or past experience so the resulting estimators tend to have high variance. Second, they are subject to the so-called 'curse of dimensionality', which practically disables their use when the number of covariates is high. Third, some identification problems emerge when differencing transformations are used to remove fixed effects.

Varying-coefficient models have the potential to overcome these shortcomings. Assuming that the regression coefficients vary depending on some exogenous continuous variables proposed by economic theory, the model is of the form

$$Y_{ij\ldots lt} = X_{ij\ldots lt}^{\top}\beta(Z_{ij\ldots lt}) + \pi_{ij\ldots lt} + v_{ij\ldots lt}, \tag{1.2}$$

where we keep the same assumptions about the random variables explained above. These models are increasingly popular because they encompass properties of other statistical models of interest such as fully nonparametric and parametric models. This fact is especially appealing for empirical studies. Specifically, a special case of the above model is a standard linear panel data regression model

$$Y_{ij\ldots lt} \quad = \quad \delta + X_{ij\ldots lt}^{\top}\beta_1 + Z_{ij\ldots lt}^{\top}\beta_2 + \pi_{ij\ldots lt} + v_{ij\ldots lt}, \tag{1.3}$$

where δ is the intercept term and β_1 and β_2 are $d\times 1$ and $q\times 1$ vectors of unknown parameters to estimate, respectively. However, despite its potential, we have not found any existing literature working on multidimensional varying coefficient panel data models.

The question of whether to use a fully parametric, nonparametric or varying coefficient specification emerges naturally. In this paper, we consider several functional form testing problems in a multidimensional varying coefficient panel data regression model. To motivate the proposed tests, we first review the estimation procedure for fully parametric multidimensional panel data models with fixed effects, and later provide a profile likelihood estimator for the varying coefficient specification which controls for all levels of unobserved heterogeneity along with the necessary architecture to conduct valid inference in this model. Nearly all of the current literature in semi/nonparametric estimation/inference focuses exclusively on a single source of unobserved heterogeneity, namely fixed effects in the individual (i) dimension. However, our approach generalizes this to any dimension of fixed effects, both individually as well as interactive fixed effects.

In order to conduct inference for this type of model, we extend the result of Hall (1984) to the multidimensional setting. This general CLT will not only be useful for our setting, but in similar settings where there are multiple indexes on outcomes of interest. We apply our CLT to three common tests of interest in a panel data setting. First, we offer a test of correct parametric specification of the conditional mean. Second, we provide a test of statistical significance for a set of the regressors. Finally, a Hausman-type test is provided to address the typical trade-off between the choice of random versus fixed effects.

The rest of the paper is structured as follows. In Section 2, we introduce the multidimensional panel data model and discuss estimation using a semiparametric smooth coefficient structure. Section 3 presents our CLT for a general r^{th}-order degenerate U-statistic while Section 4 applies the new CLT to several prominent and useful hypotheses for empirical work with multidimensional panels. Section 5 details the finite sample performance of both the estimator and the testing architecture developed here. Finally, Section 6 offers concluding remarks and directions for future research. All proofs can be found in the Appendix.

Before going further, we introduce notation frequently used in this paper: (i) C denotes a generic positive constant that may take different values at different places; (ii) we denote $\mathbb{N} = N_1 N_2 \cdots N_l$, $\sum_{ij\cdots l} = \sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\cdots\sum_{l=1}^{N_l}$, $\sum_{(i'j'\ldots l't')\neq(ij\ldots lt)} = \sum_{i'\neq i}^{N_1}\sum_{j\neq j'}^{N_2}\cdots\sum_{l'\neq l} N_l \sum_{t\neq t'}^{T}$; (iii) for $r = 1, \ldots, l$, we denote I_{N_r} and I_T as $N_r \times N_r$ and $T \times T$ identity matrices, respectively, whereas $\imath_d$, $\imath_{N_l}$, $\imath_T$ are $d \times 1$, $N_l \times 1$, and $T \times 1$ vectors of ones, respectively. Note that this notation simplifies the discussion on the asymptotic properties of the estimators and test statistics since it implicitly allows us to write $\mathbb{N} \to \infty$ when all N_r diverge to infinity, or when some of these dimensions are fixed and the others grow to infinity.

2. Empirical Examples

Many model structures are included as special cases of the specifications considered in this paper. To develop a better understanding, we detail several empirical studies which embody features of our general varying coefficient multidimensional panel data model.

2.1. **Institutions and Economic Development.** Gregg (2020) constructs factory-level panel data for Russian factories from 1894–1908 to study whether or not factories owned by corporations are more productive than unincorporated factories. As the decision to incorporate is likely related to unobserved traits of the factory, the author uses fixed-effects to control for factory, industry and time heterogeneity. This is accomplished via a three-dimensional panel data model. The results from Table 5 of Gregg (2020) can be generalized as

$$Y_{ijt} = \beta\left(X_{ijt}, Z_{ijt}; \phi\right) + \mu_{1i} + \mu_{2j} + \mu_{3t} + v_{ijt}, \quad i = 1, \ldots, N_1, \quad j = 1, \ldots, N_2, \quad t = 1, \ldots, T \quad (2.1)$$

where Y_{ijt} is one of three productivity measures (log of revenue per worker, log of total machine power or total factor productivity) for factory i in industry j in year t. $\beta(\cdot)$ is assumed to be a known function of X_{ijt} and Z_{ijt} which simply amounts to an intercept term and a binary variable indicating whether or not the factory is owned by a corporation. The coefficient vector for the two parameters is represented by ϕ. v_{ijt} is the usual idiosyncratic error term. This model includes fixed effects for each dimension so $\pi_{ijt} = \mu_{1i} + \mu_{2j} + \mu_{2t}$, where μ_{1i} are factory effects, μ_{2j} are industry effects, and μ_{3t} are time effects.

2.2. **Price Dispersion.** Gerardi and Shapiro (2009) look at the airline industry to study the effects of market structure on price dispersion. They find that price dispersion increases with competition by considering two different multidimensional panel data models with similar structures. Formally, Equations (1) and (2) in their paper can be loosely summarized by

$$Y_{ijt} = \beta\left(X_{ijt}, Z_{ijt}; \phi\right) + \gamma_{ij} + \mu_t + v_{ijt}, \quad i = 1, \ldots, N_1, \quad j = 1, \ldots, N_2, \quad t = 1, \ldots, T \quad (2.2)$$

where Y_{ijt} is the Gini log-odds ratio (their measure of price dispersion) for carrier i, across route j, in time period t. $\beta(\cdot)$ is assumed to be a known function of the vectors X_{ijt} and Z_{ijt}, which include variables such as market share and market density, ϕ is a finite dimensional parameter vector, and v_{ijt} is the usual idiosyncratic error term. This model requires fixed effects which are unique to the ij pairs so $\pi_{ijt} = \gamma_{ij} + \mu_t$, where γ_{ij} are defined as carrier-route fixed effects and μ_t are time effects.

2.3. **Migration.** Bryan and Morten (2019) study the impact of internal labor migration (across provinces) in Indonesia. They find that reducing barriers to migration leads to significant productivity boosts (with substantial heterogeneity). Their three-dimensional panel data models allow for several types of heterogeneity. For example, in their Table 2, they give the estimates for a model of the form:

$$Y_{ijt} = \beta\left(X_{ijt}, Z_{ijt}; \phi\right) + \gamma_{1it} + \gamma_{2jt} + v_{ijt}, \quad i = 1, \ldots, N_1, \quad j = 1, \ldots, N_2, \quad t = 1, \ldots, T \qquad (2.3)$$

where Y_{ijt} is the log of the average wage within the origin (i) destination (j) pair, in year t. $\beta(\cdot)$ is assumed to be a known function of the vectors X_{ijt} and Z_{ijt} which consists of the proportion of people from each origin that move to each destination in a given year as well as the distance between provinces. ϕ is the finite dimensional parameter vector and v_{ijt} is the idiosyncratic error term. $\pi_{ijt} = \gamma_{1it} + \gamma_{2jt}$, where γ_{1it} are origin-time fixed effects, and γ_{2jt} represent destination-time fixed effects.

2.4. **International Trade.** Baier and Bergstrand (2007) argue that free trade agreements are endogenous when trying to predict trade flows between nations, but that they can be addressed properly via multidimensional panel data models which include the appropriate set of fixed effects. They estimate many specifications, but we focus on their Equation (11) which consists completely of fixed effects of multiple indexes

$$Y_{ijt} = \beta\left(X_{ijt}, Z_{ijt}; \phi\right) + \gamma_{1ij} + \gamma_{2it} + \gamma_{3jt} + v_{ijt}, \quad i = 1, \ldots, N_1, \quad j = 1, \ldots, N_2, \quad t = 1, \ldots, T \qquad (2.4)$$

where Y_{ijt} is the value of trade flow from (exporter) country i to (importer) country j in time period t. $\beta(\cdot)$ is the gravity function dependent upon the vectors X_{ijt} and Z_{ijt} which include GDP of the exporting and importing nations along with distance measures, ϕ is the finite dimensional parameter vector, and v_{ijt} is the idiosyncratic error term. $\pi_{ijt} = \gamma_{1ij} + \gamma_{2it} + \gamma_{3jt}$, where γ_{1ij} are bilateral fixed effects and γ_{2it} and γ_{3jt} are exporter-time and importer-time fixed effects, respectively.

2.5. **Hedonic Price Models.** While many empirical multidimensional panel datasets are three-dimensional, more advanced data collection and monitoring can bring about higher-order features. For example, Baltagi and Etienne (2015) use a five-dimensional panel data set to study hedonic price functions. Their research looks at apartment (*flat*) sales in Paris, France where the city is divided into twenty *arrondissements*, which are each broken into four quarters (*quartiers*), and each quarter contains multiple blocks (*îlots*). More formally, their model (Equation (4) in their paper) can be written as

$$\begin{aligned} Y_{ijmlt} = \beta\left(X_{ijmlt}, Z_{ijmlt}; \phi\right) + \gamma_{jt} + \lambda_{1jmt} + \lambda_{2jmlt} + v_{ijmlt}, \qquad (2.5)\\ i = 1, 2, \ldots, N_1; j = 1, 2, \ldots, N_2; m = 1, 2, \ldots, N_3; l = 1, 2, \ldots, N_4; t = 1, 2, \ldots, T \end{aligned}$$

where Y_{ijmlt} is the log selling price of apartment i, in *arrondissesment* j, in quarter m, on block l at time period t. $\beta(\cdot)$ is the hedonic price function which is dependent upon the vectors X_{ijmlt} and Z_{ijmlt} which include attributes unique to a given *flat* such as the surface area of the apartment or the number of bedrooms; ϕ, again, is the finite dimensional parameter vector, and v_{ijmlt} is the usual idiosyncratic disturbance term.

As there is a natural ordering of the panel dimension (apartment i is on block l which is in quarter m which is in *arrondissesment* j), the fixed effects must take that information into account so $\pi_{ijmlt} = \gamma_{jt} + \lambda_{1jmt} + \lambda_{2jmlt}$, where γ_{jt} is the *arrondissement*-time fixed effect, λ_{1jmt} is the quarter-time fixed effect, which is naturally nested in the respective *arrondissement*, and λ_{2jmlt} is the block-time fixed effect, again naturally nested in the respective quarter.

3. Estimation Method

We begin this section by outlining how to estimate a general-order parametric model and highlight how to construct the matrices for various composite fixed effects parameters (including the ones in the previous section). We then propose a profile estimator for our multidimensional varying coefficient fixed effect estimator and develop its asymptotic properties.

3.1. **Fully parametric panel data model.** In this subsection, we first review the estimation procedure for a general multidimensional linear regression model (1.3) when the unobserved heterogeneity $\pi_{ij\ldots lt}$ is allowed to contain additive and/or interactive fixed effects. This result will be helpful in motivating our semiparametric estimator for $\beta(\cdot)$ in the varying coefficient specification (1.2) and for constructing our test statistics.

Model (1.3) can be written in matrix notation as

$$Y = \imath_{nT}\delta + X\beta + D_0\pi + V, \tag{3.1}$$

where $Y \equiv (Y_{11\ldots11}, \ldots, Y_{N_1N_2\ldots N_lT})^\top$ and $V \equiv (V_{11\ldots11}, \ldots, V_{N_1N_2\ldots N_lT})^\top$ are $\mathbb{N}T\times 1$ vectors of the dependent variable and disturbance terms, respectively, X is a $\mathbb{N}T \times d$ matrix of covariates, and β is the $d\times 1$ vector of slope parameters. π is the composite fixed effects parameters and D_0 the corresponding matrix of dummy variables that would have a different specification depending on the model to be considered.[2]

[2]The specific form of D_0 for the models (2.1)–(2.5) is summarized in the following table.

Model	D
(2.1)	$((I_{N_1}\otimes\imath_{N_2}\otimes\imath_T),(\imath_{N_1}\otimes I_{N_2}\otimes\imath_T),(\imath_{N_1}\otimes\imath_{N_2}\otimes I_T))$
(2.2)	$((\imath_{N_1}\otimes\imath_{N_2}\otimes I_T),(I_{N_1}\otimes I_{N_2}\otimes\imath_T))$
(2.3)	$((I_{N_1}\otimes\imath_{N_2}\otimes I_T),(\imath_{N_1}\otimes I_{N_2}\otimes I_T))$
(2.4)	$((I_{N_1}\otimes I_{N_2}\otimes\imath_T),(I_{N_1}\otimes\imath_{N_2}\otimes I_T),(\imath_{N_1}\otimes I_{N_2}\otimes I_T))$
(2.5)	$((\imath_{N_1}\otimes I_{N_2}\otimes\imath_{N_3}\imath_{N_4}\otimes I_T),(\imath_{N_1}\otimes I_{N_2}\otimes I_{N_3}\imath_{N_4}\otimes I_T),(\imath_{N_1}\otimes I_{N_2}\otimes I_{N_3}\otimes I_{N_4}\otimes I_T))$

As it is noted in Balazsi et al. (2015), when the unobserved heterogeneity is correlated with the regressors, the ordinary least-squares estimator is biased and inconsistent. To avoid this deficiency, an alternative approach is to seek an appropriate transformation to control for the unobserved effects. Different from two dimensional panel data models, in multidimensional settings, one must carefully obtain the optimal projection matrix, M_{D_0}, such that $M_{D_0}D_0\pi = 0$.

For the regression model (2.1), there are several combinations that enable us to eliminate the fixed effects, but only the optimal transformation leads to the most efficient estimator. Aiming to obtain a general optimal projection matrix, Wansbeek (1991) shows that the column space D_0 does not change by replacing I_{N_r} (or similarly I_T) with any $(G_{N_l}, \bar{\imath}_{N_l})$ orthonormal matrix of order $N_r \times (N_r - 1)$, where G_{N_r} satisfies the following conditions

$$G_{N_r}^\top \imath_{N_r} = 0, \quad \text{and} \quad G_{N_r}^\top G_{N_r} = I_{N_r-1}, \quad \text{with} \quad \bar{\imath}_{N_r} = \imath_{N_r}/\sqrt{N_r}.$$

As the matrix D_0 spans the same vector space as the orthonormal matrix, the optimal projection matrix of size $(\mathbb{N}T \times \mathbb{N}T)$ to eliminate D_0 is simply $M_{D_0} = I_n - \widetilde{D}_0\widetilde{D}_0^\top$ (see Balazsi et al. (2015) for more detail), where the specific form of $\widetilde{D}_0$ will differ depending on the unobserved effects of the model to be considered.

In Table 1, we collect different combinations of the building blocks for the three-dimensional panel data models in (2.1)–(2.4),[3] where we define $\overline{J}_{N_r} = \bar{\imath}_{N_r}\bar{\imath}_{N_r}^\top/N_r$, $\overline{J}_T = \bar{\imath}_T\bar{\imath}_T^\top/T$, $Q_{N_r} = I_{N_r} - \overline{J}_{N_r}$, and $Q_T = I_T - \overline{J}_T$. Each column of this table corresponds to one particular model and the signs indicate the way each of the building blocks has to be used to construct the appropriate $\widetilde{D}_0\widetilde{D}_0^\top$.

The removal of the unknown fixed effects is accomplished by pre-multiplying both sides of (3.1) by the corresponding projection matrix M_{D_0}. This yields

$$M_{D_0}Y = M_{D_0}X\beta + M_{D_0}V \tag{3.2}$$

as $M_{D_0}\imath_{\mathbb{N}T} = 0$ and $M_{D_0}D_0 = 0$. Applying OLS to model (3.2) gives the Least Squares Dummy Variable (LSDV) estimators for β and δ,

$$\widehat{\beta} = (X^\top M_{D_0}X)^{-1}X^\top M_{D_0}Y, \tag{3.3}$$

$$\widehat{\delta} = \overline{Y} - \overline{X}\widehat{\beta}, \tag{3.4}$$

[3]The corresponding building blocks for the five dimensional panel model in (2.5) are

$$\begin{aligned}\widetilde{D}_0\widetilde{D}_0^\top &= \overline{J}_{N_1} \otimes Q_{N_2} \otimes \overline{J}_{N_3} \otimes \overline{J}_{N_4} \otimes Q_T + \overline{J}_{N_1} \otimes Q_{N_2} \otimes Q_{N_3} \otimes \overline{J}_{N_4} \otimes Q_T \\ &\quad + \overline{J}_{N_1} \otimes Q_{N_2} \otimes Q_{N_3} \otimes Q_{N_4} \otimes Q_T + \overline{J}_{N_1} \otimes \overline{J}_{N_2} \otimes \overline{J}_{N_3} \otimes \overline{J}_{N_4} \otimes \overline{J}_T.\end{aligned}$$

Table 1. Building blocks in projection matrices for three-dimensional panel data models.

$(I_{\mathbb{N}T} - M_{D_0}) = \widetilde{D}_0\widetilde{D}_0^\top$				
Model	(2.1)	(2.2)	(2.3)	(2.4)
$Q_{N_1} \otimes Q_{N_2} \otimes Q_T$				
$Q_{N_1} \otimes Q_{N_2} \otimes \overline{J}_T$		+		+
$Q_{N_1} \otimes \overline{J}_{N_2} \otimes Q_T$				+
$\overline{J}_{N_1} \otimes Q_{N_2} \otimes Q_T$				+
$Q_{N_1} \otimes \overline{J}_{N_2} \otimes \overline{J}_T$	+	+		+
$\overline{J}_{N_1} \otimes Q_{N_2} \otimes \overline{J}_T$	+	+		+
$\overline{J}_{N_1} \otimes \overline{J}_{N_2} \otimes Q_T$	+	+		+
$\overline{J}_{N_1} \otimes \overline{J}_{N_2} \otimes J_T$	+	+		+
$Q_{N_1} \otimes \overline{J}_{N_2} \otimes I_T$			+	
$\overline{J}_{N_1} \otimes Q_{N_2} \otimes I_T$			+	
$\overline{J}_{N_1} \otimes \overline{J}_{N_2} \otimes I_T$			+	

where $\overline{Y} = (\mathbb{N}T)^{-1}\imath_{\mathbb{N}T}^\top Y$ and $\overline{X} = (\mathbb{N}T)^{-1}\imath_{\mathbb{N}T}^\top X$. Note that these LSDV estimators are actually numerically the same as the traditional fixed effects estimators for β and δ since the projection matrix proposed here leads to within-group transformations.[4]

We mention here that this estimation procedure is also valid for the case of incomplete data. However, it is necessary to be careful given that D_0 cannot be represented compactly with Kronecker products, so M_{D_0} cannot be analytically defined element-wise in general. To overcome this, it is possible to use iterative solutions to find the least-squares estimators as Guimaraes and Portugal (2010) and Carneiro et al. (2012) propose.

3.2. **Varying-coefficient panel data model.** There is a relatively large literature in semi- and nonparametric panel data models (see Henderson and Parmeter (2015), Parmeter and Racine (2019), and/or Rodriguez-Poo and Soberon (2017) and the references within, among others). However, nearly all of the previous work focuses on models with solely individual effects and the proposed techniques are not straightforwardly extended to multidimensional

[4]For model (2.1) it leads to the following within transformation

$$\widetilde{y}_{ijt} = y_{ijt} - \frac{1}{N_2}\sum_{j=1}^{N_2} y_{ijt} - \frac{1}{N_1 T}\sum_{i=1}^{N_1}\sum_{t=1}^{T} y_{ijt} - \frac{1}{N_1 N_2}\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} y_{ijt} + \frac{3}{nT}\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\sum_{t=1}^{T} y_{ijt}.$$

nonparametric panel data models. To the best of our knowledge, Henderson et al. (2021) are the first to consider nonparametric estimation of multidimensional models with fixed effects. However, their focus is on the gradient of the conditional mean (they propose to recover the unknown functions under some normalization conditions). This limits application of their method to those settings where the gradient of the conditional mean is of interest.

In order to develop a direct estimation procedure for the unknown function of a varying coefficient model, as in Equation (1.2), we extend the profile least-squares procedure of Su and Ullah (2006) and Gao and Kunpeng (2013) for a one-way error component model to the multidimensional case. In this way, we estimate $\beta(\cdot)$ as a function of the unknown fixed effects parameters. However, as Su and Ullah (2006), Gao and Kunpeng (2013), and Sun et al. (2009) note, we cannot directly estimate Equation (1.2) due to the existence of the fixed effects parameters. To overcome this, we impose $\imath^\top \pi = 0$ in the model.

Rewriting model (1.2) in matrix form, we have the following regression model:

$$Y = B\{X, \beta(Z)\} + D\pi + V, \tag{3.5}$$

where $B\{X, \beta(Z)\}$ stacks all $X_{ij\dots lt}^\top \beta(Z_{ij\dots lt})$ into a $\mathbb{N}T \times 1$ vector with the $(i, j, \dots, l, t)$ subscript matching that of the $(\mathbb{N}T \times 1)$ vector of Y, and D is a $\mathbb{N}T \times \mathbb{N}T$ dummy matrix that satisfies the identification condition.[5]

Then, for any given value of π, we estimate the unknown function $\beta(z)$, where z is an interior point in the support of Z, by minimizing

$$\min_{\beta \in \mathbb{R}^d} [Y - X\beta(z) - D\pi]^\top K_{H_z}(z)[Y - X\beta(z) - D\pi], \tag{3.6}$$

where $K_{H_z}(z) = diag\{K_{H_z}(Z_{11\dots 11}, z), \dots, K_{H_z}(Z_{N_1 N_2 \dots N_l T}, z)\}$ is a $\mathbb{N}T \times \mathbb{N}T$ diagonal matrix, with H_z a $q \times q$ symmetric positive definite bandwidth matrix and $K(\cdot)$ the product of q-univariate kernels, $k(\cdot)$: $K_{H_z}(Z_{ij\dots lt}, z) = \prod_{\ell=1}^q |H_z|^{-1} k(H_z^{-1}(Z_{ij\dots lt,\ell} - z_\ell))$.

Taking the first-order condition of the objective function with respect to $\beta(z)$ gives

$$X^\top K_{H_z}(z)[Y - X\check{\beta}(z) - D\pi] = 0.$$

[5]The structure of the matrix D depends on the form of the unobserved effects in the regression model. For example, for model (2.1) this is

$$\begin{aligned} D \;=\; & [([-\imath_{(N_1-1)} I_{N_1-1}]^\top \otimes \imath_{N_2} \otimes \dots \otimes \imath_{N_l} \otimes \imath_T), (\imath_{N_1} \otimes [-\imath_{(N_2-1)} I_{N_2-1}]^\top \otimes \dots \otimes \imath_{N_l} \otimes \imath_T), \dots, \\ & (\imath_{N_1} \otimes \imath_{N_2}^\top \otimes \dots \otimes [-\imath_{(N_l-1)} I_{N_l-1}] \otimes \imath_T), \dots, (\imath_{N_1} \otimes \imath_{N_2}^\top \otimes \dots \otimes \imath_{N_l} \otimes [-\imath_{(T-1)} I_{T-1}]), \\ & ([-\imath_{(N_1-1)} I_{N_1-1}]^\top \otimes [-\imath_{(N_2-1)} I_{N_2-1}]^\top \otimes \dots \otimes \imath_{N_l} \otimes \imath_T), \dots, \\ & (\imath_{N_1} \otimes \imath_{N_2}^\top \otimes \dots \otimes [-\imath_{(N_l-1)} I_{N_l-1}]^\top \otimes [-\imath_{(T-1)} I_{T-1}]^\top)]. \end{aligned}$$

Rearranging terms in the above equation and solving for $\breve{\beta}(z)$ yields

$$\breve{\beta}(z; H_z) = (X^\top K_{H_z}(z)X)^{-1}X^\top K_{H_z}(z)(Y - D\pi). \tag{3.7}$$

Unfortunately this estimator is infeasible given that π is unobservable. We propose to estimate the fixed effects vector, π, by substituting the estimated nonparametric function from Equation (3.7) into the objective function

$$(Y - D\pi)^\top P(Y - D\pi), \tag{3.8}$$

where $P = (I_{\mathbb{N}T} - S)^\top (I_{\mathbb{N}T} - S)$ is a $\mathbb{N}T \times \mathbb{N}T$ matrix with $S = \left[\{s_H(z_{11\ldots11})\}^\top, \ldots, \{s_H(z_{N_1N_2\ldots N_lT})\}^\top\right]^\top$ being an $\mathbb{N}T \times nt$ matrix whose $(ij\ldots lt)$-th argument is a $1 \times \mathbb{N}T$ vector of the form

$$s_{H_z}(z_{ij\ldots lt}) = X^\top_{ij\ldots lt}(X^\top K_{H_z}(z_{ij\ldots lt})X)^{-1}X^\top K_{H_z}(z_{ij\ldots lt}).$$

Let $(D^\top PD)$ be a nonsingular matrix. From Equation (3.8) we can conclude that the feasible estimator for π is

$$\widehat{\pi} = (D^\top PD)^{-1}D^\top PY. \tag{3.9}$$

Replacing π in (3.7) with (3.9) and assuming that $X^\top K_{H_z}(z)X$ is nonsingular, the feasible local-constant least-squares estimator for $\beta(\cdot)$ is

$$\widehat{\beta}(z; H_z) = g_{H_z}(z)^\top MY \equiv (X^\top K_{H_z}(z)X)^{-1}X^\top K_{H_z}(z)MY, \tag{3.10}$$

where $M = I_{\mathbb{N}T} - D(D^\top PD)^{-1}D^\top P$ and $g_{H_z}(z)^\top = (X^\top K_{H_z}(z)X)^{-1}X^\top K_{H_z}(z)$.

