Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsap/130306

Heterogeneous Three-Dimensional Panel Data Models

A. Megally[1], A. Youssef[2] and Abdel-Latif, S. H.[2,]*

[1]October University for Modern Sciences and Arts (MSA), Cairo, Egypt
[2]Department of Applied Statistics and Econometrics, Faculty of Graduate Studies for Statistical Research, Cairo University, Cairo, Egypt

Received: 30 Jun. 2023, Revised: 26 Jul. 2023, Accepted: 21 Aug. 2023.
Published online: 1 May 2024

Abstract: This paper examines panel data models, which vary in size and shape depending on their properties. When parameters vary across units and time, it is called a heterogeneous panel data model. The paper proposes suitable tests for slope homogeneity in three-dimensional panel data, using actual data from the United Nations Industrial Development Organization (UNIDO). The results showed significant evidence of cross-sectional dependence and slope heterogeneity. The paper recommends mean group, common correlated effect mean group, and augmented mean group estimators for heterogeneous three-dimensional panel data models in the case of cross-sectional dependency, with the common correlated effect mean group estimator being the most accurate.

Keywords: cross-sectional dependence, delta test, heterogeneous, slope homogeneity, three-dimensional panel data model.

1 Introduction

All panel data models, whether they are fixed effect or random effect models, assume that the slope (rate of change) of the parameter being studied is the same for all units in the dataset. Not considering differences in slope could make the results incorrect, so we should think about it [1]. The F-test is a statistical test that helps us compare the differences in errors between different regression models using cross-sectional data. It can be used to check if the slopes in these models are similar. The main issue with the later test is that it assumes the error variances are the same. The F-test doesn't work well unless there are more observations than variables. Experimental studies often do not use panels when the number of observations (T) is greater than the number of units (N) because panels are not commonly used in such cases.

In this paper, large amounts of data were used from different time periods to conduct various tests. Three tests that are used with two-dimensional panel data are proposed for use with three-dimensional panel data for the first time. Pesaran and Yamagata created the Delta test using a changed version of Swamy's test [2]. The test statistic followed a normal distribution when there was no difference in the slope. This test compares two types of statistical analysis: one that looks at different units individually, and one that combines all the units together. The Delta test allows for uneven variation because it considers the individual-specific differences in the standard errors. Furthermore, they provided a modified version of Blomquist and Westerlund's HAC test that examined both heteroskedasticity (unequal variances) and autocorrelation (relationship between data points). Cross-sectional dependence (CSD) can also be assessed using a CSD test. The researchers in [3, 4] suggested using a different test that did not include average values from different groups in the results.

In the econometrics literature, cross-sectional dependence has been extensively addressed. The researchers in [5] addressed the problem that cross-sectional dependence has rarely been allowed for in multi-level panel data models. It is defined as the interaction between cross-sectional units (such as households, enterprises, states, … etc.). It makes sense to think of reliance over "space" as the antithesis of serial correlation in time series. It may result through social interactions between people, such as consumer imitation and learning within a community, or businesses operating in the same sector. In game theory and industrial organization, this has received a lot of attention. It might also be the result of widely prevalent macroeconomic shocks or unobservable shared variables. Cross-sectional dependence among people is a problem in recent literature when cross-sectional units is large. The cross-sectional of dependence and correlation invalidates traditional t-tests and F-tests that employ standard variance-covariance estimators and leads to efficiency loss for least squares, just like serial correlation in time-series analysis. A jackknife method for determining the nature of fixed effects in three-dimensional panel data models in the case of weak serial or cross-sectional dependency among the error terms was presented by [6]. In a regression model with seemingly unrelated regressions, [7] suggests a cross-sectional dependence CSD test utilizing the pairwise average of the off-diagonal sample correlation coefficients. For many cross-sectional units N measured across T time periods, the CSD test is applicable. The CSD test is viewed as a test for weak cross-sectional

*Corresponding author e-mail: shereen_hamdy_@cu.edu.eg

dependence [8].