3.3. **Asymptotic properties.** Let $S_z \subset I\!R^q$ and $S_x \subset I\!R^d$ be the support of $Z_{ij\ldots lt}$ and $X_{ijt\ldots lt}$, respectively. In order to establish the asymptotic properties of $\widehat{\beta}(z)$, the following assumptions are necessary.

Assumption 3.1. *The continuous random variables $(X_{ij\ldots lt}, Z_{ij\ldots lt})$ are independently and identically distributed (i.i.d.) across the $(i, j, \ldots, l)$ index for each fixed t, and strictly stationary over t for each fixed $(i, j, \ldots, l)$.*

Assumption 3.2. *The unobserved fixed effects $\pi_{ij\ldots lt}$ are i.i.d. random variables self-uncorrelated with zero mean and finite variances. All the unobserved fixed effects can be correlated with the covariates.*

Assumption 3.3. *The idiosyncratic errors $v_{ij\ldots lt}$ are i.i.d. across $(i, j\ldots, l)$ for fixed t, and strictly stationary over t for each fixed $(i, j, \ldots, l)$, with zero mean and homoscedastic variance, σ^2_v. They are also independent of $X_{ij\ldots lt}$ and $Z_{ij\ldots lt}$. $\pi_{ij\ldots lt}$ and $v_{ij\ldots lt}$ are mutually uncorrelated for all $(i, j, \ldots, l)$, and t.*

Assumption 3.4. *Let $f_{Z_{ij\dots lt}}(\cdot)$ denote the density function of $Z_{ij\dots lt}$ and $f_{Z_{ij\dots lt},Z_{ij\dots lt'}}(\cdot,\cdot)$ the joint density function of $(Z_{ij\dots lt}, Z_{ij\dots lt'})$ for $t \neq t'$. Then, $0 < f_{Z_{ij\dots lt}}(\cdot) < \infty$ at any interior point $z \in S_z$. All density functions and unknown functions are uniformly bounded in the domain of Z and are all twice continuously differentiable in the neighborhood of $z \in S_z$.*

Assumption 3.5. *Let $\|A\| = \sqrt{tr(A^\top A)}$, then $E[\|X_{ij\dots lt}X_{ij\dots lt}^\top\|^2 | Z_{ij\dots lt} = z]$ is bounded and uniformly continuous in its support. The matrix function $E[X_{ij\dots lt}X_{ij\dots lt}^\top | Z_{ij\dots lt} = z]$ is positive definite, bounded and uniformly continuous at any interior point of z in the support of $f_{Z_{ij\dots lt}}(\cdot)$.*

Assumption 3.6. *$K(v) = \prod_{\ell=1}^q k(v_\ell)$ is a product kernel, and the univariate kernel function $k(\cdot)$ is a uniformly bounded, symmetric (around zero) probability density function with compact support $[-1, 1]$. $\int k(v)dv = 1$, $\int vv^\top k(v)dv = \mu_2(K)I_q$, and $\int k^2(v)dv = R(K)$, where $\mu_2(K) \neq 0$ and $R(K) \neq 0$ are scalars and I_q is a $q \times q$ identity matrix.*

Assumption 3.7. *Define $|H_z| = h_{z_1} \cdots h_{z_q}$ and $\|H_z\| = \sqrt{\sum_{\ell=1}^q h_{z_\ell}^2}$. As any element in $\mathbb{N}$ tends to infinity, $\|H_z\| \to 0$ and $\mathbb{N}_{\max}|H_z| \to \infty$.*

Assumption 3.8. *For some $\varepsilon > 0$, $E[|X_{ij\dots lt}v_{ij\dots lt}|^{(2+\varepsilon)} | Z_{ij\dots lt} = z]$ is bounded and uniformly continuous on any point of its support.*

The assumptions listed above are regularity assumptions commonly used in the nonparametric estimation literature. Assumptions 3.1–3.3 are rather standard in the panel data literature. Of course, the strict stationarity condition imposed in Assumption 3.1 can be relaxed to allow other settings of time-dependence such as strong mixing conditions or nonstationary time series and some kind of dependence between the subscripts $(i, j, \dots, l)$ can also be considered. Further, Assumption 3.3 can be relaxed to allow for heteroskedasticity. However, in this paper we consider the limiting behavior of the proposed estimators when any element in $\mathbb{N}$ tends to infinity and T is fixed, so the setup regulated by Assumptions 3.1–3.3 is sufficient. Assumptions 3.4–3.5 are smoothness and boundedness conditions on the density function and moment functionals. For the sake of simplicity in the proofs, in Assumption 3.6, it is assumed that the kernel function takes values over a closed interval. However, that condition is not essential and can be relaxed to allow a general second-order kernel function (such as a Gaussian kernel). Assumption 3.7 is quite standard in the nonparametric literature, and Assumption 3.8 is required to show that the Lyapounov conditions hold for the Central Limit Theorem (CLT).

Before presenting the limiting distribution of $\widehat{\beta}(z)$, let us introduce the following notation. Denote $D_{\beta_\kappa}(\cdot)$ the first-order derivative vector of the κth component of $\beta(\cdot)$, $\mathcal{H}_{\beta_\kappa}(z)$

the Hessian matrix of the κth component of $\beta(\cdot)$, and $D_f(z)$ the first-order derivative vector of the density function, for $\kappa = 1,\ldots,d$. Then, $diag_d\left(tr\{H_z^\top H_z D_f(z) D_{\beta_\kappa}(z)\}\right)$ and $diag_d(tr\{H_z^\top H_z \mathcal{H}_{\beta_\kappa}(z)\})$ stand for a diagonal matrix of elements $tr\{H_z^\top H_z D_f(z) D_{\beta_\kappa}(z)\}$ and $tr\{H_z^\top H_z \mathcal{H}_{\beta_\kappa}(z)\}$, respectively.

Theorem 3.1. *Suppose that Assumptions 3.1–3.8 hold. As* $\mathbb{N}_{\max} \to \infty$ *and* T *is fixed,* $\sqrt{\mathbb{N}_{\max}|H_z|}\|H_z\|^2 = O(1)$ *and we have*

$$\sqrt{\mathbb{N}_{\max}|H_z|}\left(\widehat{\beta}(z;H_z) - \beta(z) - B_{H_z}(z)\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{\sigma_v^2 R^q(K)}{T}\mathcal{B}_{XX}^{-1}(z)\right),$$

where

$$B_H(z) \;= \mu_2^q(K) diag_d\left(tr\{H_z^\top H_z D_f(z) D_{\beta_\kappa}(z)\}\right)\imath_d f_{Z_{11t}}(z)^{-1} + \tfrac{\mu_2^q(K)}{2} diag_d\left(tr\{H_z^\top H_z \mathcal{H}_{\beta_\kappa}(z)\}\right)\imath_d,$$

with $\mathcal{B}_{XX}(z) = E[X_{11\ldots11}X_{11\ldots11}^\top | Z_{11\ldots11} = z] f_{Z_{11\ldots11}}(z)$.

The result in Theorem 3.1 indicates that the estimator is consistent with the common nonparametric rate of convergence. As more curvature in $\beta(\cdot)$ exists, the bias is enlarged. The variance will be penalized when either H_z is large or the data near z are sparse. In contrast to what Lin et al. (2014) obtain in a two-dimensional panel data model, the estimation procedure proposed in this paper ensures that the fixed effects are completely removed even in small sample applications. In addition, we do not have to pay attention to how many indices grow. For the pooled least-squares estimator Sun et al. (2017) propose (for a fully nonparametric panel data model with random effects), this is a critical issue in order to avoid an inconsistent estimator with a slow rate of convergence. Our result in Theorem 3.1 is valid independent of how many indices are large.

4. A General CLT for U-Statistics

In this section, we propose a general CLT for degenerate U-statistics that will be useful for the construction and analysis of a variety of specification tests in multidimensional panel and longitudinal data models. This new CLT is necessary given the potential growth of various indices to ∞ relative to the standard setting of a single index growing to ∞.

Let U_{N_r} be an r^{th}-sample U-statistic of the form

$$U_{N_r} = \prod_{r=1}^{R}\binom{N_r}{2}^{-1}\sum_{i=1}^{N_1}\sum_{i'\neq i}^{N_1}\sum_{j=1}^{N_2}\sum_{j'\neq j}^{N_2}\cdots\sum_{l=1}^{N_l}\sum_{l'\neq l}^{N_l}\mathcal{H}_{\mathbb{N}}(\chi_{ij\ldots l}\chi_{i'j'\ldots l'}) \tag{4.1}$$

where R is the total number of individual subscripts, $\chi_{ij\ldots l}$ is an $i.i.d.$ random sample in $i,j,\ldots,l$ and $\mathcal{H}_{\mathbb{N}}(\cdot,\cdot)$ is a symmetric function in its arguments, i.e., $\mathcal{H}_{\mathbb{N}}(\chi_{ij\ldots l},\chi_{i'j'\ldots l'}) = \mathcal{H}_{\mathbb{N}}(\chi_{i'j'\ldots l'},\chi_{ij\ldots l})$. If $\mathcal{H}_{\mathbb{N}}(\chi_{ij\ldots l},\chi_{i'j'\ldots l'})$ has zero mean and $E[\mathcal{H}_{\mathbb{N}}^2(\chi_{ij\ldots l},\chi_{i'j'\ldots l'})] < \infty$, then, if a single element in $\mathbb{N}$ tends to infinity, we can resort to Hall (1984)'s CLT for degenerate U-statistics in order to show that, after proper normalization, U_{n_r} tends to a Standard Normal distribution. However, if more than one element in $\mathbb{N}$ tends to infinity, a more general CLT is necessary.

In order to handle this situation, we provide a generalization of Theorem 1 in Hall (1984) to the multidimensional setting. The following theorem describes the asymptotic distribution of an r^{th}-sample degenerate U-statistic.

Theorem 4.1. *Assume* $E[\mathcal{H}_{\mathbb{N}}(\chi_{11\ldots11}, \chi_{22\ldots22})] = 0$ *almost surely and* $E[\mathcal{H}^2_{\mathbb{N}}(\chi_{11\ldots11}, \chi_{22\ldots22})] < \infty$ *for each* $N_1, N_2, \ldots, N_l$. *If*

$$\frac{E[G^2_{\mathbb{N}}(\chi_{11\ldots11}, \chi_{22\ldots22})] + \mathbb{N}^{-1}_{\max} E[\mathcal{H}^4_{\mathbb{N}}(\chi_{11\ldots11}, \chi_{22\ldots22})]}{\{E[\mathcal{H}^2_{\mathbb{N}}(\chi_{11\ldots11}, \chi_{22\ldots22})]\}^2} \to 0,$$

as $\mathbb{N}_{\max} \to \infty$, *where* $G^2_{\mathbb{N}}(\chi_{11\ldots11}, \chi_{22\ldots22}) = E[\mathcal{H}_{\mathbb{N}}(\chi_{33\ldots33}, \chi_{22\ldots22})|\chi_{11\ldots11}, \chi_{22\ldots22}]$. *Then,*

$$U_{N_r} \quad \xrightarrow{d} \quad N\left(0, \frac{\mathbb{N}^2}{2^r} E[\mathcal{H}^2_{\mathbb{N}}(\chi_{11\ldots11}, \chi_{22\ldots22})]\right).$$

The proof of this theorem is in the Appendix. If $r = 1$, we obtain an identical result to Hall (1984). However, for $r > 1$ (multidimensional panel data models of any dimension of the fixed effects), we have a more general limiting result. While it has been assumed that the r-samples are independent, it would be possible to allow some forms of dependence between some of the samples. However, this beyond the scope of this paper.

5. Applications of the General CLT

We consider several types of specification tests for multidimensional nonparametric panel data models. The first test focuses on functional form assumptions of the regression model. The second test conducts inference based on statistical significance for a set of covariates. The third test is a Hausman test: testing the random effects specification against a fixed effects specification. Similar tests exist (see Li et al. (2002), Henderson et al. (2008), Sun et al. (2009), and Lin et al. (2014), among others). However, existing tests are constructed explicitly in the context of a two-dimensional panel data problem (i and t). To the best of our knowledge, these hypotheses in multidimensional panels do not exist.

5.1. **Test of Correct Parametric Specification.** We consider two different tests for testing the functional form assumptions of a regression. The first will enable us to test the null hypothesis of a multidimensional linear panel data model against a nonparametric alternative (H^a_0). The second compares the parametric specification against a varying coefficient alternative (H^b_0).

5.1.1. *Nonparametric Alternative.* The first testing problem statistically assess if the fully parametric model specified in Equation (1.3) is correct against an unspecified alternative model. The null and alternative hypotheses are:

$$\begin{aligned} H_0^a &: Pr\{\beta(X_{ij\dots lt}, Z_{ij\dots lt}) = \delta_0 + (X_{ij\dots lt}, Z_{ij\dots lt})^\top \beta_0\} = 1 \quad \text{for some} \quad (\delta_0, \beta_0)^\top \in \Theta \subset I\!R^{1+d+q}, \\ H_1^a &: Pr\{\beta(X_{ij\dots lt}, Z_{ij\dots lt}) = \delta + (X_{ij\dots lt}, Z_{ij\dots lt})^\top \beta\} < 1 \quad \text{for any} \quad (\delta, \beta)^\top \in \Theta, \end{aligned} \tag{5.1}$$

where Θ is a compact and convex subset on $I\!R^{1+d+q}$.

In principle, we can construct a consistent test statistic for H_0^a based on the weighted integrated squared distance between the parametric and nonparametric fits of the data, where the weight function is used to avoid the random denominator problem (e.g., see Li et al. (2002)). However, the resulting test has several bias terms which both complicates its asymptotic analysis and impacts the performance in finite samples. To overcome these hurdles, we follow Lin et al. (2014) and smooth the parametric fit instead of using the direct LSDV estimators of (δ_0, β_0). Thus, our test statistic is based on the weighted integrated squared distance between $\beta(x, z)$ and $\beta_p(x, z)$ (i.e., $\int [S_{a,\mathbb{N}}(\beta(x, z) - \beta_p(x, z))]^2 dxdz$), where $\beta(\cdot)$ is the smooth function of a fully nonparametric model, $\beta_p(\cdot)$ is the corresponding smooth function for a parametric fit, and $S_{a,\mathbb{N}}$ is the inverse matrix of the nonparametric estimator.

Since the nonparametric estimator of $\beta(x, z)$ is of the form $\widehat{\beta}(x, z) = S_{a,\mathbb{N}}^{-1} \imath_{\mathbb{N}T}^\top K_H(x, z) MY$, where $S_{a,\mathbb{N}} = \imath_{\mathbb{N}T}^\top K_H(x, z) \imath_{\mathbb{N}T}$ and $H = (H_x, H_z)$, we define a kernel smoothed least-squares estimator of $\{\delta + (X, Z)\beta\}$ given by $\widehat{\beta}_p(x, z) = S_{a,\mathbb{N}}^{-1} \imath_{\mathbb{N}T}^\top K_H(x, z)(\widehat{\delta}\imath_{\mathbb{N}T} + (X, Z)^\top \widehat{\beta})$, where $\widehat{\beta}$ and $\widehat{\delta}$ are the LSDV estimators proposed in Equations (3.3) and (3.4). It can be shown that $\widehat{\beta}(x, z) - \widehat{\beta}_p(x, z) = S_{a,\mathbb{N}}^{-1} \imath_{\mathbb{N}T}^\top K_H(x, z) M\widehat{V}_a$, where $\widehat{V} = Y - \widehat{\delta}\imath_{\mathbb{N}T} - (X, Z)\widehat{\beta}$ are the estimated parametric residuals. This allows us to obtain the test statistic

$$I_{\mathbb{N}}^a = \widehat{V}^\top M^\top \left[\int K_H(x, z) \imath_{\mathbb{N}T} \imath_{\mathbb{N}T}^\top K_H(x, z) dxdz \right] M\widehat{V},$$

where the matrix M is used to remove (asymptotically) the fixed effects term in Equation (3.1) and $K_H(x, z)$ is the product of two kernel functions defined as in Equation (3.6).

The typical element of $\int K_H(x, z) \imath_{\mathbb{N}T} \imath_{\mathbb{N}T}^\top K_H(x, z) dxdz$ can be written as twofold convolution kernels (i.e., $\overline{K}_{H_x}(X_{ij\dots t}, X_{i'j'\dots t'}) \overline{K}_{H_z}(Z_{ij\dots t}, Z_{i'j'\dots t'}) = \int K(H_x^{-1}(X_{ij\dots lt} - X_{i'j'\dots l't'}) + \omega_1) K(\omega_1) K(H_z^{-1}(Z_{ij\dots lt} - Z_{i'j'\dots l't'}) + \omega_2) K(\omega_2) d\omega_1 d\omega_2$) which preserve the local weighting property of the test statistic. Following Li et al. (2002) and Lin et al. (2014), we simplify the test replacing $\overline{K}_{H_x}(X_{ij\dots lt}, X_{i'j'\dots l't'})$ and $\overline{K}_{H_z}(Z_{ij\dots lt}, Z_{i'j'\dots l't'})$ with $K_{H_x}(X_{ij\dots lt}, X_{i'j'\dots l't'})$ and

$K_{H_z}(Z_{ij\ldots lt}, Z_{i'j'\ldots l't'})$, respectively. The resulting test statistic is

$$\widehat{I}^a_{\mathbb{N}} \;=\; \frac{1}{\mathbb{N}^2T^2}\sum_{ij\ldots lt}\sum_{i'j'\ldots l't'} \widehat{\widetilde{v}}_{ij\ldots lt}\widehat{\widetilde{v}}_{i'j'\ldots l't'}K_{H_x}(X_{ij\ldots lt}, X_{i'j'\ldots l't'})K_{H_z}(Z_{ij\ldots lt}, Z_{i'j'\ldots l't'}), \qquad (5.2)$$

where $\widehat{\widetilde{v}}_{ij\ldots lt} = \widetilde{Y}_{ij\ldots lt} - \widetilde{X}^\top_{ij\ldots lt}\widehat{\beta}_1 - \widetilde{Z}^\top_{ij\ldots lt}\widehat{\beta}_2$ are the transformed residuals obtained using the optimal projection matrix M_{D_0}, with $(\widetilde{Y}_{ij\ldots lt}, \widetilde{X}_{ij\ldots lt}, \widetilde{Z}_{ij\ldots lt})$ the corresponding within transformed variables and $(\widehat{\beta}_1, \widehat{\beta}_2)$ the resulting LSDV estimates, while H_x and H_z are symmetric positive bandwidth matrices.

To overcome the problem of asymptotically non-negligible center terms common in this type of double summation test, we use a leave-one-out estimator, which is equivalent to replacing diagonal elements of the kernel matrix with zeros. Our modified test statistic is

$$\begin{aligned}\widehat{I}^a_{\mathbb{N}} \;=\; & \prod_{r=1}^{R}\frac{1}{\mathbb{N}T(N_r-1)(T-1)}\sum_{ij\ldots lt}\;\sum_{(i'j'\ldots l't')\neq(ij\ldots lt)} \widehat{\widetilde{v}}_{ij\ldots lt}\widehat{\widetilde{v}}_{i'j'\ldots l't'}K_{H_x}(X_{ij\ldots lt}, X_{i'j'\ldots l't'}) \\ & \times K_{H_z}(Z_{ij\ldots lt}, Z_{i'j'\ldots l't'}). \qquad (5.3)\end{aligned}$$

Letting $f_{X_{ij\ldots lt},Z_{ij\ldots lt}}(\cdot,\cdot)$ denote the joint p.d.f. of $(X_{ij\ldots lt}, Z_{ij\ldots lt})$, our test statistic is a leave-one-out version of the kernel estimator of $E[\widetilde{v}_{a,ij\ldots lt}E(\widetilde{v}_{a,ij\ldots lt}|X_{ij\ldots lt}, Z_{ij\ldots lt})f(X_{ij\ldots lt}, Z_{ij\ldots lt})]$. In order to derive the asymptotic properties of $\widehat{I}^a_{\mathbb{N}}$, we can employ our CLT for degenerate U-statistics since $\widehat{I}^a_{\mathbb{N}}$ can be rewritten as in Equation (4.1), where $H_{\mathbb{N}}(\chi_{ij\ldots l}, \chi_{i'j'\ldots l'}) = \sum_{t=1}^{T}\sum_{t'\neq t}^{T} \widehat{\widetilde{v}}_{a,ij\ldots lt}\widehat{\widetilde{v}}_{a,i'j'\ldots l't'}K_{H_x}(X_{ij\ldots lt}, X_{i'j'\ldots l't'})K_{H_z}(Z_{ij\ldots lt}, Z_{i'j'\ldots l't'})$. To do so, the following assumptions are required.

Assumption 5.1. *There is an $\iota > 2$ such that $E\|X_{ij\ldots lt}\|^{(1+\iota)} \leq M < \infty$ and $E\|Z_{ij\ldots lt}\|^{(1+\iota)} \leq M < \infty$. For all $(i, j, \ldots, l)$ and $t \neq s$, $E(\beta(X_{ij\ldots lt}, Z_{ij\ldots lt})^{(1+\iota)}|X_{ij\ldots ls}, Z_{ij\ldots ls})$, $E(X^{(1+\iota)}_{ij\ldots lt}|X_{ij\ldots ls}, Z_{ij\ldots ls})$, and $E(Z^{(1+\iota)}_{ij\ldots lt}|X_{ij\ldots ls}, Z_{ij\ldots ls})$ are all uniformly bounded.*

Assumption 5.2. *Let $\widetilde{X}_{ij\ldots lt}$ be the within transformed variable for all $(i, j, \ldots, l)$, and t. $(\mathbb{N}T)^{-1}X^\top M_{D_0}X \xrightarrow{p} E(\widetilde{X}^\top_{ij\ldots l}\widetilde{X}_{ij\ldots l})$ is a finite positive definite matrix.*

Assumption 5.3. *At any interior point $(x, z) \in S$, $0 < f_{X_{ij\ldots lt},Z_{ij\ldots lt}}(\cdot,\cdot) < \infty$ and it is uniformly bounded in the domain of (X, Z).*

The asymptotic behavior of our proposed test statistic under both H^a_0 and H^a_1 is given in the following theorem.

Theorem 5.1. *Assuming $E(v^4_{ij\ldots lt}) = \tau_4 < \infty$ and suppose that Assumptions 3.1–3.5, 3.6–3.7, and 5.1–5.3 hold, as $\mathbb{N}_{\max} \to \infty$ for T fixed, we have* *a) Under* H_0^a,

$$J_{\mathbb{N}}^a = \mathbb{N}|H_x|^{1/2}|H_z|^{1/2}\frac{\widehat{I}_n^a}{\sqrt{\widehat{\Sigma}_a}} \quad \xrightarrow{d} \quad \mathcal{N}(0,1),$$

where

$$\widehat{\Sigma}_a = \prod_{r=1}^{R}\binom{N_r}{2}^{-1}\frac{1}{T^2|H_x||H_z|}\sum_{ij\ldots lt}\sum_{(i'j'\ldots l't')\neq(ij\ldots lt)}\widehat{\widetilde{v}}^2_{ij\ldots lt}\widehat{\widetilde{v}}^2_{i'j'\ldots l't'}K^2(X_{ij\ldots lt}, X_{i'j'\ldots l't}) \times K^2(Z_{ij\ldots lt}, Z_{i'j'\ldots l't'})$$

is a consistent estimator of the asymptotic variance of $\mathbb{N}|H_x|^{1/2}|H_z|^{1/2}\widehat{I}_{\mathbb{N}}^a$ *for*

$$\Sigma_a = \frac{2^{(R+1)}(T-1)\sigma_v^4 R^{(q+d)}(K)}{T^3}E[f(X_{11\ldots 1t}, Z_{11\ldots 1t})].$$

b) Under H_1^a, $Pr(J_{\mathbb{N}}^a \geq C^a) \to 1$ *as* $\mathbb{N}_{\max} \to \infty$ *for* T *fixed, where* C^a *is any positive constant.*

The proof of Theorem 5.1 is presented in the Appendix. Theorem 5.1 does not require T to be large. If T is large, the proposed test statistic works under some mild conditions and only minor modifications are needed. The limiting distribution of this test statistic is valid under two scenarios. If one element in $\mathbb{N}$ tends to infinity, we use Hall (1984)'s CLT in order to prove that $U_{N_1}^a$ tends to a Standard Normal distribution. If more than one element in $\mathbb{N}$ tends to infinity, we resort to the results in Theorem 4.1. Part (b) of Theorem 5.1 shows that, when the H_0^a is false, the probability that the $J_{\mathbb{N}}^a$ test rejects the H_0^a approaches one as $\mathbb{N}_{\max} \to \infty$, so we conclude that $J_{\mathbb{N}}^a$ is a consistent test.

5.1.2. *Varying Coefficient Alternative.* The varying coefficient model nests the parametric model. If the parametric model is correctly specified, it is more efficient. If it is not, we typically will obtained biased and inconsistent results. Here we test whether the parametric model is an adequate description of the data against the varying coefficient specification.

Let $X_{0,ij\ldots lt} = [1, X_{ij\ldots lt}]^\top$ and $\beta_0(Z_{ij\ldots lt}) = [\delta + Z_{ij\ldots lt}^\top\beta_2, \beta_1^\top]^\top$ be $(1+d)\times 1$ vectors. The fully parametric model in Equation (1.3) can be written more compactly as

$$Y_{ij\ldots lt} = X_{0,ij\ldots lt}^\top\beta_0(Z_{ij\ldots lt}) + \pi_{ij\ldots lt} + v_{ij\ldots lt}.$$

With this model in mind, we consider the following null and alternative hypotheses:

$$\begin{aligned} H_0^b &: \beta(z) - \beta_0(z) = 0 \quad \text{almost everywhere,} \\ H_1^b &: \beta(z) - \beta_0(z) \neq 0 \quad \text{on a set with positive measure.} \end{aligned} \tag{5.4}$$

Similar to before, we use a weighted integrated squared difference as the basis of our test statistic (i.e., $I^b = \int[S_{b,\mathbb{N}}(\beta(z) - \beta_0(z))dz$, where $S_{b,\mathbb{N}} = X^\top W_{H_z}(z)X)$ given that, under H_0^b, it can be proved that $I^b = 0$ and, under H_1^b, $I^b > 0$. In order to obtain a feasible statistic, we replace $\beta(z)$ with the local-constant least-squares estimator obtained in Equation (3.10) and replace $\beta_0(z)$ with the kernel smoothed least-squares estimator of $X_0\beta_0(Z)$ of the form $\widehat{\beta}_p(z) = S_{b,\mathbb{N}}^{-1}X^\top K_{H_z}(z)MX_0\widehat{\beta}_0(Z)$, where $\widehat{\beta}_0(Z) = \imath_{\mathbb{N}T}\widehat{\delta} + X\widehat{\beta}_1 + Z\widehat{\beta}_2$. The resulting test statistic is

$$I_{\mathbb{N}}^b = \frac{1}{\mathbb{N}^2T^2}\int \widehat{V}^\top K_{H_z}(z)XX^\top K_{H_z}(z)\widehat{V}dz,$$

given that $S_{b,\mathbb{N}}(z)[\widehat{\beta}(z) - \widehat{\beta}_{b,p}(z)] = X^\top K_{H_z}(z)\widehat{V}$, where $\widehat{V}$ are the corresponding parametric residuals.

The integration calculation can be simplified (due to the kernel weight function $K_{H_z}(Z_{ij\dots lt}-z)K_{H_z}(Z_{i'j'\dots l't'} - z))$ by replacing $\widehat{\beta}(z)$ with $\widehat{\beta}(Z_{ij\dots lt})$ given that only those z close to both $Z_{ij\dots lt}$ and $Z_{i'j'\dots l't'}$ are important. Taking this feature into account and leaving out the diagonal elements in the kernel matrix to remove the nonzero center terms from $I_{\mathbb{N}}^b$ under H_0^b, the resulting test statistic becomes

$$\widehat{I}_{\mathbb{N}}^b = \frac{1}{\mathbb{N}^2T^2}\sum_{ij\dots lt}\sum_{(i'j'\dots l't')\neq(ij\dots lt)} \widehat{\widetilde{v}}_{ij\dots lt}^\top X_{ij\dots lt}^\top X_{i'j'\dots l't'}\widehat{\widetilde{v}}_{i'j'\dots l't'}K_{H_z}(Z_{ij\dots lt}, Z_{i'j'\dots l't'}). \tag{5.5}$$

where $\widehat{\widetilde{v}}_{ij\dots lt}$ are the corresponding within transformed parametric residuals defined in Equation (5.2). Our test statistic can be interpreted as a leave-one-out version of the kernel estimator of $E[\widetilde{v}_{ij\dots lt}X_{ij\dots lt}^\top E(X_{ij\dots lt}\widetilde{v}_{ij\dots lt}|Z_{ij\dots lt})f(Z_{ij\dots lt})]$ and it can be also rewritten as a 'second order' U-statistic as in Equation (3.1) in order to obtain the limiting behavior of $\widehat{I}_{\mathbb{N}}^b$ under H_0^b and H_1^b.

Theorem 5.2. *Assuming $E(u_{ij\dots lt}^4) = \tau_4 < \infty$ and suppose that Assumptions 3.1–3.5, 3.6–3.7, and 5.1–5.2 hold. As $\mathbb{N}_{\max} \to \infty$ for T fixed,*

a) Under H_0^b,

$$J_{\mathbb{N}}^b = \mathbb{N}|H_z|^{1/2}\frac{\widehat{I}_{\mathbb{N}}^b}{\sqrt{\widehat{\Sigma}_b}} \quad \xrightarrow{d} \quad \mathcal{N}(0,1),$$

where

$$\begin{aligned}\widehat{\Sigma}_b \;=\; & \prod_{r=1}^{R}\binom{N_r}{2}^{-1}\binom{T}{2}^{-1}|H_z|^{-1}\sum_{ij\dots lt}\sum_{(i'j'\dots l't')\neq(ij\dots lt)}\widehat{\widetilde{v}}_{ij\dots lt}^2\widehat{\widetilde{v}}_{i'j'\dots l't'}^2X_{ij\dots lt}^\top X_{ij\dots lt}\\ & \times X_{i'j'\dots l't'}^\top X_{i'j'\dots l't'}K^2(Z_{ij\dots lt}, Z_{i'j'\dots l't'})\end{aligned}$$

is a consistent estimator of the asymptotic variance of $\mathbb{N}|H_z|^{1/2}\widehat{I}^b_{\mathbb{N}}$,

$$\Sigma_b = \frac{2^{(R+1)}(T-1)\sigma_v^4 R^q(K)}{T^3} E[E[tr\{X_{ij\ldots lt}X^\top_{i'j'\ldots l't'}\}^2|Z_{ij\ldots lt}]f(Z_{ij\ldots lt})].$$

b) *Under* H_1^b, $Pr(J^b_{\mathbb{N}} \geq C^b) \to 1$ *as* $\mathbb{N}_{\max} \to \infty$ *for* T *fixed, where* C^b *is any positive constant.*

The proof of this theorem is in the Appendix. In Theorems 5.1 and 5.2, $I^a_{\mathbb{N}}$ and $I^b_{\mathbb{N}}$ have asymptotic standard Normal distributions under H_0^a and H_0^b, respectively. However, it is well-known that kernel-based nonparametric tests do not work well for small or moderate samples (see Härdle and Mammen (1993), Li and Wang (1998), and Whang and Andrews (1993), among others). We suggest using a bootstrap procedure to better approximate the finite sample null distribution of these test statistics.

Following Henderson et al. (2008), Kapetanios (2008) and Gonçalves (2011), we propose a bootstrap procedure which consists of resampling with replacement the cross-sectional units as a whole rather than resampling within the units across time. Kapetanios (2008) shows that this resampling scheme enables superior approximation to distributions of statistics over the latter for panel data sets with large cross-sectional units and fixed or few time periods. To describe our bootstrap routine, let $\widehat{\widetilde{v}}_{ij\ldots l}$ be a $T \times 1$ vector where $\widehat{\widetilde{v}}_{ij\ldots lt} = \widetilde{Y}_{ij\ldots lt} - \widetilde{X}^\top_{ij\ldots lt}\widehat{\beta}_1 - \widetilde{Z}^\top_{ij\ldots lt}\widehat{\beta}_2$ are the within transformed residuals from a parametric regression. Using the two point wild bootstrap, our procedure is as follows:

(1) Set $\widetilde{v}^*_{ij\ldots l} = \left(\frac{1-\sqrt{5}}{2}\right)\widehat{\widetilde{v}}_{ij\ldots l}$ with probability ρ and $\widetilde{v}^*_{ij\ldots l} = \left(\frac{1+\sqrt{5}}{2}\right)\widehat{\widetilde{v}}_{ij\ldots l}$ with probability $(1-\rho)$, where $\rho = (1+\sqrt{5})/2\sqrt{5}$. Resample the entire set of temporal vector residuals for a particular cross-sectional unit ($t = 1, \ldots, T$), where $\widetilde{v}^*_{ij\ldots l}$ is obtained from a random draw from $\{\widehat{\widetilde{v}}_{ij\ldots l}\}^{N_1,N_2,\ldots,N_l}_{i,j,\ldots,l}$ with replacement.

(2) Calculate the bootstrap dependent variables as $\widetilde{Y}^*_{ij\ldots lt} = \widetilde{X}^\top_{ij\ldots lt}\widehat{\beta}_1 + \widetilde{Z}^\top_{ij\ldots lt}\widehat{\beta}_2 + \widetilde{v}^*_{ij\ldots lt}$ and use the bootstrap sample, $\{\widetilde{X}_{ij\ldots lt}, \widetilde{Z}_{ij\ldots lt}, \widetilde{Y}^*_{ij\ldots lt}\}$, to calculate the parametric LSDV estimators for β_1 and β_2, where $(\widehat{\beta}^*_1, \widehat{\beta}^*_2)$ are the LSDV estimators with $\widetilde{Y}^*_{ij\ldots lt}$ instead of $\widetilde{Y}_{ij\ldots lt}$. Calculate the bootstrap residuals as $\widehat{\widetilde{v}}^*_{ij\ldots lt} = \widetilde{Y}^*_{ij\ldots lt} - \widetilde{X}^\top_{ij\ldots lt}\widehat{\beta}^*_1 - \widetilde{Z}^\top_{ij\ldots lt}\widehat{\beta}^*_2$.

(3) Compute $J^{a*}_{\mathbb{N}}$ in the same way as $J^a_{\mathbb{N}}$ replacing $\widehat{\widetilde{v}}_{ij\ldots lt}$ with $\widehat{\widetilde{v}}^*_{ij\ldots lt}$.

(4) Repeat steps 1-4 a large number B of times and obtain the empirical distribution of $J^{a*}_{\mathbb{N}}$ to approximate the null distribution of $J^a_{\mathbb{N}}$.

Let $J^{a*}_{\mathbb{N},\alpha}$ be the α-percentile of the bootstrap distribution from step 4. Reject the null hypothesis at significance level α if $J^{a}_{\mathbb{N}} > J^{a*}_{\mathbb{N},\alpha}$. By the same reasoning as above, we can use

$$\widetilde{\Sigma}^{*}_{a} = \prod_{r=1}^{R} \binom{N_r}{2}^{-1} \binom{T}{2}^{-1} |H_z|^{-1}|H_x|^{-1} \sum_{ij\ldots lt} \sum_{(i'j'\ldots l't')\neq(ij\ldots lt)} \widehat{\widetilde{v}}^{*2}_{ij\ldots lt} \widehat{\widetilde{v}}^{*2}_{i'j'\ldots l't'} K^2(X_{ij\ldots lt}, X_{i'j'\ldots l't'}) \times K^2(Z_{ij\ldots lt}, Z_{i'j'\ldots l't'})$$

to replace $\widehat{\Sigma}^{*}_{a}$ in step 4 above.

Note that in order to show that the bootstrap method is an asymptotically valid procedure to approximate the distribution of the test statistic under H^a_0 regardless if H^a_0 holds true or not, the following assumption is required.

Assumption 5.4. *$E(\pi^{\iota}_{ij\ldots lt}|Z_{ij\ldots lt} = z)$, for $\iota = 1, \ldots, 4$, are continuous differentiable and uniformly bounded over $z \in S$.*

With the addition of this assumption, we are able to give the following result:

Theorem 5.3. *Under Assumptions 3.1–3.5, 3.6–3.7, and 5.4, as $\mathbb{N}_{\max} \to \infty$ for T fixed,*

$$\sup_{z\in I\!R} |Pr^*(J^{a*}_{\mathbb{N}} \leq z) - \Phi(z)| = o_p(1),$$

where $Pr^*(\cdot) = Pr(\cdot|\{(X_{ij\ldots lt}, Z_{ij\ldots lt}, Y_{ij\ldots lt})\}^{N_1,N_2,\ldots,N_l}_{i,j,\ldots,l})$, and $\Phi(\cdot)$ is the standard Normal cumulative distribution function.

The proof of this theorem can be accomplished by extending the proof of Theorem 3.3 in Lin et al. (2014) to the multidimensional case. For the sake of brevity, it has been omitted from the paper. A similar procedure is proposed for $J^b_{\mathbb{N}}$ replacing steps 3 and 4 with the corresponding $J^b_{\mathbb{N}}$ test statistic.

5.2. **A Test of Statistical Significance.** Removing irrelevant covariates can provide more desirable statistical properties of the proposed estimator. In a nonparametric setting, this is specially important as the curse of dimensionality is an overarching concern. With the aim of testing irrelevant covariates, we could be tempted to use the test statistic proposed in Equation (3.7). However, this test is not valid for all varying coefficient specifications given that it is unclear what happens when all Z are irrelevant. To overcome this, we propose an alternative test statistic which combines the conditional moment tests with the nonparametric estimator.

Assume we observe data from $(N_1, N_2, \ldots, N_l)$ individuals and/or groups across T periods, each with a single response variable, Y, and q-covariates, Z, that can be split up into Z_1 and Z_2. Considering a varying coefficient panel data model as in Equation (1.2) and assume that

Z_1 is a $q_1 \times 1$ vector of relevant covariates, whereas Z_2 is a $q_2 \times 1$ vector of potentially irrelevant covariates. In order to assess variable significance, we are interested in checking whether or not the residuals have a conditional mean equal to zero when the assumed irrelevant regressors are not included in the model. The null and alternative hypotheses for variable relevance are

$$\begin{aligned} H_0^c: &\quad X_{ij\ldots lt}^\top \beta(Z_{1ij\ldots lt}, Z_{2ij\ldots lt}) = X_{ij\ldots lt}^\top \beta(Z_{1ij\ldots lt}) \\ H_1^c: &\quad X_{ij\ldots lt}^\top \beta(Z_{1ij\ldots lt}, Z_{2ij\ldots lt}) \neq X_{ij\ldots lt}^\top \beta(Z_{1ij\ldots lt}), \end{aligned} \tag{5.6}$$

where we emphasize that this test is more computationally intensive than that proposed for correct parametric specification as it relies on nonparametric residuals, i.e., $\widehat{v}(z)$. However, it still possesses desirable properties of the previous tests.

We can construct a consistent test for H_0^c based on the integrated squared difference of the two nonparametric estimators (i.e., $\beta(z_1, z_2)$ and $\beta(z_1)$), which is asymptotically zero under H_0^c and positive under H_1^c. To remove the random denominator of $\widehat{\beta}(z_1, z_2)$, we pre-multiply this test by $X^\top K_{H_z}(z_1, z_2)X$, where H_z contains the bandwidth matrices H_{z_1} and H_{z_2} defined in Equation (3.6). The proposed test statistic will be based on

$$\begin{aligned} I^c &= \int [\widehat{\beta}(z_1, z_2) - \widehat{\beta}(z_1)]^\top (X^\top K_{H_z}(z_1, z_2)X)^\top X^\top K_{H_z}(z_1, z_2)X[\widehat{\beta}(z_1, z_2) - \widehat{\beta}(z_1)]dz_1 dz_2 \\ &= \int \widehat{v}(z_1)^\top K_{H_z}(z_1, z_2)XX^\top K_{H_z}(z_1, z_2)\widehat{v}(z_1)dz_1 dz_2, \end{aligned} \tag{5.7}$$

as in Section 5.1, it is easy to show that

$$X^\top K_{H_z}(z_1, z_2)[\widehat{\beta}(Z_1, Z_2) - \widehat{\beta}(Z_1)] \equiv X^\top K_{H_z}(z_1, z_2)[MY - B\{X\widehat{\beta}(Z_1)\}] \equiv X^\top K_{H_z}(z_1, z_2)\widehat{v}(z_1),$$

where $\widehat{\beta}(z_1)$ is the nonparametric estimator obtained in (3.10) and the corresponding local-constant least-squares estimator for $\beta(z_1, z_2)$, where z_1 and z_2 are interior points in the support of $Z_{1ij\ldots lt}$ and $Z_{2ij\ldots lt}$, respectively, is

$$\widehat{\beta}(z_1, z_2; H_z) = (X^\top K_{H_z}(z_1, z_2)X)^{-1} X^\top K_{H_z}(z_1, z_2)MY, \tag{5.8}$$

where $K_{H_z}(z_1, z_2)$ is the $\mathbb{N}T \times \mathbb{N}T$ matrix

$$K_{H_z}(z_1, z_2) = diag\{K_{H_z}(Z_{1,11\ldots 11} - z_1)K_{H_z}(Z_{2,11\ldots 11} - z_2), \ldots, K_{H_z}(Z_{1,N_1 N_2 \ldots N_l T} - z_1)K_{H_z}(Z_{2,N_1 N_2 \ldots N_l T} - z_2)\}.$$

The resulting feasible test statistic (removing the non-zero center term) is given by

$$\begin{aligned} \widehat{I}_{\mathbb{N}}^c &= \prod_{r=1}^{R} \frac{1}{\mathbb{N}T(N_r - 1)(T-1)} \sum_{ij\ldots lt} \sum_{(i'j'\ldots l't') \neq (ij\ldots lt)} \widehat{\widehat{v}}_{ij\ldots lt}(z_1) \widehat{\widehat{v}}_{i'j'\ldots l't'}(z_1) X_{ij\ldots lt}^\top X_{i'j'\ldots l't'} \\ &\quad \times K_{H_z}(Z_{ij\ldots lt}, Z_{i'j'\ldots l't'}), \end{aligned} \tag{5.9}$$

where $Z_{ij\ldots lt} = (Z_{1ij\ldots lt}, Z_{2ij\ldots lt})$ and $\widehat{\widetilde{v}}_{ij\ldots lt}(z_1)$ is the transformed nonparametric within residual. This test statistic can be considered as a weighted conditional-moment test, $E[\widetilde{v}_{ij\ldots lt}(z_1)X_{ij\ldots lt}^{\top}E(X_{ij\ldots lt}\widetilde{v}_{ij\ldots lt}(z_1)|Z_{1ij\ldots lt}, Z_{2ij\ldots lt})f(Z_{1ij\ldots lt}, Z_{2ij\ldots lt})]$. This test is different from what Henderson and Parmeter (2015) propose for irrelevant regressors.

The following theorem gives the asymptotic distribution of our proposed test statistic $\widehat{I}_{\mathbb{N}}^{c}$:

Theorem 5.4. *Suppose that Assumptions 3.1–3.5 and 3.6 and 3.7 hold. As* $\mathbb{N}_{\max} \to \infty$ *for* T *fixed,* $\sqrt{\mathbb{N}_{\max}|H_z|}\|H_z\|^4 = O(1)$ *we have*

a) Under H_0^c,

$$J_{\mathbb{N}}^{c} = \frac{\mathbb{N}|H_z|^{1/2}\widehat{I}_{\mathbb{N}}^{c}}{\sqrt{\widehat{\Sigma}_c}} \xrightarrow{d} \mathcal{N}(0,1),$$

where

$$\widehat{\Sigma}_c = \prod_{r=1}^{R}\binom{N_r}{2}^{-1}\binom{T}{2}^{-1}|H_z|^{-1}\sum_{ij\ldots lt}\sum_{(i'j'\ldots l't')\neq(ij\ldots lt)} \widehat{\widetilde{v}}^{2}_{ij\ldots lt}(z_1)\widehat{\widetilde{v}}^{2}_{i'j'\ldots l't'}(z_1)X_{ij\ldots lt}^{\top}X_{ij\ldots lt}X_{i'j'\ldots l't'}^{\top}X_{i'j'\ldots l't'} \times K^2(Z_{ij\ldots lt}, Z_{i'j'\ldots l't'})$$

is a consistent estimator of the asymptotic variance of $\mathbb{N}|H_z|^{1/2}\widehat{I}_{\mathbb{N}}^{c}$ *for*

$$\Sigma_c = \frac{2^{(R+1)}(T-1)\sigma_v^4 R^q(K)}{T^3}E[E(tr\{X_{ij\ldots lt}X_{i'j'\ldots l't'}^{\top}\}^2|Z_{ij\ldots lt})f(Z_{ij\ldots lt})].$$

b) Under H_1^c, $Pr(J_{\mathbb{N}}^{c} \geq C^c) \to 1$ *as* $\mathbb{N}_{\max} \to \infty$*, where* C^c *is any positive constant.*

The proof of this theorem is in the Appendix. Despite the fact that the test statistic $J_{\mathbb{N}}^{c}$ is based on nonparametric residuals, under the necessary assumptions of Theorem 5.4, the limiting properties of $J_{\mathbb{N}}^{c}$ are quite similar to what is assumed for $J_{\mathbb{N}}^{a}$ and $J_{\mathbb{N}}^{b}$, which are based on fully parametric residuals (each test is a nonparametric kernel-based test).

We recommend a bootstrap procedure for using this test in practice. Let $\widehat{\widetilde{v}}_{ij\ldots l}(z_1) = (\widehat{\widetilde{v}}_{ij\ldots l1}(z_1), \ldots, \widehat{\widetilde{v}}_{ij\ldots lT}(z_1))^{\top}$ be a $T \times 1$ vector of the within transformed nonparametric residuals from a varying coefficient regression model. We propose to follow a similar procedure as in Section 5.1 with $\widehat{\widetilde{v}}_{ij\ldots l}(z_1)$ instead of $\widehat{\widetilde{v}}_{ij\ldots l,a}$. Using the bootstrap sample to estimate $\beta(\cdot)$ via the LSDV procedure proposed before and denoting these estimates by $\widehat{\beta}^{*}(\cdot)$, the bootstrap residuals are computed as $\widehat{\widetilde{v}}^{*}_{ij\ldots lt}(z_1)$. The bootstrap test statistic $J_{\mathbb{N}}^{c*}$ is obtained as we did with $J_{\mathbb{N}}^{c}$, but with $\widehat{\widetilde{v}}^{*}_{ij\ldots lt}(z_1)$ instead of $\widehat{\widetilde{v}}_{ij\ldots lt}(z_1)$. The process is repeated B times and the empirical distribution of the B bootstrap statistics is used to approximate the distribution of the test statistic $J_{\mathbb{N}}^{c*}$ under H_0^c.

5.3. **Testing Random vs. Fixed Effects Frameworks.** In this section, we discuss how to test for the presence of random effects versus fixed effects in a varying coefficient multidimensional panel data model as in Equation (1.2). The random effects specification assumes that the unobserved heterogeneity (for individuals, group, and/or time) are uncorrelated with the regressors $X_{ij\dots lt}$ and/or $Z_{ij\dots lt}$, while for the fixed effects specification we will allow these heterogeneities to be correlated with some of the regressors in an unknown way.

We are interested in testing that all of the unobserved individual, temporal, and interactive effects are uncorrelated with all of the covariates (H_0^d) versus the alternative that they may be correlated with some of the covariates (H_1^d). The null and alternatives hypotheses can be written as

$$\begin{aligned} H_0^d: &\quad Pr\{E(\pi_{ij\dots lt}|Z_{ij\dots l1},\dots,Z_{ij\dots lT},X_{ij\dots l1},\dots,X_{ij\dots lT})=0\}=1 \quad \text{for all} \quad i,j,\dots,l \quad \text{and} \quad t \\ H_1^d: &\quad Pr\{E(\pi_{ij\dots lt}|Z_{ij\dots l1},\dots,Z_{ij\dots lT},X_{ij\dots l1},\dots,X_{ij\dots lT})\neq 0\}>0 \quad \text{for some} \quad i,j,\dots,l \quad \text{and} \quad t. \end{aligned} \tag{5.10}$$

We maintain the assumption that $E(v_{ij\dots lt}|X_{ij\dots l1},\dots,X_{ij\dots lT},Z_{ij\dots l1},\dots,Z_{ij\dots lT})=0$ under both H_0^d and H_1^d. Motivated by Sun et al. (2009), our test statistic is based on the squared difference between the fixed and random effects estimators, which is asymptotically zero under H_0^d and positive under H_1^d. Formally, we have

$$\begin{aligned} I_{\mathbb{N}}^d &= \int [\widehat{\beta}_{RE}(z)-\widehat{\beta}_{FE}(z)]^\top (X^\top K_{H_z}(z)X)^\top (X^\top K_{H_z}(z)X)[\widehat{\beta}_{RE}(z)-\widehat{\beta}_{FE}(z)]dz, \\ &= \int \widehat{v}(z)^\top K_{H_z}(z)XX^\top K_{H_z}(z)\widehat{v}(z)dz, \end{aligned}$$

given that $(X^\top K_{H_z}(z)X)[\widehat{\beta}_{RE}(z)-\widehat{\beta}_{FE}(z)] = X^\top K_{H_z}(z)\{Y - X\widehat{\beta}_{FE}(z)\} = X^\top K_{H_z}(z)\widehat{v}_d(z)$. $X^\top K_{H_z}(z)X$ has been used to avoid the random denominator problem. The corresponding nonparametric estimators are

$$\begin{aligned} \widehat{\beta}_{FE}(z) &= (X^\top K_{H_z}(z)X)^{-1}X^\top K_{H_z}(z)MY, \\ \widehat{\beta}_{RE}(z) &= (X^\top K_{H_z}(z)X)^{-1}X^\top K_{H_z}(z)Y. \end{aligned}$$

Removing the non-zero center terms and using the transformed residuals, $\widehat{\widetilde{v}}_{d,ij\dots lt}(z)$ instead of $\widehat{v}_{d,ij\dots lt}(z)$, the resulting feasible test statistic is

$$\begin{aligned} \widehat{I}_{\mathbb{N}}^d &= \prod_{r=1}^{R} \frac{1}{\mathbb{N}T(N_r-1)(T-1)} \sum_{ij\dots lt} \sum_{(i'j'\dots l't')\neq(ij\dots lt)} \widehat{\widetilde{v}}_{d,ij\dots lt}(z)\widehat{\widetilde{v}}_{d,i'j'\dots l't'}(z)X_{ij\dots lt}^\top X_{i'j'\dots l't'} \\ &\quad \times K_{H_z}(Z_{ij\dots lt},Z_{i'j'\dots l't'}). \end{aligned} \tag{5.11}$$

This leads to our theoretical result:

Theorem 5.5. *Suppose that Assumptions 3.1–3.5 and 3.6–3.7 hold. As* $\mathbb{N}_{\max} \to \infty$ *tends to infinity for* T *fixed,* $\sqrt{\mathbb{N}_{\max}|H_z|}\|H_z\|^4 = O(1)$ *and we have*

a) Under H_0^d,

$$J_{\mathbb{N}}^d = \frac{\mathbb{N}|H_z|^{1/2}\widehat{I}_{\mathbb{N}}^d}{\sqrt{\widehat{\Sigma}_d}} \xrightarrow{d} \mathcal{N}(0,1),$$

where

$$\widehat{\Sigma}_d = \prod_{r=1}^{R}\binom{N_r}{2}^{-1}\binom{T}{2}^{-1}|H_z|^{-1}\sum_{ij\ldots lt}\sum_{(i'j'\ldots l't')\neq(ij\ldots lt)}\widehat{\widetilde{v}}^2_{d,ij\ldots lt}(z)\widehat{\widetilde{v}}^2_{d,i'j'\ldots l't'}(z)X^\top_{ij\ldots lt}X_{i'j'\ldots l't'} \times X^\top_{i'j'\ldots l't'}X_{ij\ldots lt}K^2(Z_{ij\ldots lt}, Z_{i'j'\ldots l't'})$$

is a consistent estimator of the asymptotic variance of $\mathbb{N}|H_z|^{1/2}\widehat{I}_{\mathbb{N}}^d$ *for*

$$\Sigma_d = \frac{2^{(R+1)}(T-1)\sigma_v^4 R^q(K)}{T^3}E[E(\{X_{ij\ldots lt}X^\top_{i'j'\ldots l't'}\}^2|Z_{ij\ldots lt})f(Z_{ij\ldots lt})].$$

b) Under H_1^d, $Pr(J_{\mathbb{N}}^d \geq C^d) \to 1$ *as* $\mathbb{N}_{\max} \to \infty$, *where* C^d *is any positive constant.*

The proof of this theorem follows from similar reasoning as in the proof of Theorem 5.3. For the sake of brevity it has been omitted, but the detailed proof is available upon request.