Economists have been paying a lot of attention to large-scale panel data models in econometrics. Pesaran created and discovered the approximate normal distributions of the common correlated effects (CCE) estimators for panels with diverse characteristics, following reasonably broad limitations [3]. Bai studied how well the principal component analysis estimators work and found that the square root of the product of the individual and time series dimensions, NT, is reliable [9]. In simpler terms, researchers in [10] used a method called quasi-maximum likelihood estimation (QMLE) to study a type of statistical model known as dynamic linear panel data models with interactive fixed effects. They looked at a method called principal component analysis (PCA) estimation, which was previously proposed by [9], and found two reasons why it may not provide accurate results in large samples: firstly, the error term in the model may have a certain pattern over time or may have varying amounts of variability across observations; secondly, the presence of certain variables that are determined before the analysis takes place may also introduce bias in the estimation results. In addition, they examined how we can trust the estimate of PCA for panel data models when we don't know how many factors should be included as interaction fixed effects [11]. The decision about the number of factors to include is based on certain criteria for gathering information. Researchers in [12] studied a method called Pesaran's CCE approach for estimating a panel data model with several factors that affect the errors and spatial error correlations. They found that this method consistently and accurately estimated the slope coefficients.

For big panels, Pesaran and Yamagata in [2] changed Swamy's "slope homogeneity test". Models that only have factors outside of our control and have errors that follow a normal distribution will show a predictable normal pattern. The simulation tests using Monte Carlo simulations found that in panels with only independent variables, the test had the correct level of significance and ability to detect relationships. In the suggested test, autoregressive (AR) models with moderate root values did well. However, a modified version of the test was suggested for models with certain characteristics, and it works well even when the number of observations is large compared to the number of variables. They studied income patterns of individuals over time using Panel Study of Income Dynamics Data (PSID) data. They looked at models that assume the income changes are similar for everyone. Even when people with similar levels of education are grouped together, they found that the results show significant differences in how earnings change over time.

When researchers studied panel data models with many variables and a specific type of error, they faced difficulty in testing if the slopes were the same. They explored this problem [13]. The Swamy-type test was used to combine the interaction-fixed effects. The proposed test statistic can connect the things that explain something to the things that we can't see, like how much they affect each other. Based on "Monte Carlo simulations," the proposed test effectively controls sample size and has enough power.

By incorporating the heterogeneity tests and estimating techniques employed in two-dimensional models, they expanded the assumption of homogeneity in "multi-dimensional panel data models" [14]. To do this, the multidimensional model was modified with F and Swamy's tests for parameter homogeneity. When the heterogeneity was found, the four-dimensional panel data model was estimated using both the mean group and random coefficients estimators that were used in two-dimensional case. They tested the application of Okun's Law in European nations. They discovered that the countries influenced the parameter of heterogeneity. The results show that Okun's Law holds true in European nations, but that Okun's coefficient differs from one country to another.

1.1 Objective of the paper

- Applying CSD test to detect if there is any cross-sectional dependence.
- Applying Delta and HAC tests for testing slope heterogeneity.
- Estimating heterogonous slopes using Mean Group (MG), Common Correlated Effects Mean Group (CCEMG) and Augmented Mean Group (AMG) methods.
- Showing the difference between MG and CCEMG estimates for each country.

This paper is organized in this way. Section 2 looks at and talks about the economic theory that supports the different tests and estimation methods. Section 3 shows what happened when we used these tests and estimation techniques with the real data from UNIDO. A conclusion is the final part of the paper.

2. Materials and methods

In this part, the shape of a three-dimensional model will be demonstrated when its parameters are heterogeneous. the econometric theory that underlies the suggested homogeneity tests and "heterogeneous panel estimators" will also be demonstrated. A panel data model in three dimensions:

$$y_{ijt} = \beta_o + \sum_{k=1}^{K} \beta_k x_{kijt} + u_{ijt} \qquad i = 1,\dots,N,\ j = 1,\dots,M,\ t = 1,\dots,T \qquad (1)$$

The constant and slope parameters of the model are both homogeneous, and the unit dimensions in equation (1) are i, j and the time dimension is t. Given the heterogeneity of the model parameters in equation (1), we obtain:

$$y_{ijt} = \beta_{oi} + \sum_{k=1}^{K} \beta_{ki} x_{kijt} + u_{ijt} \qquad i = 1,\dots,N,\ j = 1,\dots,M,\ t = 1,\dots,T \qquad (2)$$

The constant and slope parameters in equation (2) are heterogeneous for the i^{th} dimension. Furthermore, heterogeneity may be added to all or any of the units/time in equation (2). When adopting a panel data model, it is crucial to check the homogeneity of all parameters first. If the homogeneity of the parameters is assumed to be true, it is also necessary to test the homogeneity of the constants.