We note that this test statistic enables us to detect if any of the elements in π_{ijt} are correlated with $X_{ij\ldots lt}$ and/or $Z_{ij\ldots lt}$. Nevertheless, if we are interested in checking whether one particular element in $\pi_{ij\ldots lt}$ (i.e., μ_i for example) is correlated with $X_{ij\ldots lt}$ and/or $Z_{ij\ldots lt}$, the proposed test statistic can be directly extended to accommodate this case with the appropriately defined residuals.

To implement this test in practice, we propose to use a (slightly different from above) residual-based wild bootstrap procedure. We define $\widehat{v}_{ij\ldots l}(z_1) = (\widehat{v}_{ij\ldots l1}(z_1),\ldots,\widehat{v}_{ij\ldots lT}(z_1))^\top$ as a T-dimensional vector of random effects residuals (the null), where $\widehat{v}_{ij\ldots lt}(z_t) = Y_{ij\ldots lt} - X^\top_{ij\ldots lt}\widehat{\beta}_{RE}(Z_{ij\ldots lt}; H_z)$ and $\widehat{\beta}_{RE}(\cdot; H_z)$ is the random effect nonparametric estimator of (3.5) of the form $\widehat{\beta}_{RE}(z; H_z) = g_{H_z}(z)^\top Y$. We compute the two-point wild bootstrap errors as in Section 5.1 with $\widehat{v}_{ij\ldots l}(z_1)$ instead of $\widehat{\widetilde{v}}_{ij\ldots l}$. Using the bootstrap sample, we calculate the fixed effects estimator[6] for $\beta(\cdot)$, denoted by $\widehat{\beta}^*(\cdot)$, and obtain the bootstrap residuals by $\widehat{v}^*_{ij\ldots lt}(z_1) = Y^*_{ij\ldots lt} - X^\top_{ij\ldots lt}\widehat{\beta}^*(Z_{ij\ldots lt}; H_z)$. The bootstrap test statistic $J_{\mathbb{N}}^{d*}$ is obtained as we did with $J_{\mathbb{N}}^d$ except that $\widehat{v}_{ij\ldots lt}(z_1)$ is replaced by $\widehat{v}_{ij\ldots lt}(z_1)$. This process is repeated a large

[6]Henderson et al. (2008) suggest using a random effects estimator. However, Amini et al. (2012) point out that the performance of the Hausman test is independent of the estimator used within the bootstrap, but in finite samples the use of the fixed effects estimator leads to improved size.

number (B) of times to obtain the empirical distribution of $J_{\mathbb{N}}^{d*}$. We reject the null if $J_{\mathbb{N}}^{d}$ is large relative to the empirical distribution of $J_{\mathbb{N}}^{d*}$.

6. Finite Sample Performance

We use Monte Carlo simulations to assess how our proposed estimator and test statistics perform in finite sample settings. Section 6.1 examines the finite sample behavior of the varying coefficient LSDV estimator proposed in Equation (3.10). Section 6.2 analyzes the behavior of the proposed tests, examining both their size and power properties.

6.1. **Estimation.** Given that the three dimensional case is the most common setting currently in the literature, we consider the following specification:

$$Y_{ijt} = X_{ijt}\beta(Z_{ijt}) + \pi_{ijt} + v_{ijt}, \tag{6.1}$$

where $\beta(Z_{ijt}) = sin(\pi Z_{ijt})$, Z_{ijt} is generated as an *i.i.d.* uniform $[-1, 1]$ random variable, $X_{ijt} = 0.5X_{ij(t-1)} + \xi_{ijt}$, where ξ_{ijt} is *i.i.d.* $N(0, 1)$, and v_{ijt} is *i.i.d.* $N(0, 1)$. In addition, we consider three different data generating processes (DGPs) for the unobserved heterogeneities and interactive effects

DGP1: $\pi_{ijt} = \mu_{1i} + \mu_{2j} + \mu_{3t} + v_{ijt}$,
DGP2: $\pi_{ijt} = \gamma_{1ij} + \mu_{3t} + v_{ijt}$,
DGP3: $\pi_{ijt} = \mu_{1i} + \mu_{2j} + \mu_{3t} + \gamma_{1ij} + \gamma_{2it} + v_{ijt}$,

where $\mu_{1i} = \varphi_i + c\overline{X}_i$, $\mu_{2j} = \varphi_j + c\overline{X}_j$, $\mu_{3t} = \varphi_t + c\overline{X}_t$, $\gamma_{1ij} = \varphi_{ij} + c\overline{X}_{ij}$, $\gamma_{2it} = \varphi_{it} + c\overline{X}_{it}$, and $\gamma_{3jt} = \varphi_{jt} + c\overline{X}_{jt}$, where φ_i, φ_j, φ_t, φ_{ij}, φ_{it}, φ_{jt} are each $i.i.d$ $N(0, 0.1^2)$, $\overline{X}_i = (N_2T)^{-1}\sum_{jt} X_{ijt}$, $\overline{X}_j = (N_1T)^{-1}\sum_{it} X_{ijt}$, and $\overline{X}_t = (N_1N_2)^{-1}\sum_{ij} X_{ijt}$. X_{ijt}, Z_{ijt}, v_{ijt}, and all the unobserved heterogeneities and interactive effects are mutually independent of each other, but X_{ijt} and the elements in π_{ijt} are correlated (i.e., the fixed effects case). When $c = 0$, we have to deal with the random effects case. To assess the robustness of the LSDV estimator to the presence of fixed effects, we let $c = 0.8$, 1, and 1.2.

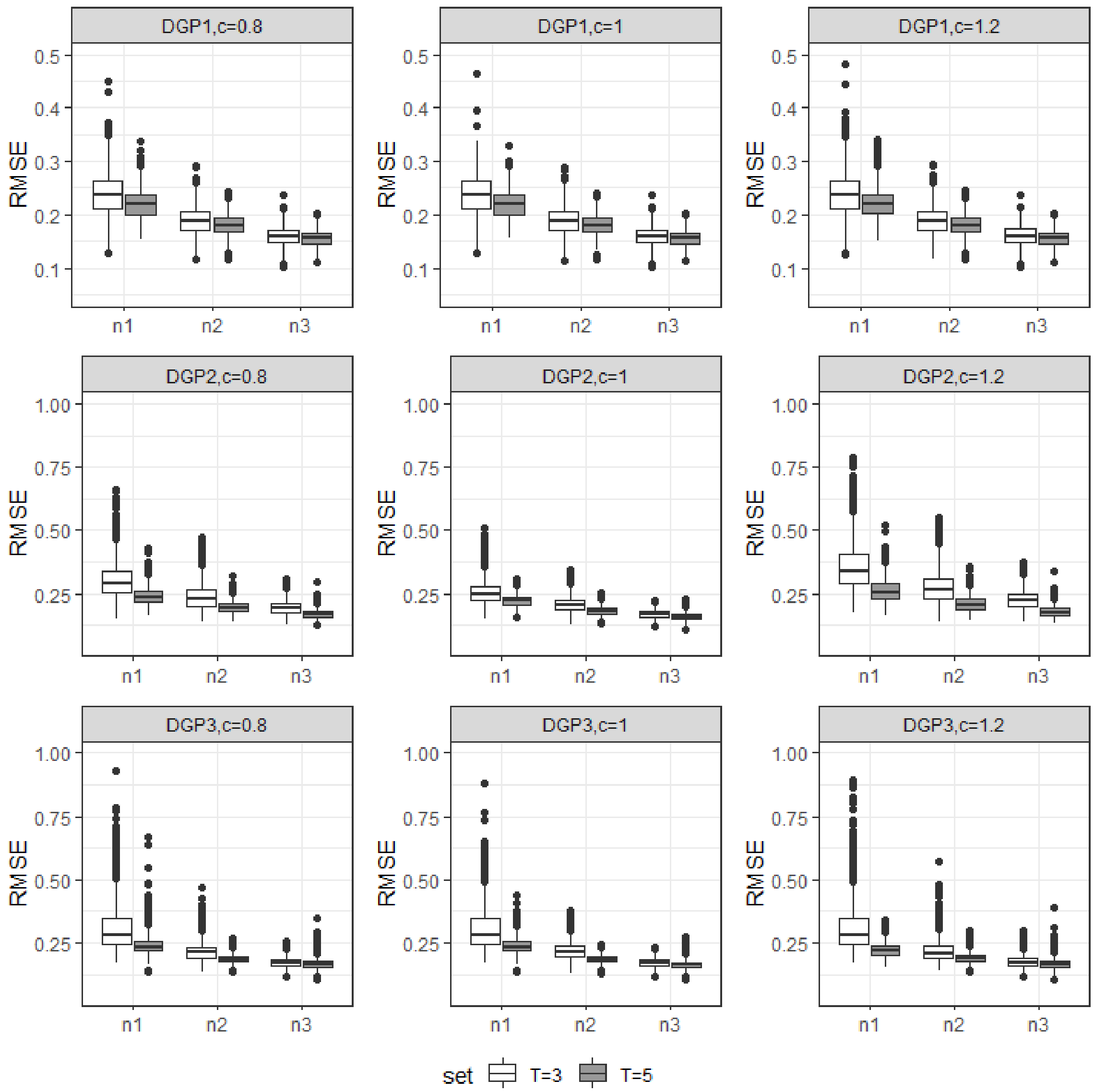
The number of Monte Carlo replications is set to 1,000, the number of cross-sections $\mathbb{N} = (N_1N_2)$ are varied from $(15, 10)$, $(20, 15)$, and $(25, 20)$, while the number of time periods is either 3 or 5. The Epanechnikov kernel function $k(u) = 0.75(1 - u^2)\mathbb{1}\{|u| \leq 1\}$ is used and the bandwidth is chosen via $H_z = h_zI_q = 2.34\widehat{\sigma}_z\mathbb{N}^{-1/5}$, where $\widehat{\sigma}_z$ is the sample standard deviation of Z_{ijt}. To evaluate the estimation accuracy of the proposed estimator, we compute the averaged mean squared errors (AMSE) of $\widehat{\beta}(z)$, defined as

$$AMSE(\widehat{\beta}(Z_{ijt}, H_z)) = \frac{1}{Q}\sum_{\varrho=1}^{Q}\frac{1}{N_1N_2T}\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\sum_{t=1}^{T}[\widehat{\beta}(Z_{ijt}, H_z) - \beta(Z_{ijt})]^2,$$

where ϱ refers to the ϱth simulation replication and Q is the total number of replications.

Figure 1 depicts boxplots of the 1,000 AMSE values of the varying coefficient LSDV estimator. Consistency of our estimator is visually shown by all AMSE values converging toward zero as either $\mathbb{N}$ or T increases. As expected, the simulation results are largely unaffected by the value of c. We argue that our LSDV estimator is robust to the presence of fixed effects in the variety of forms considered in $DGP1$–$DGP3$.

Figure 1. Boxplots of the AMSE values of the varying coefficient LSDV estimator in 1,000 independent simulations.



Note: $n_1 = \mathbb{N} = (15, 10)$, $n_2 = \mathbb{N} = (20, 15)$, and $n_3 = \mathbb{N} = (25, 20)$.

6.2. **Inference.**

6.2.1. *Specification.* In order to illustrate the performance of our proposed tests for correct functional form in finite samples, we consider the following three specifications:

$$Y_{ijt} = X_{ijt}\beta_1 + Z_{ijt}\beta_2 + \pi_{ijt} + v_{ijt}, \tag{6.2}$$

$$Y_{ijt} = \beta(X_{ijt}, Z_{ijt}) + \pi_{ijt} + v_{ijt}, \tag{6.3}$$

$$Y_{ijt} = X_{ijt}\beta(Z_{ijt}) + \pi_{ijt} + v_{ijt}, \tag{6.4}$$

where (6.2) is a fully parametric model, (6.3) represents the fully nonparametric model, and (6.4) is a varying coefficient regression model. For each specification, π_{ijt} is constructed as in DGP1-DGP3 and the other random variables are generated as in Section 6.1. Also, it is assumed $\beta(X_{ijt}, Z_{ijt}) = sin(2\pi(X_{ijt} + Z_{ijt}))$, $\beta(Z_{ijt}) = sin(2Z_{ijt})$, $\beta_1 = 5$, and $\beta_2 = 2$. We only report results for $c = 1$ given that our test statistics are invariant to different values of c, as shown in Section 6.1. We take $\mathbb{N} = (N_1 N_2)$ varying from $(10, 5)$, $(15, 10)$, and $(20, 15)$ and T equal to 3, the number of Monte Carlo replications is 1,000, and within each replication, 200 bootstrap iterations are conducted to estimate the 1%, 5% and 10% upper percentile values of the null distribution of our test statistics.

TABLE 2. Estimated size and power for the specification test $\widehat{J}^a_{\mathbb{N}}$ based on bootstrap p-values.

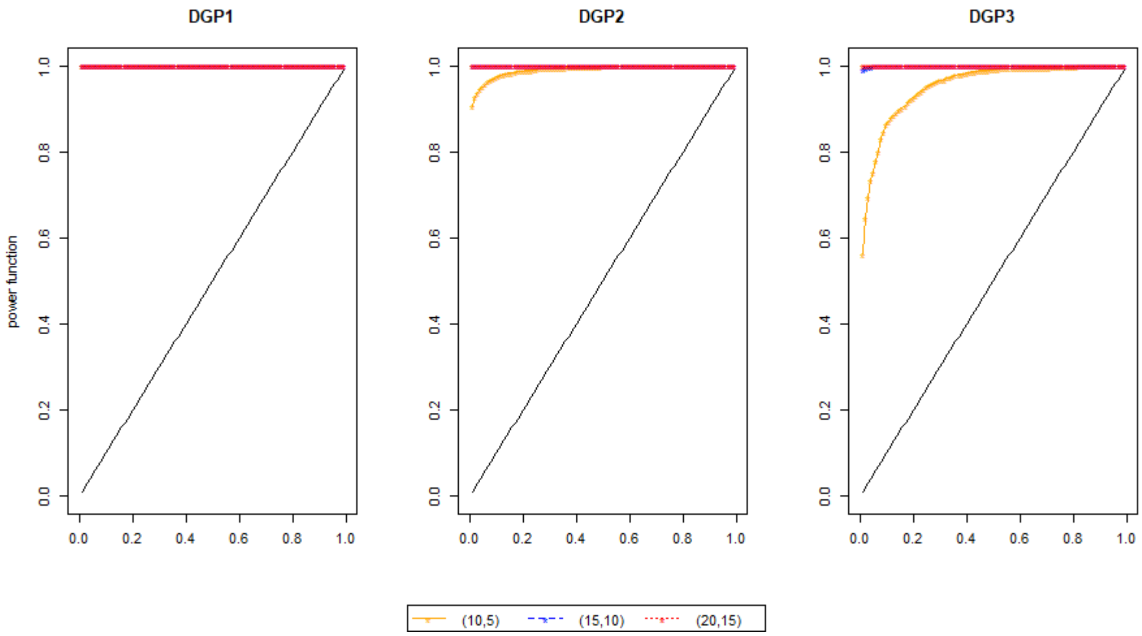
DGP	N	T	10%	5%	1%	10%	5%	1%
				Size			Power	
DGP1	(10,5)	3	0.095	0.045	0.015	1.000	1.000	1.000
	(15,10)	3	0.113	0.061	0.015	1.000	1.000	1.000
	(20,15)	3	0.112	0.061	0.009	1.000	1.000	1.000
DGP2	(10,5)	3	0.110	0.057	0.017	0.974	0.951	0.906
	(15,10)	3	0.098	0.047	0.014	1.000	1.000	1.000
	(20,15)	3	0.105	0.054	0.012	1.000	1.000	1.000
DGP3	(10,5)	3	0.104	0.055	0.009	0.865	0.753	0.563
	(15,10)	3	0.106	0.056	0.015	1.000	1.000	0.992
	(20,15)	3	0.105	0.054	0.012	1.000	1.000	1.000

Table 3. Estimated size and power for the specification test $\widehat{J}^b_{\mathbb{N}}$ based on bootstrap p-values.

DGP	N	T	10%	5%	1%	10%	5%	1%
				Size			Power	
DGP1	(10,5)	3	0.083	0.049	0.017	0.946	0.908	0.824
	(15,10)	3	0.097	0.046	0.017	0.999	0.995	0.989
	(20,15)	3	0.095	0.052	0.016	1.000	0.999	0.996
DGP2	(10,5)	3	0.115	0.059	0.018	0.904	0.845	0.699
	(15,10)	3	0.097	0.035	0.012	0.993	0.990	0.975
	(20,15)	3	0.091	0.040	0.012	0.998	0.998	0.990
DGP3	(10,5)	3	0.084	0.036	0.012	0.852	0.765	0.570
	(15,10)	3	0.098	0.052	0.009	0.994	0.987	0.955
	(20,15)	3	0.114	0.065	0.018	1.000	0.997	0.990

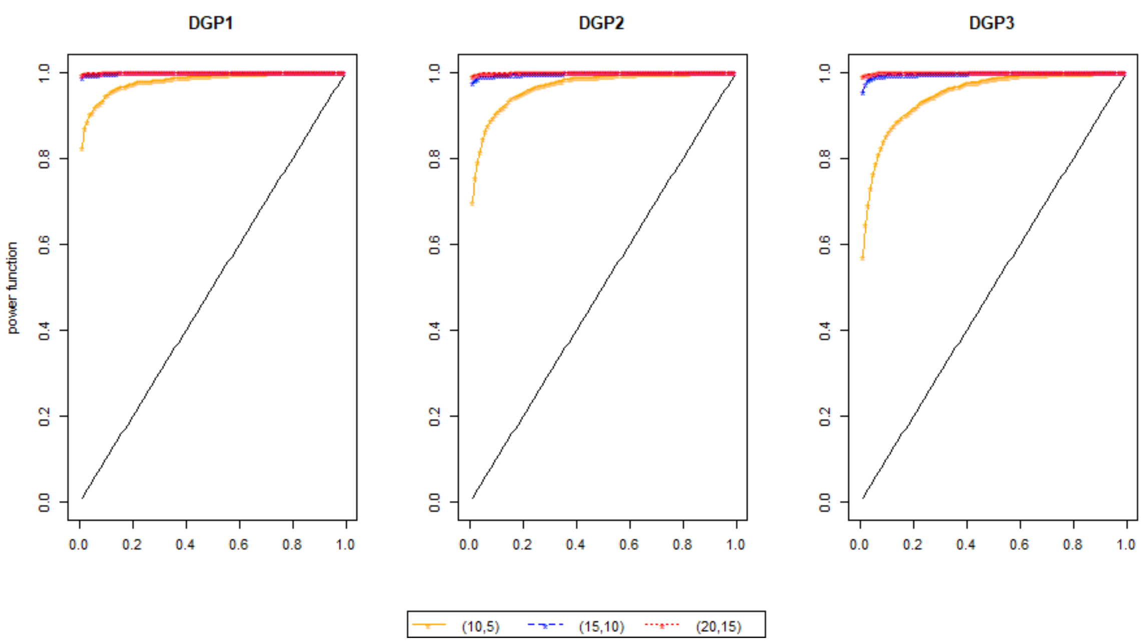
Tables 2 and 3 present the estimated size and power for the specification test for H^a_0 against H^a_1 (parametric versus nonparametric), i.e., $\widehat{J}^a_{\mathbb{N}}$, and for H^b_0 versus H^b_1 (parametric against a varying coefficient model), i.e., $\widehat{J}^b_{\mathbb{N}}$, respectively. Both test statistics have reasonable empirical size even for relatively small samples across all percentile values. The tests $\widehat{J}^a_{\mathbb{N}}$ and $\widehat{J}^b_{\mathbb{N}}$ display high power with their rejection rates increasing as the sample size increases. To better demonstrate this power, Figures 2 and 3 plot power functions against the nominal levels. The performance of these tests statistics are quite impressive, with power increasing rapidly to 1 across the three DGPs considered.

FIGURE 2. Power curves under H_1^a.



Note: Plot of power curves against nominal levels for $\mathbb{N} = (10, 5)$ *(orange line),* $\mathbb{N} = (15, 10)$ *(blue line),* $\mathbb{N} = (20, 15)$ *(red line) are depicted, respectively, with* $T = 3$ *in all cases.*

FIGURE 3. Power curves under H_1^b.



Note: Plot of power curves against nominal levels for $\mathbb{N} = (10, 5)$ *(orange line),* $\mathbb{N} = (15, 10)$ *(blue line),* $\mathbb{N} = (20, 15)$ *(red line) are depicted, respectively, with* $T = 3$ *in all cases.*

6.2.2. *Statistical significance.* For the finite sample behavior of the proposed test for statistical significance in a varying coefficient specification, we consider the following DGP

$$Y_{ijt} = X_{ijt}\beta(Z_{1ijt}, Z_{2ijt}) + \pi_{ijt} + v_{ijt}, \tag{6.5}$$

where Z_{1ijt} and Z_{2ijt} are generated as *i.i.d.* uniform $[-1,1]$ and uniform $[2,4]$ random variables, respectively, and the rest of the variables are generated as in Section 6.1. Under H_0^c, $\beta(Z_{1ijt}, Z_{2ijt}) = sin(2\pi Z_{1ijt})$, while under H_1^c, $\beta(Z_{1ijt}, Z_{2ijt}) = sin(2\pi(Z_{1ijt} + Z_{2ijt}))$. We take $\mathbb{N} = (N_1 N_2)$ varying from $(10,5)$, $(15,10)$, and $(20,15)$ and T equal to 3. The number of Monte Carlo replications is 1,000, and within each replication, 200 bootstrap iterations are conducted to estimate the 1%, 5% and 10% upper percentile values of the null distribution of $\widehat{J}_{\mathbb{N}}^c$.

Table 4 gives the proportion of rejections for the irrelevant regressors test at standard nominal levels (for DGP1-DGP3). To fulfill the conditions in Theorem 5.4 related to the bandwidth (i.e., $\mathbb{N}T|H_z| \to \infty$ and $\sqrt{\mathbb{N}T|H_z|}\|H_z\|^4 \to 0$, as $\mathbb{N} \to \infty$), we take $H_{z_1} \sim \mathbb{N}^{-\alpha_0}$ and $H_{z_2} \sim \mathbb{N}^{-\alpha_0}$, which require $\alpha_0 \in ((2/9),1)$. We use $H_z = H_{z_1} \times H_{z_2}$, where $H_{z_1} = a\widehat{\sigma}_{z_1}\mathbb{N}^{-2/7}$, $H_{z_2} = a\widehat{\sigma}_{z_2}\mathbb{N}^{-2/7}$, and $\widehat{\sigma}_{z_1}$ and $\widehat{\sigma}_{z_2}$ are the sample standard deviation of Z_{1ijt} and Z_{2ijt}, respectively. Finally, as Yao and Wang (2020) and Lavergne and Vuong (2000) point out, these type of nonparametric statistics depend heavily on the choice of bandwidth. To assess this dependence, we select a from $(0.5, 1.0, 1.5)$.

Table 4 shows that when H_0^c is true, the level of our test behaves reasonably well across all the DGPs with the rejection rate decreasing as the sample size increases. The sensitivity of the test to the bandwidth choice is evident; for DGPs 2 and 3, the test over-rejects for large values of a. In addition, when H_0^c does not hold, the behavior of the estimated power in Table 4 suggest the test is working. As $\mathbb{N}$ increases, the power of our test generally increases rapidly. The choice of bandwidth appears to have some effect on the power of our test with a larger value of a tending to yield higher power.

We also demonstrate the power performance of the test statistic $\widehat{J}_{\mathbb{N}}^c$, in Figure 4 where the power functions are plotted against the nominal levels across the three DGPs. We can conclude that the test statistic performs well in finite samples across different values of a for the bandwidth with power increasing towards 1 for all three DGPs.

6.2.3. *Random vs. Fixed Effects.* Finally, we examine the finite sample behavior of the nonparametric test proposed for detecting random effects against fixed effects in a varying coefficient specification. We consider the same three DGPS for (6.1), where $c = 0$ gives the random effects model and $c \neq 0$ leads to the fixed effects model. Our bandwidth is computed as $H_z = hI_q = a\widehat{\sigma}_z\mathbb{N}^{-1/5}$ where we take $a = (0.8, 1.0, 1.2)$ in order to check the sensibility

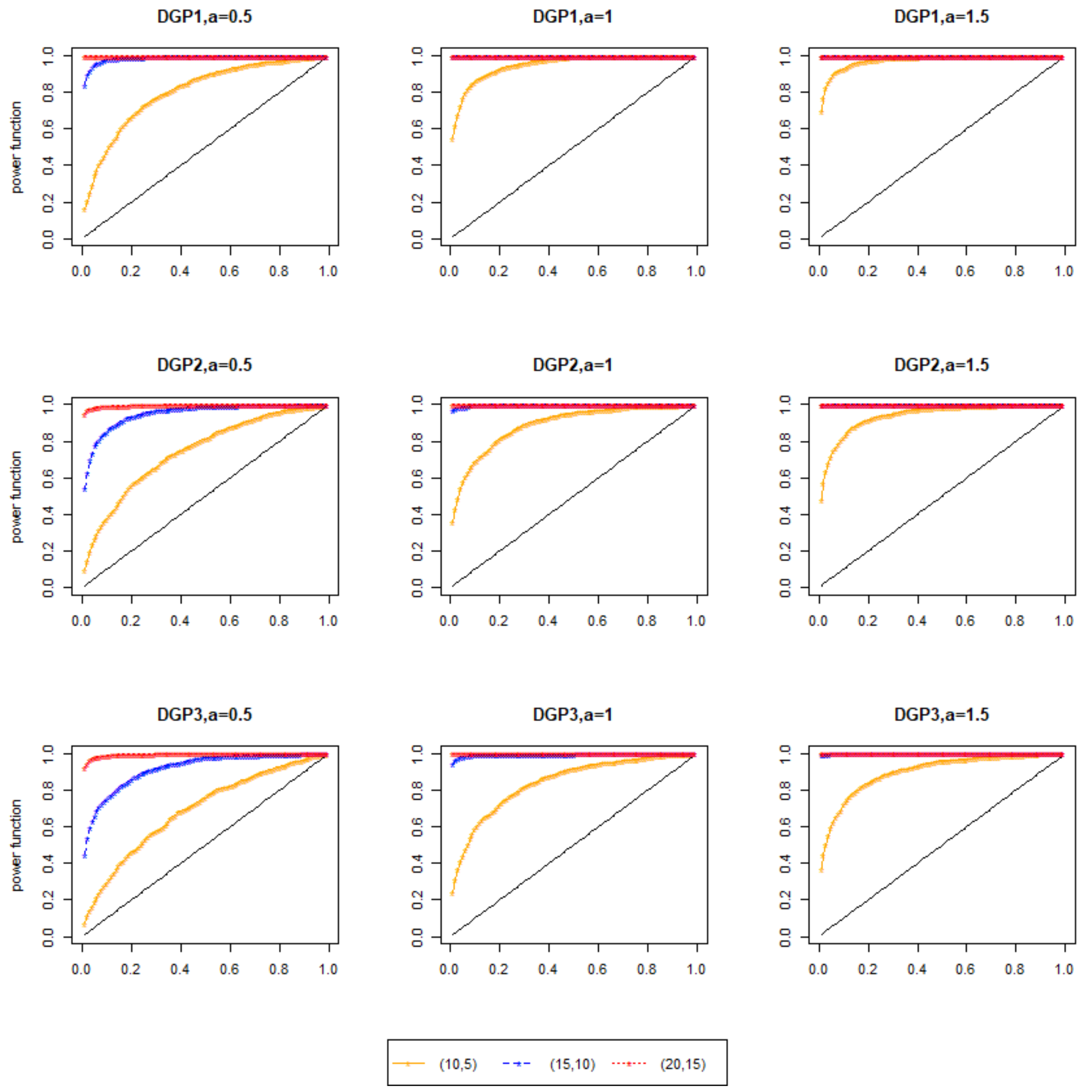
TABLE 4. Estimated size and power for the specification test $\widehat{J}^{c}_{\mathbb{N}}$ based on bootstrap p-values.

DGP	a	N	T	10%	5%	1%	10%	5%	1%
				Size			Power		
DGP1	0.5	(10,5)	3	0.098	0.057	0.017	0.478	0.347	0.164
		(15,10)	3	0.100	0.040	0.011	0.981	0.953	0.844
		(20,15)	3	0.107	0.055	0.016	1.000	0.999	0.998
	1.0	(10,5)	3	0.107	0.060	0.021	0.857	0.770	0.552
		(15,10)	3	0.106	0.052	0.014	1.000	1.000	1.000
		(20,15)	3	0.100	0.044	0.010	1.000	1.000	1.000
	1.5	(10,5)	3	0.134	0.077	0.029	0.928	0.876	0.698
		(15,10)	3	0.115	0.061	0.015	1.000	1.000	1.000
		(20,15)	3	0.095	0.057	0.015	1.000	1.000	1.000
DGP2	0.5	(10,5)	3	0.185	0.105	0.032	0.376	0.264	0.100
		(15,10)	3	0.170	0.094	0.042	0.853	0.776	0.544
		(20,15)	3	0.162	0.082	0.027	0.992	0.985	0.952
	1.0	(10,5)	3	0.228	0.144	0.060	0.691	0.579	0.357
		(15,10)	3	0.215	0.130	0.044	0.999	0.989	0.969
		(20,15)	3	0.194	0.096	0.038	1.000	1.000	1.000
	1.5	(10,5)	3	0.283	0.190	0.071	0.819	0.718	0.480
		(15,10)	3	0.266	0.172	0.048	1.000	1.000	0.999
		(20,15)	3	0.231	0.130	0.045	1.000	1.000	1.000
DGP3	0.5	(10,5)	3	0.159	0.087	0.024	0.295	0.189	0.073
		(15,10)	3	0.139	0.074	0.022	0.754	0.663	0.450
		(20,15)	3	0.132	0.068	0.024	0.991	0.979	0.926
	1.0	(10,5)	3	0.171	0.113	0.039	0.588	0.442	0.240
		(15,10)	3	0.135	0.073	0.027	0.993	0.984	0.943
		(20,15)	3	0.109	0.074	0.023	1.000	1.000	1.000
	1.5	(10,5)	3	0.201	0.124	0.050	0.724	0.598	0.372
		(15,10)	3	0.144	0.079	0.027	1.000	1.000	0.995
		(20,15)	3	0.125	0.067	0.028	1.000	1.000	1.000

of this test statistic to the bandwidth choice. $\mathbb{N} = (N_1 N_2)$ varies from $(10, 5)$, $(15, 10)$, and $(20, 15)$ while T is equal to 3, the number of Monte Carlo replications is 1,000, and within each replication, 200 bootstrap iterations are performed.

The estimated size and power (under each scenario) of the Hausman test is presented in Table 5, whereas Figures 5 and 6 graph the corresponding power functions. The estimated size of our nonparametric Hausman-type test is close to nominal size when $c = 0$, but these

FIGURE 4. Power curves under H_1^c.



Note: Plot of power curves against nominal levels for $\mathbb{N} = (10, 5)$ *(orange line),* $\mathbb{N} = (15, 10)$ *(blue line),* $\mathbb{N} = (20, 15)$ *(red line) are depicted, respectively, with* $T = 3$ *in all cases.*

results are affected by the choice of the bandwidth. In all DGPs, smaller values of a tend to lead the test to under-reject. As c gradually departs from zero, the power of the tests quickly converges to 1 as $\mathbb{N}$ increases. These power results are corroborated by the power functions in Figures 5 and 6. We conclude that the proposed nonparametric Hausman test has good finite sample size and power.

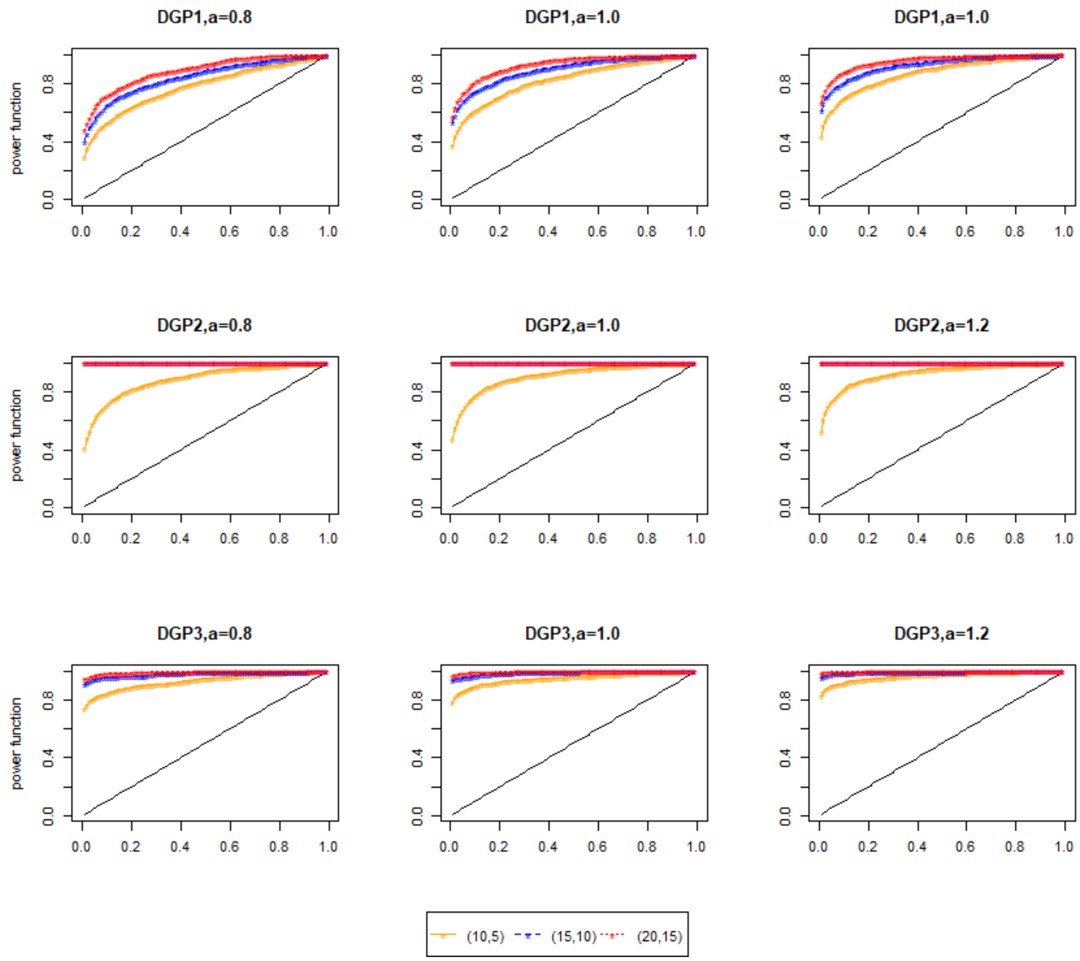
Table 5. Estimated size and power for the Hausman test $\widehat{J}^d_{\mathbb{N}}$ based on bootstrap p-values when some of the unobserved heterogeneity are fixed effects.

DGP	a	N	T	c	10%	5%	1%	c	10%	5%	1%	c	10%	5%	1%
						Size				Power				Power	
DGP1	0.8	(10,5)	3	0	0.043	0.019	0.005	0.8	0.526	0.438	0.297	1.2	0.767	0.720	0.602
		(15,10)	3		0.029	0.010	0.001		0.649	0.547	0.406		0.894	0.857	0.767
		(20,15)	3		0.019	0.005	0.001		0.713	0.629	0.481		0.956	0.924	0.849
	1.0	(10,5)	3	0	0.072	0.031	0.009	0.8	0.609	0.533	0.374	1.2	0.813	0.770	0.664
		(15,10)	3		0.062	0.030	0.005		0.742	0.677	0.534		0.926	0.903	0.840
		(20,15)	3		0.054	0.021	0.006		0.810	0.730	0.569		0.979	0.961	0.913
	1.2	(10,5)	3	0	0.141	0.060	0.010	0.8	0.694	0.606	0.442	1.2	0.869	0.819	0.716
		(15,10)	3		0.135	0.075	0.015		0.805	0.746	0.616		0.954	0.935	0.883
		(20,15)	3		0.105	0.057	0.016		0.877	0.807	0.672		0.990	0.975	0.941
DGP2	0.8	(10,5)	3	0	0.049	0.030	0.006	0.8	0.703	0.604	0.412	1.2	0.895	0.841	0.708
		(15,10)	3		0.057	0.033	0.007		0.999	0.999	0.998		1.000	1.000	1.000
		(20,15)	3		0.056	0.032	0.005		1.000	1.000	1.000		1.000	1.000	1.000
	1.0	(10,5)	3	0	0.059	0.028	0.009	0.8	0.769	0.670	0.475	1.2	0.926	0.887	0.769
		(15,10)	3		0.073	0.038	0.008		1.000	1.000	0.999		1.000	1.000	1.000
		(20,15)	3		0.069	0.036	0.007		1.000	1.000	1.000		1.000	1.000	1.000
	1.2	(10,5)	3	0	0.080	0.041	0.010	0.8	0.815	0.724	0.527	1.2	0.943	0.909	0.799
		(15,10)	3		0.092	0.050	0.015		1.000	1.000	0.999		1.000	1.000	1.000
		(20,15)	3		0.089	0.047	0.012		1.000	1.000	1.000		1.000	1.000	1.000
DGP3	0.8	(10,5)	3	0	0.053	0.026	0.008	0.8	0.843	0.812	0.742	1.2	0.922	0.902	0.857
		(15,10)	3		0.025	0.011	0.002		0.957	0.943	0.909		0.987	0.982	0.973
		(20,15)	3		0.021	0.011	0.004		0.982	0.966	0.943		0.999	0.998	0.988
	1.0	(10,5)	3	0	0.095	0.044	0.010	0.8	0.890	0.855	0.783	1.2	0.945	0.928	0.894
		(15,10)	3		0.065	0.031	0.009		0.968	0.955	0.935		0.992	0.990	0.979
		(20,15)	3		0.052	0.027	0.009		0.988	0.982	0.967		0.999	0.998	0.993
	1.2	(10,5)	3	0	0.128	0.077	0.018	0.8	0.915	0.894	0.827	1.2	0.958	0.949	0.915
		(15,10)	3		0.145	0.069	0.016		0.985	0.976	0.953		0.995	0.994	0.987
		(20,15)	3		0.117	0.058	0.018		0.995	0.992	0.981		1.000	1.000	0.998

7. Conclusion

Multidimensional panel models are becoming more popular in various applied economic domains given the availability of detailed datasets and practitioners awareness that various forms of heterogeneity are important to account for in economic modelling. Most of these multidimensional models have focused on the parametric setting, which exposes the analysis to various forms of misspecification. To broaden the scope of available methods, we have

FIGURE 5. Power curves under H_1^d when $c = 0.8$.



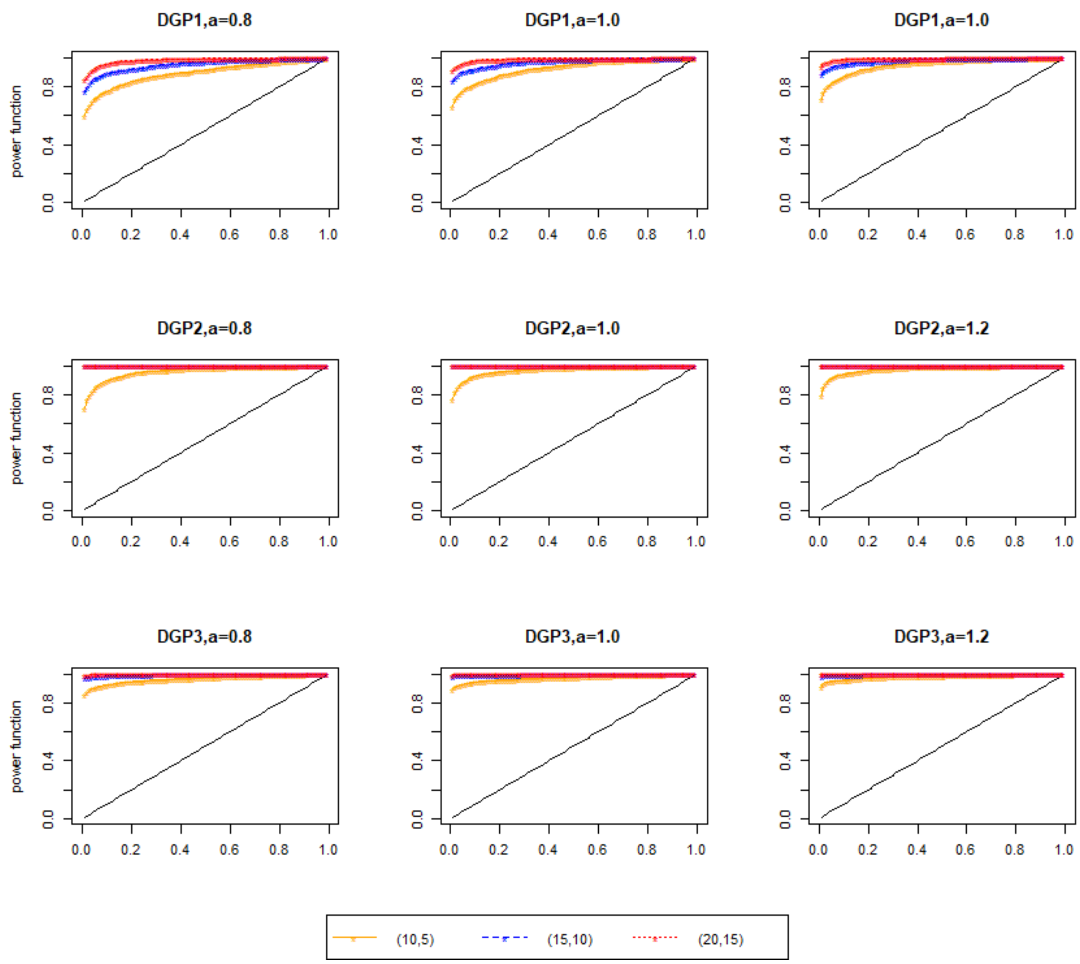
Note: Plot of power curves against nominal levels for $\mathbb{N} = (10, 5)$ *(orange line),* $\mathbb{N} = (15, 10)$ *(blue line),* $\mathbb{N} = (20, 15)$ *(red line) are depicted, respectively, with* $T = 3$ *in all cases.*

presented a semiparametric varying coefficient model tailored to the multidimensional panel data setting.

The estimator was built on the profile least squares approach and shown to posses standard asymptotic properties for this class of estimators. To make the estimator more appealing for the applied community, we also provided three important tests statistics. This required nontrivial extensions to Hall (1984)'s CLT for a second-order degenerate U-statistic. Our extension covers the case that more than one of the indices capturing heterogeneity in the multidimensional panel can grow unbounded.

With this new central limit theory, we were able to characterize the limiting distribution for test statistics for model misspecification, statistical significance and if the unobserved

FIGURE 6. Power curves under H_1^d when $c = 1.2$.



Note: Plot of power curves against nominal levels for $\mathbb{N} = (10, 5)$ *(orange line),* $\mathbb{N} = (15, 10)$ *(blue line),* $\mathbb{N} = (20, 15)$ *(red line) are depicted, respectively, with* $T = 3$ *in all cases.*

heterogeneity was correlated with the regressors. To demonstrate the finite sample behavior of our estimator and tests, we conducted a range of Monte Carlo simulations. The simulation results suggest that our estimator behaves well in reasonable sample sizes, while the proposed tests all had accurate empirical size and desirable power.

∎

APPENDIX

Remember that $g_{H_z}(z)^\top = (X^\top K_{H_z}(z)X)^{-1}X^\top K_{H_z}(z)$ is a $1 \times \mathbb{N}T$ vector with a typical element given by $g_{H_z,ij\ldots lt}(z)$. Before proceeding to the asymptotic analysis of the proposed estimator and tests, we first establish some Lemmas that will be key to prove that results. Let us denote $K(Z_{ij...lt}, Z_{i'j'...l't'}) = K(H_z^{-1}(Z_{ij...lt} - Z_{i'j'...l't'}))$, $K_{H_z}(Z_{ij...lt}, Z_{i'j'...l't'}) = |H_z|^{-1}K(H_z^{-1}(Z_{ij...lt} - Z_{i'j'...l't'}))$, and similar definitions for $K(X_{ij...lt}, X_{i'j'...l't'})$ and $K_{H_x}(X_{ij...lt}, X_{i'j'...l't'})$.

Lemma 7.1. *Under Assumptions 3.1, 3.4, and 3.6–3.7,*

$$g_{H_z,ij...lt}(z) = \Psi^{-1}(z)X_{ij...lt}K_{H_z}(Z_{ij...lt}, z) = \frac{1}{\mathbb{N}T}\mathcal{B}_{XX}^{-1}(z)X_{ij...lt}K_{H_z}(Z_{ij...lt}, z)\{1 + o_p(1)\}.$$

where $\Psi(z) = X^\top K_{H_z}(z)X$ *and* $\mathcal{B}_{XX}(z) = E[X_{ij...lt}X_{ij...lt}^\top | Z_{ij...lt} = z]f_{Z_{ij...lt}}(z)$.

Proof of Lemma 7.1: Considering the denominator term, it can be shown that, as any element in n tends to infinity,

$$\frac{1}{\mathbb{N}T}\Psi(z) = \mathcal{B}_{XX}(z)\{1 + o_p(1)\}. \quad \text{(I.1)}$$

In order to show this result, under Assumptions 3.1, 3.4 and 3.6, the law of iterated expectation (LIE) yields

$$\frac{1}{\mathbb{N}T}E(\Psi(z)) = E(X_{ij...lt}X_{ij...lt}^\top | Z_{ij...lt} = z)f_{Z_{ij...lt}}(z)\int K(u)du + o_p(\|H_z\|).$$

To conclude this proof, it is necessary to apply the Central Limit Theorem for which we need to show that $(\mathbb{N}T)^{-2}Var(\Psi) \to 0$, as $\mathbb{N}_{\max} \to \infty$. Under Assumption 3.1,

$$\begin{aligned}
\frac{1}{\mathbb{N}^2T^2}Var(\Psi) &= \frac{1}{\mathbb{N}T}Var(X_{ij...lt}X_{ij...lt}^\top K_{H_z}(Z_{ij...lt}, z)) \\
&+ \frac{1}{\mathbb{N}T^2}\sum_{\kappa=1}^{T-1}(\kappa - T)Cov(X_{ij...l1}X_{ij...l1}^\top K_{H_z}(Z_{ij...l1}, z), X_{ij...l(1+\kappa)}X_{ij...l(1+\kappa)}^\top K_{H_z}(Z_{ij...l(1+\kappa)}, z)) \\
&= O_p\left(\frac{1}{\mathbb{N}|H_z|}\right) + o_p\left(\frac{1}{\mathbb{N}|H_z|}\right).
\end{aligned}$$

Under Assumption 3.7, $\mathbb{N}|H_z| \to \infty$, so the variance term tends to zero. Then, using these results and by the Slutsky theorem, the proof of the Lemma is done.

∎

Lemma 7.2. *Let* $\mathcal{J}_{N_1-1} = \imath_{N_1-1}\imath_{N_1-1}^\top$ *and let* $(\mathcal{J}_{N_2-1}, \ldots, \mathcal{J}_{N_l-1}, \mathcal{J}_{T-1})$ *be defined in a similar way. Under Assumptions 3.1–3.7, it can be shown*

$$(D^\top PD)^{-1} = (D^\top D)^{-1} - O_p(\nu_\mathbb{N}), \quad \nu_\mathbb{N} = tr\{H_z^2\} + \frac{1}{\sqrt{\mathbb{N}|H_z|}},$$

where the specific structure of the matrix D depends on the form of the unobserved effects of the regression model.

Proof Lemma 7.2: This lemma can be proved adapting the proof scheme in Lemma A.5 in Gao and Kunpeng (2013) to this general case with composite fixed effects parameters. Note that the proof has been omitted for the sake of brevity.

∎

Proof of Theorem 3.1: Replacing (3.5) in (3.10) and rearranging terms, the local-constant least-squares estimator to analyze can be written as

$$\begin{aligned}\widehat{\beta}(z;H_z) &= (X^\top K_{H_z}(z)X)^{-1}X^\top K_{H_z}(z)[B\{X,\beta(Z)\}+D_0\theta+V]\\ &= \widehat{\beta}_1(z;H_z)+\widehat{\beta}_2(z;H_z)+\widehat{\beta}_3(z;H_z),\end{aligned}$$

where $\widehat{\beta}_1(z;H_z)=g_{H_z}(z)^\top MB\{X,\beta(Z)\}$, $\widehat{\beta}_2(z;H_z)=g_{H_z}(z)^\top MD_0\pi$, and $\widehat{\beta}_3(z;H_z)=g_{H_z}(z)^\top MV$. To derive the limiting results of $\widehat{\beta}(z)$, it is enough to show

$$\sqrt{\mathbb{N}|H_z|}(\widehat{\beta}(z;H_z)-\beta(z))-\sqrt{\mathbb{N}|H_z|}(\widehat{\beta}_1(z;H_z)+\widehat{\beta}_2(z;H_z)-\beta(z))=\sqrt{\mathbb{N}|H_z|}\widehat{\beta}_3(z;H_z), \tag{I.2}$$

where we will demonstrate that $(\widehat{\beta}_1(z;H_z)+\widehat{\beta}_2(z;H_z)-\beta(z))$ contributes to the asymptotic bias, whereas the right-hand side of (I.2) is asymptotically normal.