2.1. Homogeneity tests:

The authors talked about how to collect and analyze data from different sources using panel data models that cover a long period of time [1]. Wald, Swamy, and Pesaran et al. did something in [15], the Likelihood Ratio test was used to check if things were similar or different. The Delta, HAC, and CSD tests were suggested in this paper because they have not been used before in three-dimensional panel data models.

2.1.1. The standard Delta test:

This test was introduced to analyze heterogeneous panel data with large dimensions, and it was based on Swamy's slope homogeneity test [2]. The test has the following assumptions: the variance is allowed to be heterogeneous, and the distributions of $\varepsilon_{i,t}$ and $\varepsilon_{j,s}$ are independent in the cases where the two values disagree. The test is computed as follows:

$$\tilde{\Delta} = \frac{1}{\sqrt{N}} \left(\frac{\sum_{i=1}^{N} \tilde{d}_i - k}{\sqrt{2k}} \right) \qquad (3)$$

This test statistic follows the normal distribution asymptotically. According to Eq. (3), the "$\tilde{d}_i$" value shows how much the estimate for a cross-sectional specific unit differs from the pooled estimate when we consider the weights of each unit:

$$\tilde{d}_i = (\hat{\beta}_{ki} - \hat{\beta}_{KWFE})' \ \frac{X'_{ki} M_{1i} X_{ki}}{\tilde{\sigma}_i^2} (\hat{\beta}_{ki} - \hat{\beta}_{KWFE}) \qquad (4)$$

Where $X_{ki} = (X_{ki,1},\dots,X_{ki,T_i})'$ $M_{1i} = I_{T_i} - Z_{1i}(Z'_{1i}Z_{1i})^{-1}Z'_{1i}$ and $Z_{1i} = (\tau_{T_i}, X_{ki})$ with τ_{T_i}represents the constant with a $(T_i \times 1)$ vector of ones. $\hat{\beta}_{ki}$ and $\hat{\beta}_{KWFE}$ are defined as follows:

$$\hat{\beta}_{ki} = (X'_{ki} M_{1i} X_{ki})^{-1} X'_{ki} M_{1i} y_i \qquad (5)$$

$$\hat{\beta}_{KWFE} = \left(\sum_{i=1}^{N} \frac{X'_{ki} M_{1i} X_{ki}}{\tilde{\sigma}_i^2} \right)^{-1} \sum_{i=1}^{N} \frac{X'_{ki} M_{1i} y_i}{\tilde{\sigma}_i^2} \qquad (6)$$

Where $y_i = (y_{i,1},\dots,y_{i,T_i})$

$$\tilde{\sigma}_i^2 = \frac{(y_i - X_{ki}\hat{\beta}_{FE})' M_{1i} (y_i - X_{ki}\hat{\beta}_{FE})}{T_i - 1}, \qquad (7)$$

$$\hat{\beta}_{FE} = \left(\sum_{i=1}^{N} X'_{ki} M_{1i} X_{ki} \right)^{-1} \sum_{i=1}^{N} X'_{ki} M_{1i} y_i \qquad (8)$$

The unimportant regressors and the constant (α_i) are considered to be heterogeneous and then are accumulated in (Z_{1i}) and partially separated utilizing the projection matrix (M_{1i}). The following assumption $(N,T) \overset{j}{\rightarrow} \infty$ underpins the test statistic's asymptotic properties, so that $\frac{\sqrt{N}}{T^2} \rightarrow 0$. When we change Eq. (1) to a regular first AR model, the results shown by [2] stay the same. The mean-variance bias adjusted $\tilde{\Delta}$, when the errors follow a normal distribution, can be expressed as:

$$\tilde{\Delta}_{adj} = \sqrt{N} \left(\frac{N^{-1} \sum_{i=1}^{N} \tilde{d}_i - k}{\sqrt{var(\tilde{z}_i, T_i)}} \right) \qquad (9)$$