Analyzing the asymptotic behavior of $\widehat{\beta}_1(z;H_z)$, it can be decomposed into the following terms

$$\widehat{\beta}_1(z;H_z)=g_{H_z}(z)^\top B\{X,\beta(z)\}-g_{H_z}(z)^\top D(D^\top PD)^{-1}D^\top PB\{X,\beta(Z)\}. \tag{I.3}$$

Focusing on the first element and using a Taylor expansion yields

$$\begin{aligned}&g_{H_z}(z)^\top B\{X,\beta(z)\}-\beta(z)\\ &=(X^\top K_{H_z}(z)X)^{-1}\sum_{ij\ldots lt}X_{ij\ldots lt}K_{H_z}(X_{ij\ldots lt},x)X_{ij\ldots lt}^\top\{\beta(Z_{ij\ldots lt})-\beta(z)\}\\ &=\sum_{ij\ldots lt}g_{H_z,ij\ldots lt}(z)\Big\{X_{ij\ldots lt}^\top\otimes(Z_{ij\ldots lt}-z)^\top D_\beta(z)+\frac{1}{2}X_{ij\ldots lt}^\top\otimes(Z_{ijt}-z)^\top\mathcal{H}_\beta(z)(Z_{ij\ldots lt}-z)\\ &+X_{ij\ldots lt}^\top R_\beta(z)\Big\},\end{aligned}$$

where $R_\beta(z)$ is a vector of Taylor series remainder terms such as

$$R_\beta(Z_{ij\ldots lt},z)=\int_0^1\left[\frac{\partial^2\beta}{\partial z\partial z^\top}(z+\omega(Z_{ij\ldots lt}-z))-\frac{\partial^2\beta(z)}{\partial z\partial z^\top}\right](1-\omega)d\omega$$

and ω is a weight function. Therefore, the expression to analyze is

$$\widehat{\beta}_1(z;H_z)=\beta(z)+\mathcal{B}_{11}(z)+\mathcal{B}_{12}(z)-\mathcal{B}_{13}(z), \tag{I.4}$$

where

$$\begin{aligned}
\mathcal{B}_{11}(z) &= \sum_{ij\ldots lt} g_{H_z,ij\ldots lt}(z) X_{ij\ldots lt}^\top \otimes (Z_{ij\ldots lt}-z)^\top D_\beta(z) \\
&+ \frac{1}{2}\sum_{ij\ldots lt} g_{H_z,ij\ldots lt}(z) X_{ij\ldots lt}^\top \otimes (Z_{ij\ldots lt}-z)^\top \mathcal{H}_\beta(z)(Z_{ij\ldots lt}-z), \\
\mathcal{B}_{12}(z) &= \sum_{ij\ldots lt} g_{H_z,ij\ldots lt}(z) X_{ij\ldots lt}^\top \otimes (Z_{ij\ldots lt}-z)^\top R_\beta(Z_{ij\ldots lt},z)(Z_{ij\ldots lt}-z), \\
\mathcal{B}_{13}(z) &= g_{H_z}(z)^\top D(D^\top D)^{-1} D^\top B\{X,\beta(Z)\}.
\end{aligned}$$

Considering first the behavior of $\mathcal{B}_{11}(z)$ and using Lemma 7.1, under Assumptions 3.1, 3.4, and 3.6, the LIE yields

$$\begin{aligned}
E[\mathcal{B}_{11}(z)] &= \mathcal{B}_{XX}^{-1}(z) E[X_{ij\ldots lt}X_{ij\ldots lt}^\top | Z_{ij\ldots lt}=z] f_{Z_{ij\ldots lt}}(z) \otimes \int tr\{H_z^2 uu^\top K(u) du D_{\beta_r}(z) D_f(z)\} f_{Z_{ij\ldots lt}}^{-1}(z) \\
&+ \mathcal{B}_{XX}^{-1}(z) E[X_{ij\ldots lt}X_{ij\ldots lt}^\top | Z_{ij\ldots lt}=z] f_{Z_{ij\ldots lt}}(z) \otimes \int tr\{uu^\top H_z^2 \mathcal{H}_\beta(z)\} K(u) du \\
&+ o_p(\|H_z\|^2) \\
&= \mu_2^q(K) diag_d\left(tr\{H_z^2 D_f(z) D_{\beta_\kappa}(z)\}\right) \imath_d f_{Z_{ij\ldots lt}}^{-1}(z) + \frac{\mu_2^q(K)}{2} diag_d(tr\{H_z^2 \mathcal{H}_{\beta_\kappa}(z)\})\imath_d \\
&+ o_p(\|H_z\|^2).
\end{aligned}$$

Following a similar proof scheme as in (I.1), it is easy to show that $Var(\mathcal{B}_{11}(z)) = o_p\left(\frac{1}{\mathbb{N}|H_z|}\right)$, so we can conclude

$$\begin{aligned}
\mathcal{B}_{11}(z) &= \mu_2^q(K) diag_d(tr\{H_z^2 D_f(z) D_{\beta_\kappa}(z)\})\imath_d f_{Z_{ijt}}^{-1}(z) + \frac{\mu_2^q(K)}{2} diag_d(tr\{H_z^2 \mathcal{H}_{\beta_\kappa}(z)\})\imath_d \\
&+ o_p(\|H_z\|^2) + o_p\left(\frac{1}{\sqrt{\mathbb{N}|H_z|}}\right). \qquad \text{(I.5)}
\end{aligned}$$

Similarly, by Lemma 7.1 and the Slutsky theorem, under Assumptions 3.1, 3.4, and 3.6, as $\mathbb{N}_{\max} \to \infty$,

$$\mathcal{B}_{12} = o_p(\|H_z\|^2). \qquad \text{(I.6)}$$

In order to prove this, using the LIE and the law of large numbers,

$$\begin{aligned}
\mathcal{B}_{12}(z) &= \int \mathcal{B}_{XX}^{-1}(z) E(X_{ij\ldots lt}X_{ij\ldots lt}^\top | Z_{ij\ldots lt} = z + H_z u) \otimes (H_z u)^\top R_\beta(z+H_z u, z) H_z u K(u) \\
&\quad \times f(z+H_z u) du.
\end{aligned}$$

Let $\zeta(\eta)$ be the modulus of continuity of $\mathcal{H}_{\beta}(z)$, $R_{\beta}(z+H_z u, z) \leq \int_0^1 \zeta(\omega\|H_z u\|)(1-\omega)d\omega$. Using this property and by boundedness of $f_{Z_{ij\ldots lt}}(\cdot)$ and $\mathcal{B}_{XX}(\cdot)$, we get

$$\begin{aligned} |\mathcal{B}_{12}(z)| &\leq C\int\int_0^1 |(H_z u)^\top||\zeta(\omega\|H_z u\|)|(1-\omega)|H_z u|K(u)dud\omega \\ &\leq Ctr\left\{\int |u^\top H_z^2 u|K(u)du\right\}\int_0^1 |\zeta(\omega\|H_z u\|)|d\omega = o_p(\|H_z\|^2), \end{aligned}$$

given that it is assumed that as $\mathbb{N}_{\max}\to\infty$, $\zeta(\omega\|H_z u\|)\to 0$. This expression tends to zero by the dominated convergence at the rate $o_p(\|H_z\|^2)$ and so (I.5) is proved.

Focusing on the behavior of $\mathcal{B}_{13}(z)$ and following a similar reasoning as in Lemmas A.6-A.7 in Gao and Kunpeng (2013), $\mathcal{B}_{13}(z)$ can be written as

$$\begin{aligned} \mathcal{B}_{13}(z) &= gH_z(z)^\top D(D^\top D)^{-1}D^\top PB\{X,\beta(Z)\} + (s.o.) \\ &= \sum_{r=1}^{l}\frac{1}{N_r}\left[E[\beta(Z_{ij\ldots lt})] + O_p\left(\frac{1}{\sqrt{\mathbb{N}}}\right)\right] + \sum_{r=1}^{l}\sum_{r'\neq r}^{l}\frac{1}{N_r N_{r'}}\left[E[\beta(Z_{ij\ldots lt})] + O_p\left(\frac{1}{\sqrt{\mathbb{N}}}\right)\right] \\ &+ (s.o.). \end{aligned} \tag{I.7}$$

Using (I.5)-(I.7) in (I.4), we obtain

$$\begin{aligned} \widehat{\beta}_1(z;H_z) &= \mu_2^q(K)diag_d(tr\{H_z^2 D_f(z)D_{\beta_r}(z)\})\imath_d f_{Z_{ijt}}^{-1}(z) + \frac{\mu_2^q(K)}{2}diag_d(tr\{H_z^2\mathcal{H}_{\beta_r}(z)\})\imath_d \\ &+ o_p(\|H_z\|^2) + o_p\left(\frac{1}{\sqrt{\mathbb{N}|H_z|}}\right). \end{aligned} \tag{I.8}$$

Considering now the limiting result of $\widehat{\beta}_2(z;H_z)$, it is easy to show that $MD\pi = 0$, where $M = I_{\mathbb{N}T} - D(D^\top PD)^{-1}D^\top P$. Hence, by adding and subtracting terms and using Lemma 7.2, we can write

$$\begin{aligned} \widehat{\beta}_2(z;H_z) &= g_{H_z}(z)^\top MD_0\pi_0 \\ &= g_{H_z}(z)^\top[I_{\mathbb{N}T} - D(D^\top PD)^{-1}D^\top P][D_0\pi_0 - D\pi + D\pi] \\ &= g_{H_z}(z)^\top(D_0\pi_0 - D\pi) - g_{H_z}(z)^\top D(D^\top D)^{-1}D^\top P(D_0\pi_0 - D\pi) + (s.o.). \end{aligned}$$

Let $D_0\pi_0 - D\pi = \mathbb{N}T\overline{\pi}e_{\mathbb{N}T}$, where $\overline{\pi}$ is the corresponding cross-sectional or temporal average and $e_{\mathbb{N}T}$ is a $\mathbb{N}T\times 1$ vector having 1 in the entries related with the index of π_{ijt} and 0 otherwise. Using this result and following a similar reasoning as in (I.7), by strict stationarity condition,

$$\begin{aligned} \widehat{\beta}_2(z;H_z) &= \frac{\overline{\pi}}{\mathbb{N}T}\sum_{ij\ldots lt}\mathcal{B}_{XX}^{-1}(z)X_{1j\ldots lt}K_H(Z_{1j\ldots lt},z) + (s.o.) \\ &= \overline{\pi}B_{XX}^{-1}(z)E(X_{ij\ldots lt}|Z_{ij\ldots lt}=z)f_{Z_{ij\ldots lt}}(z)(1+o_p(1)) + (s.o.). \end{aligned} \tag{I.9}$$

Considering the limiting behavior of $\widehat{\beta}_2(z; H_z)$, using the result obtained in (I.9) and Assumption 3.2, it is possible to conclude that

$$\sqrt{\mathbb{N}|H_z|}\widehat{\beta}_2(z; H_z) = \sqrt{T|H_z|}\mathcal{B}_{XX}^{-1}(z)T^{-1}E(X_{ij\ldots lt}|Z_{ij\ldots lt} = z)f_{Z_{ij\ldots lt}}(z)\sqrt{\mathbb{N}}\overline{\pi}(1 + o_p(1)) \quad \text{(I.10)}$$

and it converges in probability to zero when $T|H_z| \to 0$ for T fixed. See Lemma A.9 in Gao and Kunpeng (2013) for further details.

Now we focus on the limiting behavior of $\widehat{\beta}_3(z; H_z)$. Using the results in Lemma 7.2, the expression to analyze is

$$\widehat{\beta}_3(z; H_z) = g_{H_z}(z)^\top V - g_{H_z}(z)^\top D(D^\top D)^{-1}D^\top V + (s.o.) \quad \text{(I.11)}$$

Under Assumptions 3.1, $E(\widehat{\beta}_3(z)) = 0$. Following a similar reasoning as in I.7, it is easy to show, as $\mathbb{N}_{\max} \to \infty$, the second term of the right-hand side is asymptotically negligible. Therefore, considering the variance term of $\widehat{\beta}_3(z; H_z)$, the expression to analyze is

$$\begin{aligned} \mathbb{N}|H_z|Var(\widehat{\beta}_3(z; H_z)) &= \mathbb{N}|H_z|E[(X^\top K_{H_z}(z)X)^{-1}X^\top K_{H_z}(z)VV^\top K_{H_z}(z)X(X^\top K_{H_z}(z)X)^{-1}] \\ &+ (s.o.). \end{aligned} \quad \text{(I.12)}$$

By Assumption 3.3, the $v_{ij\ldots lt}$'s are *i.i.d.* in the subscripts $(i, j, \ldots, l)$. Similar to above, by the law of iterated expectations, it can be shown that the middle term of (I.9) is

$$\begin{aligned} &\frac{|H_z|}{\mathbb{N}T^2}\sum_{ij\ldots l}\sum_{tt'} E[X_{ij\ldots lt}v_{ij\ldots lt}v_{ij\ldots lt'}X_{ij\ldots lt'}^\top K_{H_z}(Z_{ij\ldots lt}, z)K_{H_z}(Z_{ij\ldots lt'}, z)] \\ &= \frac{\sigma_v^2 R^q(K)}{T}\mathcal{B}_{XX}(z)(1 + o_p(1)). \end{aligned} \quad \text{(I.13)}$$

Replacing (I.13) and Lemma 7.1 in (I.12), the Slutsky's theorem yields

$$Var(\widehat{\beta}_3(z; H_z)) = \frac{\sigma_u^2 R^q(K)}{\mathbb{N}T|H_z|}\mathcal{B}_{XX}^{-1}(z)(1 + o_p(1)). \quad \text{(I.14)}$$

To complete the proof of the theorem, it is necessary to show that as any element in $\mathbb{N}$ tends to infinity,

$$\sqrt{\mathbb{N}|H_z|}\widehat{\beta}_3(z; H_z) \xrightarrow{d} N\left(0, \frac{\sigma_v^2 R^q(K)}{T}\mathcal{B}_{XX}^{-1}(z)\right), \quad \text{(I.15)}$$

for which the Lyapounov's condition has to be checked. With this aim, we write

$$\frac{\sqrt{\mathbb{N}|H_z|}}{\mathbb{N}T}\sum_{ij\ldots lt} X_{ij\ldots lt}v_{ij\ldots lt}K_{H_z}(Z_{ij\ldots lt}, z) = \frac{1}{T\sqrt{\mathbb{N}}}\sum_{ij\ldots l}\xi_{ij\ldots l},$$

where $\overline{\xi}_{ij\ldots l} = T^{-1}\sum_{t=1}^T \xi_{ij\ldots lt}$ for $\xi_{ij\ldots lt} = |H_z|^{1/2}X_{ij\ldots lt}v_{ij\ldots lt}K_{H_z}(Z_{ij\ldots lt}, z)$. Under the assumptions of the theorem, $\overline{\xi}_{ij\ldots l}$ is a sequence of *i.i.d.* random variables, given that T is assumed to be fixed. To prove the Lyapounov condition, we have to analyze

$$\prod_{r=1}^{l} \lim_{N_r \to \infty} s_{\mathbb{N}}^{-(2+\varepsilon)} \sum_{ij\dots l} |\bar{\xi}_{ij\dots lt}|^{(2+\varepsilon)} = 0,$$

where $s_{\mathbb{N}} = Var\left(\sum_{ij\dots l} \bar{\xi}_{ij\dots lt}\right) = \mathbb{N}T^{-1}Var(\xi_{ij\dots lt}) + \mathbb{N}T^{-1}Cov(\xi_{ij\dots l1}, \xi_{ij\dots lt})$.

In equation (I.13), it has been proved that, as $H \to 0$,

$$\begin{aligned} Var(\xi_{ij\dots lt}) &= \sigma_v^2 R^q(K)\mathcal{B}_{XX}(z)(1 + o_p(1)), \\ Cov(\xi_{ij\dots l1}, \xi_{ij\dots lt}) &= o_p(1). \end{aligned}$$

By Minkowski's inequality and Assumption 3.8, following a similar method to that above, it can be proved that, for some $\varepsilon > 0$, $E|\bar{\xi}_{ij\dots l}|^{(2+\varepsilon)} \leq CT^{(2+\varepsilon)/2}|H_z|^{-\varepsilon/2}$. Then, Assumption 3.7 yields

$$\mathbb{N}^{-\varepsilon/2} E|\xi_{ij\dots lt}|^{(2+\varepsilon)} \leq C(\mathbb{N}|H_z|)^{-\varepsilon/2} \to 0$$

so the Lyapounov's condition holds. Finally, using (I.8), (I.10), and (I.15) in (I.2), the proof of the theorem is done.

∎

Proof of Theorem 4.1: Let $\mathcal{V}_{\mathbb{N},ij\dots l} = \sum_{i'=1}^{i-1} \sum_{j'=1}^{j-1} \dots \sum_{l'=1}^{l-1} \mathcal{H}_{\mathbb{N}}(\chi_{ij\dots l}, \chi_{i'j'\dots l'})$, for $2 \leq i \leq N_1, 2 \leq j \leq N_2, \dots, 2 \leq l \leq N_l$. In order to prove this theorem, the following two conditions have to be checked. As any element in $\mathbb{N}$ tends to infinity, for each $\varepsilon > 0$,

a) $r_{\mathbb{N}}^{-2} \sum_{ij\dots l} E[\mathcal{V}_{\mathbb{N},ij\dots l}^2 \mathbb{1}(|\mathcal{V}_{\mathbb{N},ij\dots l}| > \varepsilon r_{\mathbb{N}})] \longrightarrow 0$,

b) $r_{\mathbb{N}}^{-2} \mathcal{V}_{\mathbb{N}}^2 \xrightarrow{p} 1$,

where $r_{\mathbb{N}}^2 = E(U_{\mathbb{N}}^2)$ and $\mathcal{V}_{\mathbb{N}}^2 = \sum_{i=2}^{N_1} \sum_{j=2}^{N_2} \dots \sum_{l=1}^{N_l} E\left(\mathcal{V}_{\mathbb{N},ij\dots l}^2 | \chi_{11\dots 1}, \dots, \chi_{(i-1)(j-1)\dots(l-1)}\right)$.

If these two conditions hold, we can resort to the CLT of Brown (1971) in the same way as Hall (1984) did for a one-sample degenerate U-statistics. Then, it is possible to conclude that $r_{\mathbb{N}}^{-1} U_{\mathbb{N}} \sim N(0, 1)$.

In order to check the condition a), we need to obtain the bounds of $E[\mathcal{V}_{\mathbb{N}}^4]$ and prove that $r_{\mathbb{N}}^{-4} \sum_{i=2}^{N_1} \sum_{j=2}^{N_2} \dots \sum_{l=1}^{N_l} E[\mathcal{V}_{\mathbb{N},ij\dots l}^4] \xrightarrow{p} 0$, as any element in $\mathbb{N}$ tends to infinity. Then, the

Lindeberg condition collected in (I) is obtained immediately from

$$\begin{aligned}\lim_{\mathbb{N}\to\infty} r_{\mathbb{N}}^{-2}\sum_{i=2}^{N_1}\sum_{j=2}^{N_2}\dots\sum_{l=2}^{N_l} E[\mathcal{V}_{\mathbb{N},ij\dots l}^2 \mathbb{1}(|\mathcal{V}_{\mathbb{N},ij\dots l}| > \varepsilon r_{\mathbb{N}})] \quad &\leq \quad \lim_{\mathbb{N}\to\infty} r_{\mathbb{N}}^{-2}\sum_{i=2}^{N_1}\sum_{j=2}^{N_2}\dots\sum_{l=2}^{N_l} E\left[\mathcal{V}_{\mathbb{N},ij\dots l}^2\left(\frac{\mathcal{V}_{\mathbb{N},ij\dots l}^2}{\varepsilon^2 r_{\mathbb{N}}^2}\right)\right] \\ &= \varepsilon^{-2}\lim_{\mathbb{N}\to\infty} r_{\mathbb{N}}^{-4}\sum_{i=2}^{N_1}\sum_{j=2}^{N_2}\dots\sum_{l=1}^{N_l} E[\mathcal{V}_{\mathbb{N},ij\dots l}^4] = 0.\end{aligned} \tag{I.16}$$

To analyze $r_{\mathbb{N}}^2 = E[U_{\mathbb{N}}^2]$, we use Assumption 3.1 and the LIE, obtaining

$$\begin{aligned} r_{\mathbb{N}}^2 &= \sum_{i=2}^{N_1}\sum_{j=2}^{N_2}\dots\sum_{l=2}^{N_l}\sum_{i'=1}^{i-1}\sum_{j'=1}^{j-1}\dots\sum_{l'=1}^{l-1} E[H_{\mathbb{N}}^2(\chi_{ij\dots l},\chi_{i'j'\dots l'})] \\ &+ \sum_{i=2}^{N_1}\sum_{j=2}^{N_2}\dots\sum_{l=2}^{N_l}\sum_{i'=1}^{i-1}\sum_{j'=1}^{j-1}\dots\sum_{l'=1}^{l-1}\sum_{i''\neq i'}^{i-1}\sum_{j''\neq j'}^{j-1}\dots\sum_{l''\neq l'}^{l-1} E[H_{\mathbb{N}}(\chi_{ij\dots l},\chi_{i'j'\dots l'})E[H_{\mathbb{N}}(\chi_{ij\dots l}\chi_{i''j''\dots l''}|\chi_{ij\dots l},\chi_{i'j'\dots l'})]] \\ &= \sum_{i=2}^{N_1}\sum_{j=2}^{N_2}\dots\sum_{l=2}^{N_l}(i-1)(j-1)\times\dots\times(l-1)E[H_{\mathbb{N}}^2(\chi_{11\dots 1},\chi_{22\dots 2})] \\ &= \frac{\mathbb{N}}{2^R}\prod_{r=1}^{l}(N_r-1)E[H_{\mathbb{N}}^2(\chi_{11\dots 1},\chi_{22\dots 2})],\end{aligned}$$

given that $E[H_{\mathbb{N}}(\chi_{ij\dots l},\chi_{i''j''\dots l''})|\chi_{ij\dots l},\chi_{i'j'\dots l'}] = 0$ by the degeneracy of the U-statistic,[7] so we can say

$$r_{\mathbb{N}}^4 = \frac{\mathbb{N}^2}{2^{2R}}\prod_{r=1}^{l}(N_r-1)^2 E[H_{\mathbb{N}}^2(\chi_{11\dots 1},\chi_{22\dots 2})]^2, \tag{I.17}$$

where R is the total number of cross-sectional samples.

If we consider the behavior of $E[\mathcal{V}_{\mathbb{N},ij\dots l}^4]$, it is possible to show that, based on the law of iterated expectations and the degeneracy of the U-statistic,

$$\begin{aligned} E[\mathcal{V}_{\mathbb{N},ij\dots l}^4] &= \sum_{i'=1}^{i-1}\sum_{j'=1}^{j-1}\dots\sum_{l'=1}^{l-1} E[H_{\mathbb{N}}^4(\chi_{ij\dots l},\chi_{i'j'\dots l'})] \\ &+ 3\sum_{i'=1}^{i-1}\sum_{j'=1}^{j-1}\dots\sum_{l'=1}^{l-1}\sum_{i''=1}^{i-1}\sum_{j''=1}^{j-1}\dots\sum_{l''=1}^{l-1} E[H_{\mathbb{N}}^2(\chi_{ij\dots l},\chi_{i'j'\dots l'})H_{\mathbb{N}}^2(\chi_{ij\dots l},\chi_{i''j''\dots l''})] \\ &= (i-1)(j-1)\times\dots\times(l-1)E[H_{\mathbb{N}}^4(\chi_{11\dots 1},\chi_{22\dots 2})] \\ &+ 3(i-1)(i-2)(j-1)(j-2)\times\dots\times(l-1)(l-2)E[H_{\mathbb{N}}^2(\chi_{11\dots 1},\chi_{22\dots 2})H_{\mathbb{N}}^2(\chi_{11\dots 1},\chi_{33\dots 3})],\end{aligned}$$

[7] $U_{\mathbb{N}}$ is a degenerate U-statistic given that, using Assumption 3.6, it can be shown that $H_{\mathbb{N}}(\chi_{ij\dots l},\chi_{i'j'\dots l'})$ is a symmetric function, $E[H_{\mathbb{N}}(\chi_{ij\dots l},\chi_{i'j'\dots l'})] = 0$, and $E[H_{\mathbb{N}}^2(\chi_{ij\dots l},\chi_{i'j'\dots l'})] \leq \infty$ as any element in $\mathbb{N}$ tends to infinity.

so it yields

$$\sum_{ij\ldots l} E[\mathcal{V}^4_{\mathbb{N},ij\ldots l}] \leq C\mathbb{N}^2 E[H^4_{\mathbb{N}}(\chi_{11},\chi_{22})] + C\mathbb{N}^3 E[H^2_{\mathbb{N}}(\chi_{11},\chi_{22})H^2_{\mathbb{N}}(\chi_{11},\chi_{33})]$$
$$\leq C\mathbb{N}^2(1+\mathbb{N})E[H^4_{\mathbb{N}}(\chi_{11\ldots 1},\chi_{22\ldots 2})], \quad \text{(I.18)}$$

where the Cauchy-Schwarz inequality was used on the second element of the right-hand side of the above expression.

Therefore, using (I.17)-(I.18) we can conclude that, as any element in $\mathbb{N}$ tends to infinity,

$$r_{\mathbb{N}}^{-2}\sum_{i=2}^{N_1}\sum_{j=2}^{N_2}\cdots\sum_{l=2}^{N_l} E[\mathcal{V}^2_{\mathbb{N},ij\ldots l}\mathbb{1}(|\mathcal{V}_{\mathbb{N},ij\ldots l}| > \varepsilon r_{\mathbb{N}})] \leq \frac{CE[H^4_{\mathbb{N}}(\chi_{11\ldots 1},\chi_{22\ldots 2})]}{\mathbb{N}E[H^2_{\mathbb{N}}(\chi_{11\ldots 1},\chi_{22\ldots 2})]^2} \xrightarrow{p} 0, \quad \text{(I.19)}$$

so it is proved that condition a) holds.