Where $var(\tilde{z}_i, T_i) = \frac{2k(T_i - k - 1)}{T_i - k + 1}$ (10)

2.1.2. A heteroskedastic and autocorrelation HAC robust test:

The researchers in [2, 16] modified a new version of the Delta test and it is known as the HAC test. The test statistic is a number used to analyze data in a test:

$$\tilde{\Delta}_{HAC} = \sqrt{N} \left(\frac{N^{-1} S_{HAC} - K}{\sqrt{2k}} \right), \qquad (11)$$

where $S_{HAC} = \sum_{i=1}^{N} T_i(\hat{\beta}_{ki} - \hat{\beta}_{KHAC})'(\hat{Q}_{i,T_i}\hat{V}_{i,T_i}^{-1}\hat{Q}_{i,T_i})(\hat{\beta}_{ki} - \hat{\beta}_{KHAC})$ (12)

$$\hat{\beta}_{KHAC} = (\sum_{i=1}^{N} T_i\hat{Q}_{i,T_i}\hat{V}_{i,T_i}^{-1}\hat{Q}_{i,T_i})^{-1}\sum_{i=1}^{N}\hat{Q}_{i,T_i}\hat{V}_{i,T_i}^{-1}X'_{ki}M_{1i}y_i \quad (13)$$

Where $\hat{\beta}_{ki}$ is the ordinary least square estimator for each i, M_{1i} as previously explained and

$\hat{Q}_{i,T_i} = T_i^{-1}(X'_{ki}M_{1i}X_{ki})$, The HAC adjustment is performed using the estimator below:

$$\hat{V}_{i,T_i} = \hat{\Omega}_i(0) + \sum_{j=1}^{T_i-1} k(j/B_{i,T_i})[\hat{\Omega}_i(j) + \hat{\Omega}_i(j)'], \quad (14)$$

$\hat{\Omega}_i(j) = T_i^{-1}\sum_{t=j+1}^{T_i}\hat{\check{u}}_{i,t}\hat{\check{u}}'_{i,t-j}$ and $\hat{\check{u}}_{i,t} = (\check{X}_{ki,t} - \bar{\bar{\check{X}}}_{ki,t})\hat{\varepsilon}_{i,t}$ With $\bar{\bar{\check{X}}}_{ki,t} = T_i^{-1}\sum_{t-1}^{T_i}\check{X}_{ki,t}$ where $\check{X}_{ki,t}$ is t^{th} element of $X_{ki}M_{1i}$. $\hat{\varepsilon}_{i,t}$ is a residual calculated from a usual regression method where M_{1i} is used as the projection matrix. In Eq. (13), B_{i,T_i} represents the bandwidth parameter, which determines the width of a function, and k represents the kernel function, which is a mathematical function used for smoothing or estimating data.

2.1.3 A cross-sectional dependence test:

In panels that have many different groups to study and a lot of time periods, there is a possibility that cross-sectional dependence is presented. There are two kinds of cross-sectional dependence: weak and strong [17]. A weak CD test means that when there are many observations over time, the relationship between them becomes less strong. When there is strong cross-dependence, the correlation becomes stable and does not change much. Spatial approaches are often employed to approximate cross-sectional dependency when it is weak. Strong cross-sectional dependence is formed by factor loading γ_i and a shared time-specific factor f_t. The same variables affect all cross-sectional units:

$$y_{i,t} = \mu_i + \beta'_{ki}X_{ki,t} + u_{i,t} \quad (15)$$

$$u_{i,t} = \gamma'_i f_t + \varepsilon_{i,t}$$

The unknown factor loadings are represented by γ_i which is a vector of size $(m \times 1)$. The common factor f_t is also unknown and represented by a vector of size $(m \times 1)$.

When you leave out explanatory variables and the common factors that are connected, it can cause a bias in the results due to the missing information. To estimate the factors that are commonly shared, you can use the average values from a specific time [3] or a statistical technique called principal components [9]. The CCE estimator developed by [3] can be used without needing to know how many common components will be used beforehand. The second technique can be used to eliminate strong cross-sectional dependence.

The researchers in [4] created a modified model that includes the lagged cross-sectional averages. The suggestion is to employ PCSA= $[T^{1/3}]$ when dealing with regressors that have weak exogeneity. Following that, we have the option to represent the equation (15) using cross-sectional notation.

$$y_{i,t} = \mu_i + \beta'_{1i}x_{1i,t} + \beta'_{2i}x_{2i,t} + \sum_{l=1}^{PCSA}\gamma_{i,l}\bar{v}_t + \varepsilon_{i,t} \quad (16)$$

$$\bar{v}_t = \frac{1}{N}\sum_{j=1}^{N}(x_{1j,t}, x_{2j,t}, y_{j,t}) \quad (17)$$

The term $\bar{v}_t$ means the average values taken across different sections and the variables $x_{1i,t}$ and $x_{2i,t}$ can include the dependent variable after getting the lag. There are two models that can be used with the CCE estimator: the pooled and mean group. As a result, it is easy to expand the current delta test to include average values from different sections and give advice on whether to use a pooled or mean group model. One suggestion was to break down the average data to remove the influence of any significant connection between different sections. This would help make the cross-sectional dependence robust delta test more reliable. Let's say the average values and their previous values are put in a matrix called $\bar{V}_t$. The dividing out is then done by:

$$\tilde{V}_t = \frac{1}{N}\sum_{j=1}^{N}(x_{1j,t}, x_{2j,t}, y_{j,t}),\ \bar{V}_t = (\tilde{V}_t, \dots\dots, \tilde{V}_{t-PCSA}) \quad (18)$$

$$M_{\bar{V}_t} = I_T - \bar{V}_t(\bar{V}_t'\bar{V}_t)^{-1}\bar{V}_t' \quad (19)$$

$$\check{y}_i = y_i M_{\bar{V}_t} \quad (20)$$

$$\check{X}_{1i} = X_{1i}M_{\bar{V}_t} \text{ and } \check{X}_{2i} = X_{2i}M_{\bar{V}_t} \quad (21)$$

Then, the variables are used to build $\tilde{\Delta}_{CSA}$ following Eq. (3) and Eq. (11) for the HAC test.

2.2. Heterogeneous panel estimators:

In the early methods of panel time-series econometrics, such as empirical estimators, cointegration and unit-root tests, it was assumed that the individuals in the panel were independent of each other in terms of their characteristics [18]. The second generation of techniques [9, 3] addressed the problem of panel member correlation in-depth. The three empirical estimators presented in this paper show that the assumption of panel member homogeneity has been weakened.

2.2.1. The mean group estimator:

The cross-section dependence in the MG estimator produced by [1] was not considered. All MG type estimators employ very similar mechanisms as follows:

1. Create N simple linear equations using the ordinary least-squares method for each group.

2. Put together the expected coefficients from all groups.

$$\hat{\beta}_{MG} = N^{-1} \sum_{i=1}^{N} \hat{\beta}_i \tag{22}$$

$$\hat{\beta}_i = (X_i' M_T X_i)^{-1} X_i' M_T y_i \tag{23}$$

2.2.2. The common correlated effects mean group estimator:

Using Pesaran's CCEMG estimator, cross-sectional dependency and time-varying unobservable with various effects on panel members can be considered [3]. The estimator of CCEMG includes the average values of both the dependent and independent variables ($\bar{y}_t$ and $\bar{x}_t$ respectively) in the additional factors used for regression, along with an intercept and the variable x_{it}. These additional regressors are the "group-specific regression equation". The averages are calculated from data gathered for the entire group, and then used as independent variables in each of the regression models. Among the panel members, the calculated coefficients $\hat{\beta}_i$ were averaged using various weights as necessary. For example, in [19], the Pesaran's CCE technique is used by considering that the first layer regressors have a factor structure.