Considering now the condition b), we need to prove the convergence in squared mean of $r_{\mathbb{N}}^{-2}\mathcal{V}^2_{\mathbb{N}}$ to 1, which implies that $r_{\mathbb{N}}^{-2}\mathcal{V}_{\mathbb{N}} \xrightarrow{p} 1$, by obtaining bounds for $E[\mathcal{V}^4_{\mathbb{N}}]$. Let us denote $G_{\mathbb{N}}(x,y) = E[H_{\mathbb{N}}(\chi_{11}\ldots 1, x)H_n(\chi_{11\ldots 1}, y)]$ and $V^2_{\mathbb{N}} = \sum_{i=2}^{N_1}\sum_{j=2}^{N_2}\cdots\sum_{l=2}^{N_l}\upsilon_{\mathbb{N},ij\ldots l}$, where

$$\upsilon_{\mathbb{N},ij\ldots l} = \sum_{i'=1}^{i-1}\sum_{j'=1}^{j-1}\cdots\sum_{l'=1}^{l-1} G_{\mathbb{N}}(\chi_{i'j'\ldots l'},\chi_{i'j'\ldots l'}) + 2\sum\sum_{2\leq i'<i''\leq i-1}\sum\sum_{2\leq j'<j''\leq j-1}\cdots\sum\sum_{2\leq l'<l''\leq l-1} G_{\mathbb{N}}(\chi_{i'j'\ldots l'},\chi_{ii'j''\ldots l''}).$$

If $(i_1 \leq i_2)$, $(j_1 \leq j_2)$ and so on,

$$\begin{aligned} E[\upsilon_{\mathbb{N},i_1j_1}\upsilon_{\mathbb{N},i_2j_2}] &= \sum_{i_1'=1}^{i_1-1}\sum_{j_1'=1}^{j_1-1}\cdots\sum_{l_1'=1}^{l_1-1} E[G^2_{\mathbb{N}}(\chi_{11\ldots 1},\chi_{11\ldots 1})] + \sum_{i_1'=1}^{i_1-1}\sum_{i_2'>i_1'}^{i_2-1}\sum_{j_1'=1}^{j_1-1}\sum_{j_2'>j_1'}^{j_2-1}\cdots\sum_{l_1'=1}^{l_1-1}\sum_{l_2'>l_1'}^{l_2-1}\{E[G_{\mathbb{N}}(\chi_{11\ldots 1},\chi_{11\ldots 1})]\}^2 \\ &+ 4\sum\sum_{2\leq i_1'<i_1''\leq(i_1-1)}\sum\sum_{2\leq j_1'<j_1''\leq(j_1-1)}\cdots\sum\sum_{2\leq l_1'<l_1''\leq(l_1-1)} E[G^2_{\mathbb{N}}(\chi_{11\ldots 1},\chi_{22\ldots 2})] \end{aligned}$$

given that, if $(i_1' \leq i_1'', j_1' \leq j_1'', \ldots, l_1' \leq l_1'')$, $(i_2' \leq i_2'', j_2' \leq j_2'', \ldots, l_2' \leq l_2'')$,

$$E[G_{\mathbb{N}}(\chi_{i_1'j_1'\ldots l_1'},\chi_{i_1''j_1''\ldots l_1''})G_{\mathbb{N}}(\chi_{i_2'j_2'\ldots l_2'},\chi_{i_2''j_2''\ldots l_2''})]$$
$$= \begin{cases} E[G^2_{\mathbb{N}}(\chi_{11\ldots 1},\chi_{11\ldots 1})] & \text{if } i_1'=i_1'', j_1'=j_1'',\ldots,l_1'=l_1'', i_2'=i_2'', j_2'=j_2'',\ldots l_2'=l_2'', i_1'=i_2', j_1'=j_2',\ldots l_1'=l_2' \\ \{E[G_{\mathbb{N}}(\chi_{11\ldots 1},\chi_{11\ldots 1})]\}^2 & \text{if } i_1'=i_1'', j_1'=j_1'',\ldots,l_1'=l_1'', i_2'=i_2'', j_2'=j_2'',\ldots,l_2'=l_2'', i_1'\neq i_2', j_1'\neq j_2',\ldots l_1'\neq l_2' \\ E[G^2_{\mathbb{N}}(\chi_{11\ldots 1},\chi_{22\ldots 2})] & \text{if } i_1'=i_1'', j_1'=j_1'',\ldots l_1'=l_1'', i_2'=i_2'', j_2'=j_2'',\ldots,l_2'=l_2'', i_1'<i_2', j_1'<j_2',\ldots,l_1'<l_2' \\ 0 & \text{otherwise.} \end{cases}$$

Using these results and rearranging terms, we get

$$\begin{aligned} E[\mathcal{V}^4_{\mathbb{N}}] &\leq C\mathbb{N}^3 Var[G_{\mathbb{N}}(\chi_{11\ldots 1},\chi_{11\ldots 1})] + C\mathbb{N}^4\{G_{\mathbb{N}}(\chi_{11\ldots 1},\chi_{22\ldots 2})\}^2 + C\mathbb{N}^4 E[G^2_{\mathbb{N}}(\chi_{11\ldots 1},\chi_{22\ldots 2})] \\ &\leq C\mathbb{N}^3 E[G^2_{\mathbb{N}}(\chi_{11\ldots 1},\chi_{11\ldots 1})] + C\mathbb{N}^4 E[\chi^2_{\mathbb{N}}(\chi_{11\ldots 1},\chi_{22\ldots 2})] + C\mathbb{N}^3(\mathbb{N}-1)\{E[G_{\mathbb{N}}(\chi_{11\ldots 1},\chi_{11\ldots 1})]\}^2, \end{aligned}$$

and $E[\mathcal{V}_{\mathbb{N}}^2 - r_{\mathbb{N}}^2] \leq C\mathbb{N}^4 E[G_{\mathbb{N}}^2(\chi_{11...1}, \chi_{22...2})] + C\mathbb{N}^3 E[H_{\mathbb{N}}^4(\chi_{11...1}, \chi_{22...2})]$, where we have used the fact that $Var[G_{\mathbb{N}}(\chi_{11...1}, \chi_{11...1})] = E[G_{\mathbb{N}}^2(\chi_{11...1}, \chi_{11...1})] - \{E[G_{\mathbb{N}}(\chi_{11...1}, \chi_{11...1})]\}^2$ and $E[G_{\mathbb{N}}^2(\chi_{11...1}, \chi_{11...1})] = E[H_{\mathbb{N}}^4(\chi_{11...1}, \chi_{22...2})]$.

Using the bound of $E[\mathcal{V}_{\mathbb{N}}^4]$ and $E[\mathcal{V}_{\mathbb{N}}^2] = r_{\mathbb{N}}^2$, it results that, as $\mathbb{N}_{\max}\infty$,

$$\begin{aligned} E[(r_{\mathbb{N}}^{-2}\mathcal{V}_{\mathbb{N}}^2 - 1)^2] &= r_{\mathbb{N}}^{-4} E[(\mathcal{V}_{\mathbb{N}}^2 - r_{\mathbb{N}}^2)^2] \\ &\leq \frac{CE[G_{\mathbb{N}}^2(\chi_{11...1}, \chi_{22...2})] + \mathbb{N}^{-1}E[H_{\mathbb{N}}^4(\chi_{11...1}, \chi_{22...2})]}{\{E[H_{\mathbb{N}}^2(\chi_{11...1}, \chi_{22...2})]\}^2} \to 0. \end{aligned} \tag{I.20}$$

Then, $r_{\mathbb{N}}^{-2}\mathcal{V}_{\mathbb{N}}^2$ converges to 1 in squared mean, which implies $r_{\mathbb{N}}^{-2}\mathcal{V}_{\mathbb{N}}^2 \xrightarrow{d} 1$ and proves the condition b).

∎

Proof of Theorem 5.1: Let $\widetilde{W}_{ij...lt} = [\widetilde{X}_{ij...lt}, \widetilde{Z}_{ij...lt}]$ be a $(d+q) \times 1$ vector. Under H_0^a, the transformed residuals can be written as $\widehat{\widetilde{v}}_{ij...lt} = \widetilde{v}_{ij...lt} - \widetilde{W}_{ij...lt}^{\top}(\widehat{\beta}_0 - \beta)$. The proposed test statistic can then be written as

$$\begin{aligned} \widehat{I}_{\mathbb{N}}^a &= \frac{1}{\mathbb{N}^2 T^2} \sum_{ij...lt} \sum_{i'j'...l't'} \widetilde{v}_{ij...lt}\widetilde{v}_{i'j'...l't'} K_{H_x}(X_{ij...lt}, X_{i'j'...l't'}) K_{H_z}(Z_{ij...lt}, Z_{i'j'...l't'}) \\ &- \frac{2}{\mathbb{N}^2 T^2} \sum_{ij...lt} \sum_{i'j'...l't'} \widetilde{v}_{ij...lt}\widetilde{W}_{i'j'...l't'}^{\top}(\widehat{\beta}_0 - \beta) K_{H_x}(X_{ij...lt}, X_{i'j'...l't'}) K_{H_z}(Z_{ij...lt}, Z_{i'j'...l't'}) \\ &+ \frac{1}{\mathbb{N}^2 T^2} \sum_{ij...lt} \sum_{i'j'...l't'} (\widehat{\beta}_0 - \beta)^{\top}\widetilde{W}_{ij...lt}\widetilde{W}_{i'j'...l't'}^{\top}(\widehat{\beta}_0 - \beta) K_{H_z}(X_{ij...lt}, X_{i'j'...l't'}) K_{H_z}(Z_{ij...lt}, Z_{i'j'...l't'}) \\ &= \mathbb{I}_{\mathbb{N}_1}^a - 2\mathbb{I}_{\mathbb{N}_2}^a + \mathbb{I}_{\mathbb{N}_3}^a, \end{aligned} \tag{I.21}$$

where the definitions of $\mathbb{I}_{\mathbb{N}_s}^a$, $s = 1, 2, 3$ should be apparent from the context. Using $\widehat{\beta}_0 - \beta = O_p(\mathbb{N}^{-1/2})$, it is easy to show that, under H_0^a, $\mathbb{I}_{\mathbb{N}_2}^a = O_p(\mathbb{N}^{-1})$ and $\mathbb{I}_{\mathbb{N}_3}^a = O_p(\mathbb{N}^{-1})$. Hence, $\mathbb{I}_{\mathbb{N}_1}^a$ is the leading term of $\widehat{I}_{\mathbb{N}}^a$ under H_0^a.

In order to obtain the asymptotic distribution of $\widehat{I}_{\mathbb{N}}^a$ under H_0^a, we write (I.21) as a general second-order U-statistic of the form

$$I_{\mathbb{N}_1}^a = \prod_{r=1}^{R} \binom{N_r}{2}^{-1} \sum_{i=1}^{N_1-1} \sum_{i'=1+i}^{N_1} \sum_{j=1}^{N_2-1} \sum_{j'=1+j}^{N_2} \cdots \sum_{l=1}^{N_l-1} \sum_{l'=1+l}^{N_l} H_{\mathbb{N}}^a(\chi_{ij...l}, \chi_{i'j'...l'})|H_x|^{-1}|H_z|^{-1}, \tag{I.22}$$

where now $\chi_{ij...l} = (X_{ij...l}, Z_{ij...l})$ and $H_{\mathbb{N}}^a(\chi_{ij...l}, \chi_{i'j'...l'})$ is a symmetric, centered and degenerate variable of the form

$$H_{\mathbb{N}}^a(\chi_{ij...l}, \chi_{i'j'...l'}) = \binom{T}{2}^{-1} \sum_{t=1}^{T-1} \sum_{t'=1+t}^{T} \widetilde{v}_{ij...lt}\widetilde{v}_{i'j'...l't'} K(X_{ij...lt}, X_{i'j'...l't'}) K(Z_{ij...lt}, Z_{i'j'...l't'}).$$

Under Assumptions 3.1 and 3.2, $\mathbb{N}|H_x|^{1/2}|H_z|^{1/2}\widehat{I}^a_{\mathbb{N}} = \mathbb{N}|H_x|^{1/2}|H_z|^{1/2}\mathbb{I}^a_{\mathbb{N}_1} + O_p(|H_z|^{1/2}|H_x|^{1/2})$ given that it can be shown that $\mathbb{I}^a_{\mathbb{N}_1}$ has zero mean and asymptotic variance given by

$$\begin{aligned} Var(\mathbb{I}^a_{\mathbb{N}_1}) &= \prod_{r=1}^{R} \frac{2^{(R+1)}}{\mathbb{N}T(N_r-1)(T-1)|H_x|^2|H_z|^2} E[\widetilde{v}^2_{ij\ldots lt}\widetilde{v}^2_{i'j'\ldots l't'}K^2(X_{ij\ldots lt}, X_{i'j'\ldots l't'})K^2(Z_{ij\ldots lt}, Z_{i'j'\ldots l't'})] + (s.o.) \\ &= \prod_{r=1}^{R} \frac{2^{(R+1)}(T-1)R^{2q}(K)\sigma_v^4}{\mathbb{N}(N_r-1)T^3|H_x||H_z|} \int f^2(X_{ij\ldots lt}, Z_{ij\ldots lt})dX_{ij\ldots lt}dZ_{ij\ldots lt} \int K^2(u_1)K^2(u_2)du_1du_2 \\ &\quad \times(1+o_p(1)) \\ &= \prod_{r=1}^{R} \frac{2^{(R+1)}(T-1)R^{(q+d)}(K)\sigma_v^4}{\mathbb{N}(N_r-1)T^3|H_x||H_z|} E[f(X_{11\ldots 1t}, Z_{11\ldots 1t})](1+o_p(1)). \end{aligned}$$

Let $G^a_{\mathbb{N}}(\chi_{11\ldots 1}, \chi_{22\ldots 2}) = E[H^a_{\mathbb{N}}(\chi_{11\ldots 1}, \chi_{ij\ldots l})H^a_{\mathbb{N}}(\chi_{22\ldots 2}, \chi_{ij\ldots l})]$. Straightforward calculations yield

$$\frac{E[\{G^a_{\mathbb{N}}(\chi_{11\ldots 1}, \chi_{22\ldots 2})\}^2] + \mathbb{N}^{-1}E[\{H^a_{\mathbb{N}}(\chi_{11\ldots 1}, \chi_{22\ldots 2})\}^4]}{\{E[\{H^a_{\mathbb{N}}(\chi_{11\ldots 1}, \chi_{22\ldots 2})\}^2]\}^2} = \frac{O_p(|H_x|^3|H_z|^3) + O_p(\mathbb{N}^{-1}|H_x||H_z|)}{O_p(|H_x|^2|H_z|^2)} \to 0$$

provided that $|H_x| \to 0$, $|H_z| \to 0$, and $\mathbb{N}|H_x||H_z| \to \infty$ as $\mathbb{N}_{\max} \to \infty$. Applying the CLT in Theorem 5.1, we can conclude

$$\mathbb{N}_{\max}|H_x|^{1/2}|H_z|^{1/2}\widehat{I}^a_{\mathbb{N}} \xrightarrow{d} N(0, \Sigma_a), \tag{I.23}$$

where $\Sigma_a = \frac{2^{(R+1)}(T-1)R^{(q+d)}(K)\sigma_v^4}{T^3|H_x||H_z|}E[f(X_{11\ldots 1t}, Z_{11\ldots 1t})]$.

Following a similar reasoning as in (I.21), it can be shown that the leading term of $\widehat{\Sigma}_a$ under H^a_0 is given by

$$\Sigma^*_a = \prod_{r=1}^{R} \frac{4}{\mathbb{N}(N_r-1)} \sum_{i=1}^{N_1-1}\sum_{i'=1+i}^{N_1}\sum_{j=1}^{N_2-1}\sum_{j'=1+j}^{N_2} \cdots \sum_{l=1}^{N_l-1}\sum_{l'=1+l}^{N_l} H^a_{\mathbb{N}}(\chi_{ij\ldots l}, \chi_{i'j'\ldots l'}), \tag{I.24}$$

with $H^a_{\mathbb{N}}(\chi_{ij\ldots l}, \chi_{i'j'\ldots l'})$ being a second-order U-statistic of the form

$$H^a_{\mathbb{N}}(\chi_{ij\ldots l}, \chi_{i'j'\ldots l'}) = \binom{T}{2}^{-1} \sum_{t=1}^{T-1}\sum_{t'=1+t}^{T} \widetilde{v}^2_{ij\ldots lt}\widetilde{v}^2_{i'j'\ldots l't'}K^2(X_{ij\ldots lt}, X_{i'j'\ldots l't'})K_H(Z_{ij\ldots lt}, Z_{i'j'\ldots l't'})|H_x|^{-1}|H_z|^{-1}.$$

In order to establish the asymptotic equivalence of $\widehat{\Sigma}_a$ to Σ_a, we need to check the limiting behavior of $E[\|H^a_{\mathbb{N}}(\chi_{ij\ldots l}, \chi_{i'j'\ldots l'})\|^2]$. Using the law of iterated expectations, Assumption 3.2, and the strict stationarity assumption, we get

$$\begin{aligned}
&E[\|H^a_{\mathbb{N}}(\chi_{ij\ldots l},\chi_{i'j'\ldots l'})\|^2]\\
&=\binom{T}{2}^{-2}\frac{1}{|H_x|^2|H_z|^2}\sum_{t=1}^{T-1}\sum_{t'=1+t}^{T}\sum_{s=1}^{T-1}\sum_{s'=1+s}^{T}E[E(\widetilde{v}^2_{ij\ldots lt}\widetilde{v}^2_{i'j'\ldots l't'}\widetilde{v}^2_{ij\ldots ls}\widetilde{v}^2_{i'j'\ldots l's'})K^2(X_{ij\ldots lt},X_{i'j'\ldots l't'})\\
&\times K^2(Z_{ij\ldots lt},Z_{i'j'\ldots l't'})K^2(X_{ij\ldots ls},X_{i'j'\ldots l's'})K^2(Z_{ij\ldots ls},Z_{ij\ldots l's'})]\\
&=\frac{2\{E(\widetilde{v}^4_{ij\ldots lt})\}^2}{T(T-1)|H_x||H_z|}\int K^4(u_1)K^4(u_2)du_1du_2\int f(X_{i'j'\ldots l't'},Z_{i'j'\ldots l't'})dX_{i'j'\ldots l't'}dZ_{i'j'\ldots l't'}\\
&\times(1+o_p(1))\\
&=O_p(|H_x|^{-1}|H_z|^{-1})=o_p(\mathbb{N}), \qquad \text{(I.25)}
\end{aligned}$$

with a finite T as $E(v^4_{ij\ldots lt})=\tau_4<\infty$. Using this result and following a similar reasoning as in Lemma 3.1 of Powell et al. (1989), we can conclude that $\sqrt{\mathbb{N}_{\max}}(\widehat{\Sigma}_a-\Sigma_a)=o_p(1)$. Hence, it is easy to show that $E(\Sigma^*_a)=\Sigma_a+o_p(1)$, which completes the first part of the proof.

Focusing now on the second part of the theorem, the transformed residuals under the H_1 can be written as $\widehat{\widetilde{v}}_{ij\ldots lt}=\widetilde{v}_{ij\ldots lt}+\widetilde{\beta}(X_{ij\ldots lt},Z_{ij\ldots lt})-\widetilde{W}^{\top}_{ij\ldots lt}\widehat{\beta}$, where $\widetilde{\beta}(X_{ijt},Z_{ijt})$ is the within transformed unknown function. Assumptions 3.1 and 3.2 and 5.1 and 5.2 ensure $\widehat{\beta}=\overline{\beta}+O_p(\mathbb{N}^{-1/2})$, where $\overline{\beta}$ is the probability limit of $\widehat{\beta}$. We can replace $\widehat{\beta}$ by $\overline{\beta}$ in analyzing the behavior of $\widehat{I}^a_{\mathbb{N}}$ under H^a_1. Using the fact that $v_{ij\ldots lt}$ is independent of $(X_{ij\ldots lt},Z_{ij\ldots lt})$, the expression to analyze is

$$\begin{aligned}
\widehat{I}^a_{\mathbb{N}} &= \frac{1}{\mathbb{N}^2T^2|H_x||H_z|}\sum_{ij\ldots lt}\sum_{i'j'\ldots l't'}\widetilde{v}_{ij\ldots lt}\widetilde{v}_{i'j'\ldots l't'}K(X_{ij\ldots lt},X_{i'j'\ldots l't'})K(Z_{ij\ldots lt},Z_{i'j'\ldots l't'})\\
&+\frac{1}{\mathbb{N}^2T^2|H_x||H_z|}\sum_{ij\ldots lt}\sum_{i'j'\ldots l't'}[\widetilde{\beta}(X_{ij\ldots lt},Z_{ij\ldots lt})-\widetilde{W}^{\top}_{ij\ldots lt}\overline{\beta}]\\
&\quad\times[\widetilde{\beta}(X_{i'j'\ldots l't'},Z_{i'j'\ldots l't'})-\widetilde{W}^{\top}_{i'j'\ldots l't'}\overline{\beta}]K(X_{ij\ldots lt},X_{i'j'\ldots l't'})K(Z_{ij\ldots lt},Z_{i'j'\ldots l't'})+O_p\left(\frac{1}{\sqrt{\mathbb{N}}}\right)\\
&= I\!\!I^a_{\mathbb{N}_1}+\widetilde{I\!\!I}^a_{\mathbb{N}_2}+O_p\left(\frac{1}{\sqrt{\mathbb{N}}}\right), \qquad \text{(I.26)}
\end{aligned}$$

where the definitions of $I\!\!I^a_{\mathbb{N}_1}$ and $\widetilde{I\!\!I}^a_{\mathbb{N}_1}$ should be apparent from the context.

$I\!\!I^a_{\mathbb{N}_1}$ has been analyzed previously, so we can conclude that $I\!\!I^a_{\mathbb{N}_1}=O_p((\mathbb{N}|H_x|^{1/2}|H_z|^{1/2})^{-1})$. Considering now $\widehat{I\!\!I}^b_{\mathbb{N}_2}$ and following a similar reasoning as in (I.25), it is easy to show that $E[\|\widetilde{H}^a_{\mathbb{N}}(\chi_{ij\ldots l},\chi_{i'j'\ldots l'})\|^2]=O_p(|H_x|^{-1}|H_z|^{-1})=o_p(\mathbb{N})$. Then, by Lemma 3.1 of Powell et al. (1989), we obtain $\widetilde{I\!\!I}^a_{\mathbb{N}_2}=E[\mathcal{H}^a_{\mathbb{N}}(\chi_{ij\ldots l},\chi_{i'j'\ldots l'})]+o_p(1)$, where we denote $\widetilde{\psi}^a_{ij\ldots lt}=$

$\widetilde{\beta}(X_{ij\ldots lt}, Z_{ij\ldots lt}) - \widetilde{W}_{ij\ldots lt}^{\top}\overline{\beta}$ and

$$\widetilde{H}_{\mathbb{N}}^{a}(\chi_{ij\ldots l}, \chi_{i'j'\ldots l'}) = \binom{T}{2}^{-1} \sum_{t=1}^{T-1} \sum_{t'=1+t}^{T} \widetilde{\psi}_{ij\ldots lt}^{a} \widetilde{\psi}_{i'j'\ldots l't'}^{a} K(X_{ij\ldots lt}, X_{i'j'\ldots l't'}) K(Z_{ij\ldots lt}, Z_{i'j'\ldots l't'}) |H_x|^{-1} |H_z|^{-1}.$$

Under Assumption 3.1, $(X_{ij\ldots lt}, Z_{ij\ldots lt})$ are independent of $(X_{i'j'\ldots l't'}, Z_{i'j'\ldots l't'})$, for $i \neq i', j \neq j', \ldots, l' \neq l'$. Then, by the law of iterated expectations and the strict stationarity condition,

$$\begin{aligned} E[\mathcal{H}_{\mathbb{N}}^{a}(\chi_{ij\ldots l}, \chi_{i'j'\ldots l'})] &= \binom{T}{2}^{-1} \frac{1}{|H_x||H_z|} \sum_{t=1}^{T-1} \sum_{t'=1+t}^{T} \int E[\widetilde{\psi}_{ij\ldots lt}^{a} | X_{ij\ldots lt}, Z_{ij\ldots lt}] E[\widetilde{\psi}_{i'j'\ldots l't'}^{a} | X_{i'j'\ldots l't'}, Z_{i'j'\ldots l't'}] \\ &\quad \times K(X_{ij\ldots lt}, X_{i'j'\ldots l't'}) K(Z_{ij\ldots lt}, Z_{i'j'\ldots l't'}) f(X_{ij\ldots lt}, Z_{ij\ldots lt}) f(X_{i'j'\ldots l't'}, Z_{i'j'\ldots l't'}) \\ &\quad \times dX_{ij\ldots lt} dZ_{ij\ldots lt} dX_{i'j'\ldots l't'} dZ_{i'j'\ldots l't'} \\ &= E\left[\{M(X_{ij\ldots lt}, Z_{ij\ldots lt})\}^2 f(X_{ij\ldots lt}, Z_{ij\ldots lt})\right], \end{aligned} \quad \text{(I.27)}$$

where $M(X_{ij\ldots lt}, Z_{ij\ldots lt}) = E[\widetilde{\psi}_{ij\ldots lt}^{a} | X_{ij\ldots lt}, Z_{ij\ldots lt}]$. It follows that

$$\widetilde{\mathbb{I}}_{n_2}^{a} = E[\{M(X_{ij\ldots lt}, Z_{ij\ldots lt})\}^2 f(X_{ij\ldots lt}, Z_{ij\ldots lt})] + o_p(1),$$

given that it is assumed that $E[\{M(X_{ij\ldots lt}, Z_{ij\ldots lt})\}^2 f(X_{ij\ldots lt}, Z_{ij\ldots lt})]$ is a positive constant.