The estimator's main goal for the observable variables is to generate accurate estimates of their parameter values. When using cross-sectional averages, the calculated coefficients and their average estimations are difficult to understand and merely serve to counteract the common factor's biassing effects. Local spillover effects as well as global shocks, like the recent financial crisis, can be easily managed by the CCEMG technique. It has a finite number of "weak" components and a few "strong" components. The estimator is also not affected by common variables that change over time [17].

2.2.3. The augmented mean group estimator:

In place of Pesaran's CCEMG estimator, Teal and Eberhardt employed the AMG estimator [20]. The CCEMG estimator considers the invisible common factor f_t as something that doesn't matter and is just a distraction for practical research. However, in cross-country production functions the unobservable reflect overall total factor productivity (TFP). The AMG technique is divided into three stages:

a) A model is created that combines data from different years. We use a method called "the first difference OLS" to estimate this model. We then calculate the coefficients on the "differenced year dummies and sum them up". They have made an estimated average of the TFP across different groups as time went on. This process is described as "the common dynamic process".

b) The estimated TFP process is then included in the group-specific regression model either as a variable or as a coefficient that affects each group member. This process is done by subtracting the estimated values from the dependent variable. Just like in the MG example, each regression model has a starting point that represents constant factors that don't change over time (TFP levels).

c) The model's parameters that are specific to a particular group are combined together for the entire panel, similar to the MG and CCEMG estimators. We may also assign weights to these parameters.

Both the estimation methods AMG and CCEMG did well in Monte Carlo simulations when dealing with panels that have "nonstationary variables" and "multifactor error terms". This was measured using root mean squared error or bias [21].

2.3. Data:

Data used in this paper is about the industrial sector which is available for the years 2005 and later in the (UNIDO's Industrial Demand-Supply Balance) database. The data were obtained from production data provided by "National Statistical Offices" as well as "UNIDO estimates for ISIC-based international trade data" (COMTRADE), according to the

"United Nations Commodity Trade database". Data on 137 industrial categories were submitted by a global panel of 74 nations between 2005 and 2013. The panel data is unbalanced as there are many missing values. A few of the trade-related topics covered in this study are output, imports, and apparent consumption. The impact of industrial output and imports on consumption was explored using three-dimensional panel data models that are (countries, industries, and years) in this paper. The three-dimensional panel data model according to this data is written as follows:

$$C_{ijt} = \beta_{oi} + \beta_{1i}O_{ijt} + \beta_{2i}I_{ijt} + u_{ijt} \quad (24)$$

Where, i: country, j: industry and t: year. C_{ijt}, O_{ijt} and I_{ijt} are the consumption, output and imports respectively of an industry i in a country j at year t.

3. Results and discussion:

In this section, we stated the results of our study. First, we applied the Pesaran's CSD test to check if there is cross-sectional dependence [7]. Second, Delta and HAC tests were applied to check the slope heterogeneity of the parameters. We found that the parameters were heterogonous, so we proposed the MG, CCEMG and AMG as appropriate estimators for these heterogonous parameters. Finally, we discussed the limitations of this study and a future plan for the coming research.

3.1. The result of cross-sectional dependence test:

The hypotheses of the CSD test:

H_o: Cross-sectional independence

H_A: Cross-sectional dependence

Table 1: Results for CD-test

Variables	CD-test	P-value
Output	**95.38**	**0.000**
Imports	**105.84**	**0.000**
Consumption	**103.44**	**0.000**

Notes: Under the null hypothesis of cross-section independence CD ~ N(0,1)

The CD test was employed to find "cross-sectional dependence" in the variables included in this data. Table (1) shows that the rejection of the null hypothesis of cross-section independence, which indicates the existence of cross-section dependence at 1% significance level.

3.2. The result of slope homogeneity test:

In this paper, heterogeneity is considered for only countries and β_i is heterogeneous parameter for countries. Firstly, parameter homogeneity was tested using the standard Delta test and HAC test. The results are summarized in the table below:

Table 2: Results for slope homogeneity tests

Delta test		HAC test	
Delta	**P-value**	**Delta**	**P-value**
-13.449	**0.000**	**-12.270**	**0.000**
Adj. -32.942	**0.000**	**Adj. -30.055**	**0.000**

Note: These tests automatically assume a heterogeneous constant. Following Andrews and Monahan [22], a Bartlett kernel and automatically selected bandwidth were used in the HAC test.

The Delta test score shows that the idea that the slopes are the same is not true. When we use the "HAC robust test" suggested by Blomquist and Westerlund [16] to consider autocorrelation in the leftover data, we find out that the assumption of homogeneity is incorrect. The results of the Delta and HAC tests agree, and the factors specific to each country are different. This means that the MG, CCEMG, and AMG estimators are the most suitable ones to use for this three-dimensional panel data model.

3.3. The result of estimation methods:

Table 3: the outcomes of the estimation methods

	MG	CCEMG	AMG
Output	**1.001** **(1938075)****	**0.938** **(37.70)****	**1.00** **(5418.95)****
Imports	**0.999** **(1002.36)****	**0.806** **(21.90)****	**0.999** **(2301.60)****
(c_d_p)	**-**	**-**	**0.696** **(1.00)**
Intercept	**-6638.67** **(-1.32)**	**-28071.62** **(-2.41)***	**-2962.71** **(-1.51)**
RMSE	**224.68**	**0.000**	**36.36**

*Note: "Statistical significance at the 5% and 1% level is indicated with * and ** respectively and t statistics are reported in brackets".*

The estimates of the variables for the MG and AMG estimators have similar values, as shown in Table (3). The variable estimates provided by CCEMG differ slightly from those offered by the other estimators. In contradiction of CCEMG estimator, which has the lowest RMSE, MG estimator has the greatest RMSE. While the Pesaran's CCEMG estimate allows for cross-section reliance and "time-variant" unobservable with heterogeneous influence across panel members, the MG estimator excludes any cross-section dependence.

Table 4: Slope estimates by countries

Methods / Country	$\hat{\beta}_1$		$\hat{\beta}_2$		Mean Consumption
	MG	**CCEMG**	**MG**	**CCEMG**	
Azerbaijan	**.9999**	**.9071**	**1.000**	**.9052**	**170195.4**
Austria	**1.000**	**.9910**	**.9999**	**.6529**	**2886544**
Belgium	**1.012**	**1.018**	**.9762**	**.8117**	**4004905**
Bulgaria	**1.000**	**.9865**	**.9999**	**.8427**	**492620.4**
Cyprus	**.9999**	**.9991**	**1.000**	**.7576**	**205119.7**
Estonia	**.9999**	**.9671**	**1.000**	**.8471**	**226442.5**
Germany	**.9999**	**1.013**	**1.000**	**.8408**	**2.25e+07**
Greece	**.9999**	**1.025**	**1.000**	**.9777**	**1438204**
Hungary	**1**	**1.137**	**1**	**.6830**	**1037748**
Italy	**.9999**	**.8604**	**1.000**	**.8949**	**1.58e+07**
Lithuania	**.9999**	**.9998**	**1.000**	**.7777**	**437900.9**
Portugal	**1**	**1.082**	**1**	**.9094**	**1700270**
Romania	**.9999**	**1.070**	**1.000**	**.9437**	**1393854**
Spain	**.9999**	**.9746**	**1.000**	**.7127**	**1.04e+07**
United Kingdom	**.9999**	**1.073**	**1.000**	**.8161**	**1.28e+07**

Note: All estimates are statistically significant at 1% level.

4. Conclusion:

It is advised in this paper to extend the assumption of homogeneity in multidimensional panel data models and to apply heterogeneity tests and estimating methods like those used in two-dimensional models. The Delta and HAC tests used to test parameter homogeneity were adapted to the three-dimensional model for this purpose. After heterogeneity was found, the MG, CCEMG, and AMG estimators used in the two-dimensional models were used in the "three-dimensional panel data model". Because there is no test and estimation approach in the literature that takes into consideration the variability of multidimensional panel data models, it is expected that this study will lead to new studies of original value in this area.

Advantages of our paper:

We applied familiar two-dimensional panel data tests in the case of three-dimensional to extend the assumption of using these tests or other ones in the multidimensional case and the same for estimation methods. We proved that the CCEMG estimators are the most accurate for heterogenous parameters in the case of cross-sectional dependence.