Finally, we have to show that $\widehat{\Sigma}_a$ converges to a positive constant under H_1^a. With this aim, $\widehat{\Sigma}_a$ can be written as

$$\widehat{\Sigma}_a = \prod_{r=1}^{R} \binom{N_r}{2}^{-1} \sum_{i=1}^{N_1-1} \sum_{i'=1+i}^{N_1} \sum_{j=1}^{N_2-1} \sum_{j'=1+j}^{N_2} \cdots \sum_{l=1}^{N_l-1} \sum_{l'=1+l}^{N_l} \mathcal{B}_{\mathbb{N}}^{a}(\chi_{ij\ldots l}, \chi_{i'j'\ldots l'}) + o_p(1), \quad \text{(I.28)}$$

where $\mathcal{B}_{\mathbb{N}}^{a}(\chi_{ij\ldots l}, \chi_{i'j'\ldots l'})$ is a second-order U-statistic defined as

$$\begin{aligned} \mathcal{B}_{\mathbb{N}}^{a}(\chi_{ij\ldots l}, \chi_{i'j'\ldots l'}) &= \binom{T}{2}^{-1} \frac{1}{|H_x||H_z|} \sum_{t=1}^{T-1} \sum_{t'=1+t}^{T} (\widetilde{v}_{ij\ldots lt} + \widetilde{\psi}_{ij\ldots lt}^{a})^2 (\widetilde{v}_{i'j'\ldots l't'} + \widetilde{\psi}_{i'j'\ldots l't'}^{a})^2 K^2(X_{ij\ldots lt}, X_{i'j'\ldots l't'}) \\ &\quad \times K^2(Z_{ij\ldots lt}, Z_{i'j'\ldots l't'}). \end{aligned}$$

Following a similar procedure as above, it can be shown that $E[\|\mathcal{B}_{\mathbb{N}}^{a}(\chi_{ij\ldots l}, \chi_{i'j'\ldots l'})\|^2] = O_p(|H_x||H_z|) = o_p(\mathbb{N})$. Using Lemma 3.1 of Powell et al. (1989), we can conclude that $\widehat{\Sigma}_a \xrightarrow{p} \Sigma_a$, given that it can be shown that $\mathcal{B}_{\mathbb{N}}^{a}(\chi_{ij\ldots l}, \chi_{i'j'\ldots l'})$ converges in probability to a finite positive number as any element in $\mathbb{N}$ tends to infinity (as does $\widehat{\Sigma}_a$). Finally, combining the above results, we obtain that, under H_1^a, the standardized test statistic $\widehat{J}_{\mathbb{N}}^{a} = \mathbb{N}|H_x|^{1/2}|H_z|^{1/2}\widehat{I}_{\mathbb{N}}^{a}/\sqrt{\widehat{\Sigma}_a}$ diverges to $+\infty$ at the rate $\mathbb{N}|H_x|^{1/2}|H_z|^{1/2}$. This completes the proof of the Theorem.

∎

Proof of Theorem 5.2: The proof of the result for $\widehat{I}^b_{\mathbb{N}}$ under H^b_0 is similar to Theorem 5.1. $\widehat{I}^b_{\mathbb{N}}$ can be written as a second-order U-statistic of the form

$$\widehat{I}^b_{\mathbb{N}} = \prod_{r=1}^{R}\binom{N_r}{2}^{-1}\sum_{i=1}^{N_1-1}\sum_{i'=1+i}^{N_1}\sum_{j=1}^{N_2-1}\sum_{j'=1+j}^{N_2}\cdots\sum_{l=1}^{N_l-1}\sum_{l'=1+l}^{N_l} H^b_{\mathbb{N}}(\chi_{ij\ldots l},\chi_{i'j'\ldots l'})|H_z|^{-1} + O_p\left(\frac{1}{\mathbb{N}}\right), \quad (I.29)$$

where $H^b_{\mathbb{N}}(\chi_{ij\ldots l},\chi_{i'j'\ldots l'}) = \binom{T}{2}^{-1}\sum_{t=1}^{T-1}\sum_{t'=1+t}^{T}\widetilde{v}_{ij\ldots lt}\widetilde{v}_{i'j'\ldots l't'}X^\top_{ij\ldots lt}X_{i'j'\ldots l't'}K(Z_{ij\ldots lt},Z_{i'j'\ldots l't'})$. Let $G^b_{\mathbb{N}}(\chi_{11\ldots 1},\chi_{22\ldots 2}) = E[H^b_{\mathbb{N}}(\chi_{11\ldots 1},\chi_{ij\ldots l})H^b_{\mathbb{N}}(\chi_{22\ldots 2},\chi_{ij\ldots l})]$. It can be shown that, as any element in $\mathbb{N}$ tends to infinity,

$$\frac{E[\{G^b_{\mathbb{N}}(\chi_{11\ldots 1},\chi_{22\ldots 2})\}^2] + \mathbb{N}^{-1}E[\{H^b_{\mathbb{N}}(\chi_{11\ldots 1},\chi_{11\ldots 1})\}^4]}{\{H^b_{\mathbb{N}}(\chi_{11\ldots 1},\chi_{22\ldots 2})\}^2} = \frac{O_p(|H_z|^3) + O_p(\mathbb{N}^{-1}|H_z|)}{O_p(|H_z|^2)} \to 0. \quad (I.30)$$

The results in Theorem 4.1 can be used to obtain the asymptotic distribution of the test statistic $\widehat{I}^b_{\mathbb{N}}$ under H^b_0 following a similar reasoning as in the proof of Theorem 5.1. For sake of brevity, it has been omitted.

Focusing now on the proof of the second part of the theorem, the transformed residuals under H^b_1 are written as $\widehat{\widetilde{v}}_{ij\ldots lt} = \widetilde{v}_{ij\ldots lt} + \widetilde{\beta}(X_{ij\ldots lt},Z_{ij\ldots lt}) - \widetilde{W}^\top_{ij\ldots lt}\widehat{\beta}$, where now $\widetilde{\beta}(X_{ij\ldots lt},Z_{ij\ldots lt})$ is the corresponding within transformation of $X^\top_{ij\ldots lt}\beta(Z_{ij\ldots lt})$ and $\widetilde{\pi}^b_{ij\ldots lt} = \widetilde{\beta}(X_{ij\ldots lt},Z_{ij\ldots lt}) - \widetilde{W}^\top_{ij\ldots lt}\overline{\beta}$. Following a similar reasoning as in (I.26), the expression to analyze is

$$\begin{aligned}\widehat{I}^b_{\mathbb{N}} &= \frac{1}{\mathbb{N}^2T(T-1)|H_z|}\sum_{ij\ldots lt}\sum_{i'j'\ldots l't'}\widetilde{v}_{ij\ldots lt}\widetilde{v}_{i'j'\ldots l't'}X^\top_{ij\ldots lt}X_{ij\ldots lt}K(Z_{ij\ldots lt},Z_{i'j'\ldots l't'}) \\ &+ \frac{1}{\mathbb{N}^2T(T-1)|H_z|}\sum_{ij\ldots lt}\sum_{i'j'\ldots l't'}\widetilde{\psi}^b_{ij\ldots lt}\widetilde{\psi}^b_{i'j'\ldots l't'}X^\top_{ij\ldots lt}X_{i'j'\ldots l't'}K(Z_{ij\ldots lt},Z_{i'j'\ldots l't'}) + O_p\left(\frac{1}{\sqrt{\mathbb{N}}}\right) \\ &= I\!\!I^b_{\mathbb{N}_1} + I\!\!I^b_{\mathbb{N}_2} + O_p\left(\frac{1}{\mathbb{N}}\right), \quad (I.31)\end{aligned}$$

where the definitions of $I\!\!I^b_{\mathbb{N}_1}$ and $I\!\!I^b_{\mathbb{N}_2}$ should be apparent.

Under Assumption 3.3, it can be shown that $I\!\!I^b_{\mathbb{N}_1}$ has zero mean and variance

$$\begin{aligned}Var(I\!\!I^b_{n_1}) &= \prod_{r=1}^{R}\frac{2^{(R+1)}}{\mathbb{N}T(N_l-1)(T-1)|H_z|^2}E[\widetilde{v}^2_{ij\ldots lt}\widetilde{v}^2_{i'j'\ldots l't'}X^\top_{ij\ldots lt}X_{i'j'\ldots l't'}X^\top_{i'j'\ldots l't'}X_{ij\ldots lt}K^2(Z_{ij\ldots lt},Z_{i'j'\ldots l't'})] \\ &= \prod_{r=1}^{R}\frac{2^{(R+1)}(T-1)\sigma^4_v}{\mathbb{N}T^3(N_r-1)|H_z|}\int E(tr\{X_{ij\ldots lt}X^\top_{i'j'\ldots l't'}\}^2|Z_{ij\ldots lt})f(Z_{ij\ldots lt})f(Z_{ij\ldots lt}+H_zu)dZ_{ij\ldots lt} \\ &\quad\times\int K^2(u)du \\ &= \prod_{r=1}^{R}\frac{2^{(R+1)}(T-1)\sigma^4_uR^q(K)}{\mathbb{N}T^3(N_r-1)|H_z|}E[E(tr\{X_{ij\ldots lt}X^\top_{i'j'\ldots l't'}\}^2|Z_{ij\ldots lt})f(Z_{ij\ldots lt})](1+o_p(1)).\end{aligned}$$

Following a similar procedure as in (I.27), it can be shown that $I\!I^b_{\mathbb{N}_2}$ converges in probability to a finite number as $\mathbb{N}_{\max} \to \infty$ under H^b_1. To complete the proof, it is necessary to show that $\widehat{\Sigma}_b$ converges to a positive constant under H^b_1. With this aim, $\widehat{\Sigma}_b$ can be rewritten as

$$\widehat{\Sigma}_b = \prod_{r=1}^{R} \binom{N_r}{2}^{-1} \sum_{i=1}^{N_1-1} \sum_{i'=1+i}^{N_1} \sum_{j=1}^{N_2-1} \sum_{j'=1+j}^{N_2} \cdots \sum_{l=1}^{N_l-1} \sum_{l'=1+l}^{N_l} \mathcal{B}^b_{\mathbb{N}}(\chi_{ij\ldots l}, \chi_{i'j'\ldots l'}),$$

where

$$\begin{aligned}\mathcal{B}^b_{\mathbb{N}}(\chi_{ij\ldots l}, \chi_{i'j'\ldots l'}) &= \binom{T}{2}^{-1} \frac{1}{|H_z|} \sum_{t=1}^{T-1} \sum_{t'=1+t}^{T} (\widetilde{v}_{ij\ldots lt} + \widetilde{\psi}^b_{ij\ldots lt})^2 (\widetilde{v}_{i'j'\ldots l't'} + \widetilde{\psi}^b_{i'j'\ldots l't'})^2 tr\{X_{ij\ldots lt} X^\top_{i'j'\ldots l't'}\}^2 \\ &\quad \times K^2(Z_{ij\ldots lt}, Z_{i'j'\ldots l't'})\end{aligned}$$

is a standard second-order U-statistic. Under a similar reasoning as in (I.28), it is easy to show that $E[\|\mathcal{B}^b_{\mathbb{N}}(\chi_{ij\ldots l}, \chi_{i'j'\ldots l'})\|^2] = O_p(|H_z|) = o_p(\mathbb{N})$, so we can conclude that $\widehat{\Sigma}_b \xrightarrow{p} \Sigma_b$ given that, as any element in $\mathbb{N}$ tends to infinity, we can prove that $\mathcal{B}^b_{\mathbb{N}}(\chi_{ij\ldots l}, \chi_{i'j'\ldots l'})$ converges in probability to a finite positive constant. Combining these results, it is proved that the standardized test statistic, $J^b_{\mathbb{N}}$, diverges to $+\infty$ at the rate of $\mathbb{N}|H_z|^{1/2}$, which completes the proof of this theorem.

∎

Proof of Theorem 5.3: Define $B^*(X_{ij\ldots lt}, \beta(Z_{ij\ldots lt}), \widehat{\beta}(Z_{ij\ldots lt})) = B(X_{ij\ldots lt}, \widehat{\beta}(Z_{ij\ldots lt}; H_z)) - B(X_{ij\ldots lt}, \beta(Z_{ij\ldots lt}; H_z))$, where $B(X_{ij\ldots lt}, \beta(Z_{ij\ldots lt})) \equiv X^\top_{ij\ldots lt}\beta(Z_{ij\ldots lt})$ and $B(X_{ij\ldots lt}, \widehat{\beta}(Z_{ij\ldots lt})) \equiv X^\top_{ij\ldots lt}\widehat{\beta}(Z_{ij\ldots lt}; H_z)$. Under H^c_0, the transformed nonparametric residuals can be written as $\widehat{v}_{ij\ldots lt}(z_1) = \widetilde{v}_{ij\ldots lt} - \widetilde{B}^*(X_{ij\ldots lt}, \beta(Z_{ij\ldots lt}), \widehat{\beta}(Z_{ij\ldots lt}))$, where $\widetilde{B}^*(X_{ij\ldots lt}, \beta(Z_{ij\ldots lt}), \widehat{\beta}(Z_{ij\ldots lt}))$ is the corresponding transformed expression of $B^*(X_{ij\ldots lt}, \beta(Z_{ij\ldots lt}), \widehat{\beta}(Z_{ij\ldots lt}))$. The expression to analyze is

$$\begin{aligned}I\!I^c_{\mathbb{N}} &= \frac{1}{\mathbb{N}^2 T^2} \sum_{ij\ldots lt} \sum_{i'j'\ldots l't'} \widetilde{v}_{ij\ldots lt} \widetilde{v}_{i'j'\ldots l't'} X^\top_{ij\ldots lt} X_{i'j'\ldots l't'} K_{H_z}(Z_{ij\ldots lt}, Z_{i'j'\ldots l't'}) \\ &+ \frac{1}{\mathbb{N}^2 T^2} \sum_{ij\ldots lt} \sum_{i'j'\ldots l't'} \widetilde{B}^*(X_{ij\ldots lt}, \beta(Z_{ij\ldots lt}), \widehat{\beta}(Z_{ij\ldots lt})) \widetilde{B}^*(X_{i'j'\ldots l't'}, \beta(Z_{i'j'\ldots l't'}), \widehat{\beta}(Z_{i'j'\ldots l't'})) \\ &\quad \times K_{H_z}(Z_{ij\ldots lt}, Z_{i'j'\ldots l't'}) \\ &- \frac{2}{\mathbb{N}^2 T^2} \sum_{ij\ldots lt} \sum_{i'j'\ldots l't'} \widetilde{v}_{ij\ldots lt} \widetilde{v}_{i'j'\ldots l't'} \widetilde{B}^*(X_{i'j'\ldots l't'}, \beta(Z_{i'j'\ldots l't'}), \widehat{\beta}(Z_{i'j'\ldots l't'})) K_{H_z}(Z_{ij\ldots lt}, Z_{i'j'\ldots l't'}) \\ &= I\!I^c_{\mathbb{N}_1} + I\!I^c_{\mathbb{N}_2} - 2 I\!I^c_{\mathbb{N}_3}, \qquad \text{(I.32)}\end{aligned}$$

where the definitions $I\!I^c_{\mathbb{N}_s}$, for $s = 1, 2, 3$, should be apparent from the context.

Considering the behavior of $I\!I^c_{\mathbb{N}_1}$, it can be written as a second-order U-statistic of the form

$$I\!I^c_{\mathbb{N}_1} = \prod_{r=1}^{R} \binom{N_r}{2}^{-1} \sum_{i=1}^{N_1-1} \sum_{i'=1+i}^{N_1} \sum_{j=1}^{N_2-1} \sum_{j'=1+j}^{N_2} \cdots \sum_{l=1}^{N_l-1} \sum_{l'=1+l}^{N_l} H^c_{\mathbb{N}}(\chi_{ij\ldots l}, \chi_{i'j'\ldots l'})|H_z|^{-1},$$

where now $\chi_{ij\ldots l} = (Z_{1ij\ldots l}, Z_{2ij\ldots l})$ and $H^c_{\mathbb{N}}(\chi_{ij\ldots l}, \chi_{i'j'\ldots l'})$ is a symmetric, centered and degenerate variable of the form

$$H^c_{\mathbb{N}}(\chi_{ij\ldots l}, \chi_{i'j'\ldots l'}) = \binom{T}{2}^{-1} \sum_{t=1}^{T-1} \sum_{t'=1+t}^{T} \widetilde{v}_{ij\ldots lt} \widetilde{v}_{i'j'\ldots l't'} K(Z_{ij\ldots lt}, Z_{i'j'\ldots l't'}).$$

Let $G^c_{\mathbb{N}}(\chi_{11\ldots 1}, \chi_{ij\ldots l}) = E[H^c_{\mathbb{N}}(\chi_{11\ldots 1}, \chi_{ij\ldots l}) H^c_{\mathbb{N}}(\chi_{22\ldots 2}, \chi_{ij\ldots l})]$, straightforward calculations yield a similar result as in (I.30). Under Assumption 3.3, it can be shown that $I\!I^c_{\mathbb{N}_1}$ has zero mean and

$$Var(I\!I^c_{\mathbb{N}_1}) = \frac{2^{(R+1)}(T-1)\sigma_v^4 R^q(K)}{\mathbb{N}^2 T^3 |H_z|} E[E[tr\{X_{ij\ldots lt} X^\top_{i'j'\ldots l't'}\}^2 | Z_{ij\ldots lt}] f(Z_{ij\ldots lt})](1 + o_p(1)).$$

Previously it was shown that $\widehat{\beta}(z; H_z) - \beta(z) = O_p\left(\|H_z\| + \sqrt{\frac{ln\mathbb{N}}{\mathbb{N}|H_z|}}\right)$. Following similar arguments as in Li and Wang (1998) and Zheng (1996), it is easily shown that the test statistic to analyze is

$$\mathbb{N}|H_z|^{1/2} \widehat{I}^c_{\mathbb{N}_1} \;=\; \mathbb{N}|H_z|^{1/2} I\!I^c_{\mathbb{N}_1} + O_p(\sqrt{\mathbb{N}|H_z|}\|H_z\|) + O_p(|H_z|^{1/2}), \tag{I.33}$$

given that we can prove that $\mathbb{N}|H_z|^{1/2} I\!I_{\mathbb{N}_2} = O_p(\mathbb{N}|H_z|^{1/2}\|H_z\|^4)$ and $\mathbb{N}|H_z|^{1/2} I\!I_{\mathbb{N}_3} = O_p(\|H_z\|^2) + O_p(\mathbb{N}^{-1/2}|H_z|^{-1/2})$, whereas $\mathbb{N}|H_z|^{1/2} I\!I_{\mathbb{N}_2} = O_p(\mathbb{N}|H_z|^{1/2})$ and $\mathbb{N}|H_z|^{1/2} I\!I_{\mathbb{N}_3} = O_p(1)$ under H^c_1.

Using the CLT in Theorem 4.1 in (I.33), we can conclude that, under Assumption 3.7 and as $\mathbb{N}_{\max} \to \infty$,

$$\mathbb{N}|H_z|^{1/2} \widehat{I}^c_{\mathbb{N}} \xrightarrow{d} N(0, \Sigma_c), \tag{I.34}$$

where $\Sigma_c = \frac{2^{(R+1)}(T-1)\sigma_v^4 R^q(K)}{T^3} E[E[tr\{X_{ij\ldots lt} X^\top_{i'j'\ldots l't'}\}^2 | Z_{ij\ldots lt}] f(Z_{ij\ldots lt})]$. Following a similar reasoning as in (I.28) and using the fact that $\sup_{z \in I\!R} |\widehat{\beta}(Z_{1ij\ldots lt}; H_z) - \beta(Z_{1ij\ldots lt})| = O_p\left(\|H_z\|^2 + (\mathbb{N}|H_z|)^{-1/2}\right)$ and $v_{ij\ldots lt}$ is an independent random variable of $(X_{ij\ldots lt}, Z_{ij\ldots lt})$, it is straightforward to show that $\widehat{\Sigma}^c$ converges in probability to a finite positive constant. Combining the above results, we are able to prove that $J^c_{\mathbb{N}}$ diverges to $+\infty$ at the rate $\mathbb{N}|H_z|^{1/2}$ and the proof of the theorem is finished.

∎

References

Ai, C. and Q. Li. 2008. *Semi-parametric and non-parametric methods in panel data models*, The econometrics of panel data: Fundamentals and recent developments in theory and practice, pp. 451–478.

Amini, S., M.S. Delgado, D.J. Henderson, and C.F. Parmeter. 2012. *Fixed vs random: the Hausman test four decades later*, Essays in Honor of Jerry Hausman (Advances in Econometrics), vol. 29, pp. 479–513.

Baier, S.L. and J.H. Bergstrand. 2007. *Do free trade agreements actually increase members' international trade?*, Journal of International Economics **71**, 72–95.

Balazsi, L., L. Matyas, and T. Wansbeek. 2015. *The estimation of multi-dimensional fixed effects panel data models*, Econometric Reviews **37**, 212–227.

Baltagi, G., B.H. an Bresson and J.M. Etienne. 2015. *Hedonic housing prices in Paris: An unbalanced spatial lag pseudo-panel model with nested random effects*, Journal of Applied Econometrics **30**, 509–528.

Brown, B.M. 1971. *Martingale central limit theorems*, Annals of Mathematical Statistics **42**, 59–66.

Bryan, G. and M. Morten. 2019. *The aggregate productivity effects of internal migration: Evidence from Indonesia*, Journal of Political Economy **127**, 2229–2268.

Carneiro, A., P. Guimaraes, and P. Portugal. 2012. *Real wages and the business cycle: accounting for worker and firm heterogeneity*, American Economic Journal: Macroeconomics **4**, 133–152.

Galvao, A.F. and G.V. Montes-Rojas. 2017. *Multi-dimensional panels in quantile regression models*, The Econometrics Multi-dimensional Panels, pp. 239–262.

Gao, Y. and L. Kunpeng. 2013. *Nonparametric estimation of fixed effects panel data models*, Journal of Nonparametric Statistics **25**, 679–693.

Gerardi, K.S. and A.H. Shapiro. 2009. *Does competition reduce price dispersion? New evidence from the airline industry*, Journal of Political Economy **117**, 1–37.

Gonçalves, S. 2011. *The moving blocks bootstrap for panel linear regression models with individual fixed effects*, Econometric Theory **27**, 1048–1082.

Gregg, A.G. 2020. *Factory productivity and the concession system of incorporation in late Imperial Russia*, American Economic Review **110**, 401–427.

Guimaraes, P. and P. Portugal. 2010. *A simple feasible procedure to fit models with high-dimensional fixed effects*, Stata Journal **10**, 628–649.

Hall, P. 1984. *Central limit theorem for integrated square error of multivariate nonparametric density estimators*, Journal of Multivariate Analysis **14**, 1–16.

Härdle, W. and E. Mammen. 1993. *Comparing nonparametric versus parametric regression fits*, Annals of Statistics **21**, 1926–1947.

Henderson, D. J., R. J. Carroll, and Q. Li. 2008. *Nonparametric estimation and testing of fixed effects panel data models*, Journal of Econometrics **144**, 257–275.

Henderson, D. J. and C. F. Parmeter. 2015. *Applied nonparametric econometrics*, Cambridge University Press, New York.

Henderson, D.J., J.M. Rodriguez-Poo, and A. Soberon. 2021. *Nonparametric multi-dimensional fixed effects panel data models*, Econometric Reviews **0**, 1–38.

Kapetanios, G. 2008. *A bootstrap procedure for panel data sets with many cross-sectional units*, Econometrics Journal **11**, 377–395.

Lavergne, P. and Q. Vuong. 2000. *Nonparametric significant testing*, Econometric Theory **16**, 576–601.

Li, Q., C.J. Huang, D. Li, and T. Fu. 2002. *Semiparametric smooth coefficient models*, Journal of Business and Economic Statistics **20**, 412–422.

Li, Q. and S. Wang. 1998. *A simple consistent bootstrap test for a parametric regression function*, Journal of Econometrics **87**, 145–165.

Lin, Z., Q. Li, and Y. Sun. 2014. *A consistent nonparametric test of parametric regression functional form in fixed effects panel data models*, Journal of Econometrics **178**, 167–179.

Matyas, L. 2017. *The econometrics of multi-dimensional panels: Theory and applications*, Advanced Studies in Theoretical and Applied Econometrics, 50, Springer.

Parmeter, C. F. and J. S. Racine. 2019. *Nonparametric estimation and inference for panel data models*, Paned data econometrics.

Powell, J.L., J.H. Stock, and T.M. Stoker. 1989. *Semiparametric estimation of index coefficients*, Econometrics **57**, 1403–1430.

Rodriguez-Poo, J.M. and A. Soberon. 2017. *Nonparametric and semiparametric panel data models: recent developments*, Journal of Economic Surveys **31**, 923–960.

Su, L. and A. Ullah. 2006. *Profile likelihood estimation of partially linear panel data models with fixed effects*, Economics Letters **92**, 75–81.

______. 2011. *Nonparametric and semiparametric panel econometric models: Estimation and testing*, Handbook of Empirical Economics and Finance, pp. 455–497.

Sun, Y., R. J. Carroll, and D. Li. 2009. *Semiparametric estimation of fixed effects panel data varying coefficient models*, Advances in Econometrics **25**, 101–130.

Sun, Y., W. Lin, and Q. Li. 2017. *Nonparametric models with random effects*, The Econometrics of Multidimensional Panels. Advanced Studies in Theoretical and Applied Econometrics, vol. 50, pp. 195–238.

Wansbeek, T. 1991. *Singular value decomposition of design matrices in balanced ANOVA models*, Statistics & Probability Letters **11**, 33–36.

Whang, Y. and D.W.K. Andrews. 1993. *Tests of specification of parametric and semiparametric models*, Journal of Econometrics **57**, 277–318.

Yao, F. and T. Wang. 2020. *A nonparametric test of significant variables in gradients*, Econometric Theory **0**, 1–45.

Zheng, X. 1996. *A consistent test of functional form via nonparametric estimation technique*, Journal of Econometrics **75**, 263–289.

University of Alabama and University of Miami and Universidad de Cantabria