Conflict of interest

The authors say they do not have any conflicts of interest that could affect the research and its results. This includes any conflicts related to money, personal matters, who gets credit for the research, or anything else.

Financing

The study was performed without financial support.

Data availability

Data will be available upon on request.

Acknowledgments

We express our gratitude to Faculty of Graduate Studies for Statistical Research – Cairo University.

References

[1] M. H. Pesaran and R. Smith. Estimating long-run relationships from dynamic heterogeneous panels. *Journal of econometrics*, 68(1), 79-113, (1995).

[2] M. H. Pesaran and T. Yamagata. Testing slope homogeneity in large panels. *Journal of econometrics*, 142(1), 50-93, (2008).

[3] M. H. Pesaran. Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica*, 74(4), 967-1012, (2006).

[4] A. Chudik and M. H. Pesaran. Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors. *Journal of econometrics*, 188(2), 393-420, (2015b).

[5] G. Feng, J. Gao and B. Peng. Multi-Level Panel Data Models: Estimation and Empirical Analysis. *Monash University Journal Contribution*, (2022).

[6] X. Lu, K. Miao and L. Su. Determination of different types of fixed effects in three-dimensional panels. *Econometric Reviews*, 40(9), 867-898, (2021).

[7] M. H. Pesaran. General diagnostic tests for cross-sectional dependence in panels. *Empirical Economics*, 60(1), 13-50, (2021).

[8] A. Chudik and M. H. Pesaran. Large Panel Data Models with Cross-Sectional Dependence: a survey. *CAFE Research Paper*, 13.15, (2013).

[9] J. Bai. Panel data models with interactive fixed effects. *Econometrica*, 77(4), 1229-1279, (2009).

[10] H. R. Moon and M. Weidner. Dynamic linear panel regression models with interactive fixed effects. *Econometric Theory*, 33(1), 158-195, (2017).

[11] H. R. Moon and M. Weidner. Linear regression for panel with unknown number of factors as interactive fixed effects. *Econometrica*, 83(4),1543-1579, (2015).

[12] M. H. Pesaran and E. Tosetti. Large panels with common factors and spatial correlation. *Journal of Econometrics*, 161(2), 182-202, (2011).

[13] T. Ando and J. Bai. A simple new test for slope homogeneity in panel data models with interactive effects. *Economics Letters*, 136, 112-117, (2015).

[14] F. Y. Tatoğlu and H. İçen. *Heterogeneous Multi-Dimensional Panel Data Models: Okun's Law for NUTS2 Level in Europe1*, in Selected Topics in Applied Econometrics, Ebru Çağlayan Akay,Özge Kork, Ed. Peter Lang Publishing, Inc., Berlin: pp.31-45, (2019).

[15] M. H. Pesaran, A. Ullah and T. Yamagata. A bias-adjusted LM test of error cross-section independence. *The Econometrics Journal*, 11(1), 105-127, (2008).

[16] J. Blomquist and J. Westerlund. Testing slope homogeneity in large panels with serial correlation. *Economics Letters*, 121(3), 374-378, (2013).

[17] A. Chudik, M. H. Pesaran and E. Tosetti. Weak and strong cross-section dependence and estimation of large panels. *The Econometrics Journal*, 14(1), C45-C90, (2011).

[18] A. Levin, C. F. Lin and C. S. J. Chu. Unit root tests in panel data: asymptotic and finite-sample properties. *Journal of econometrics*, 108(1), 1-24, (2002).

[19] G. Kapetanios, L. Serlenga and Y. Shin. Estimation and inference for multi-dimensional heterogeneous panel datasets with hierarchical multi-factor error structure. *Journal of Econometrics*, 220(2), 504-531, (2021).

[20] F. Teal and M. Eberhardt. Productivity analysis in global manufacturing production. *Department of Economics Discussion Paper Series*, University of Oxford, (2010).

[21] M. Eberhardt and S. Bond. Cross-section dependence in nonstationary panel models: a novel estimator, *MPRA Paper 17692*, University Library of Munich, Germany, (2009).

[22] D. W. Andrews and J. C. Monahan. An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. *Econometrica*: *Journal of the Econometric Society*, 953-966, (1992).