European Journal of Economics, Finance and Business Development
ISSN (E): 2938-3633
Volume 2, Issue 6, June - 2024

ECONOMETRIC MODELING BASED ON PANEL DATA

Muminova Maxbuba Abduvafoyevna
Associate Professor, "Econometrics" Department, TSEU

Abstract

In this article, panel data, analysis, data representing the activities of a set of enterprises by quarters, annual socio-economic indicators of a specific group of countries, the construction of flexible and comprehensive models of panel data, and obtaining answers within the framework of spatial models are examined. The importance of panel data in analyzing the impact of various factors on the development of events over time, and in researching econometric and socio-economic processes, is discussed.

Keywords: panel data, time series, regression model, production functions, production logarithm, identification.

Introduction

Panel Data - Multivariate observation data that is repeatedly measured over time. If the sample contains data for several objects, each observed at multiple time intervals, such data is called panel data.

Panel data is a special type of data widely used in econometrics and socio-economic process research. It is a collection of observations for each research object at various time points. This data structure allows for the analysis of the dynamics and changes of events over time, as well as accounting for the individual characteristics of each object.

Panel data is used in economics, sociology, political science, and many other fields. It allows researchers to account not only for temporal changes but also for differences between research objects. This is particularly useful for analyzing the impact of various factors on the development of events over time.

For example, panel data can be used to analyze the impact of education on workers' wages. Researchers can collect data on wages and education levels for each employee over several years and analyze how these indicators change over time and how education affects wages.

Panel data is a set of observations that monitor the same variable or a set of variables over different periods or events. This data can be collected over several years, quarters, months, or days.

Panel data is a special form of data collected over time. Their main difference from "ordinary" data is that they allow for the analysis of changes in variables over time and account for these changes when conducting statistical tests and models.

This data is often used in economic and sociological research to study the dynamics and interrelationships of various factors. They help identify the impact of different factors on a variable and also predict the future values of variables based on past observations.

Panel data can include information on household income, commodity prices, unemployment rates, or consumer preferences. For each of these variables, data is collected at different times, and various factors affect their changes.

In econometrics, panel data often refers to multivariate data that includes measurements over a certain period. A panel data set can be time-based and include observations of individuals within a specific sample, or it can contain information about each individual in the sample.

For example, we might look at the sales and employee ratios of 50 firms over five years or, if panel data is taken geographically, we might consider the GDP and money supply ratio of 20 countries over 20 years.

Panel data consists of information obtained from consecutive observations of an economic subject or object (households, firms, regions, countries, etc.). Examples include studying the annual budget of several households, data reflecting the activities of a set of enterprises by quarters, or annual socio-economic indicators of a group of countries. Thus, panel data combines information similar to time series data with spatial data: spatial data is available for each economic object and corresponds to each object. Input data forms one or several time series. This special structure of panel data allows us to create more flexible and comprehensive models and answer questions that cannot be addressed within spatial models. Specifically, it enables the accounting and analysis of individual differences between economic units that cannot be addressed within standard regression models.

For example, by studying the per capita GDP value for a specific country, one can observe the inflation rate, investment volume, money supply, and more for each period. However, alongside these, there are factors that cannot be observed or quantified but significantly impact the studied indicator (eg, the object's geographic location, history, cultural and educational traditions). Such factors' effects can be considered constant (ie, independent of time) for each national economy. For only a few countries, spatial data can identify the impact of simple economic factors on per capita GDP, but it cannot assess differences in specific characteristics between countries. The assessment can be conducted based on an analysis of several years of observation results.

One of the constant issues at the micro-level is indicating households' expenditures on a specific type of product, such as personal care products. By including economic and socio-demographic living conditions, you can collect spatial data and find the relationship between household income and the price of the studied product. However, panel data may show that income is not always significant in household expenditures, as differences in expenditures may be influenced by family traditions, cultural level, and other factors that are not always measurable or observable.

Often, the individual factors of objects are correlated with other influencing variables. For instance, the overall cultural level of a household is naturally related to its income level. In regression, a specific factor indicates the importance of the object in the model, and ignoring it in the model leads to incorrect estimation of other parameters. In other words, models based on panel data lead to correct parameter estimation. In this case, panel data includes observation data of the same objects at different times. The assumption of independence between these observations can lead to a distortion of reality, requiring more precise methods than the OLS method for analyzing these models.

European Journal of Economics, Finance and Business Development
ISSN (E): 2938-3633
Volume 2, Issue 6, June - 2024

Panel data consists of information from observations of n different objects at various T times. Indexes can be used to indicate the numbers of the studied objects when recording inter-object data. For example,

y_{it} is the value of the dependent variable for the i-th object.

In panel data, it is necessary to specify the object number and time. This requires additional numbering, where index \(i \) represents the observation number, and index \(t \) represents the observation time. Thus,

y_{it} denotes the value of the dependent variable at time t for object i (a k-dimensional vector),

- x_{it} denotes the value of the independent variable at time t for object i ,
- e_{it} denotes the corresponding errors (i = 1,2,… ,n; t=1,2,..,T).

Even if some observations are missing in panel data, it still allows for representing the real situation. In balanced panels, all observations are known, meaning that all variables have observation values for each period and subject. If data for one period is missing for a particular object, such panels are called unbalanced.

The simplest model is the ordinary linear regression model:

$$y_{it} = x'_{it} + \varepsilon_{it} \qquad (1)$$

or by introducing the following notations

$$y_i = \begin{bmatrix} y_{i1} \\ y_{i2} \\ . \\ . \\ y_{iT} \end{bmatrix}, \; x_i = \begin{bmatrix} x'_{i1} \\ x'_{i2} \\ . \\ . \\ x'_{iT} \end{bmatrix}, \; \varepsilon = \begin{bmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \\ . \\ . \\ \varepsilon_{iT} \end{bmatrix};$$

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ y_n \end{bmatrix}, \; X = \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ x_n \end{bmatrix}, \; \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ \varepsilon_n \end{bmatrix}.$$

we write it in matrix form as follows:

$$Y = X \cdot \beta + \varepsilon. \qquad (2)$$

As mentioned, panel data allows for accounting for unique individual differences between economic units (objects). One possible way to implement such ideas is to express the model in the following form:

$$y_{it} = \alpha_i + x'_{it}\beta + \varepsilon_{it}. \qquad (3)$$

Here, i represents the specific effect of object i that is independent of time t. Accounting for the unique individual effect of the object allows for more accurate conclusions. For example, the usual approach to evaluating production functions involves estimating the equation.

$$y_{it} = \mu + x_{it}^{'}\beta + \varepsilon_{it}. \quad (4)$$

Here, ln(Y) represents the logarithm of production, and ln(X) is a k-dimensional vector of the logarithms of the production functions.

To have an accurate understanding of production functions (especially for smaller firms), it is appropriate to include the quality of management as part of the production factors and take it into account, meaning:

$$y_{it} = \mu + x_{it}^{'}\beta + q_i\beta_{k+1} + \varepsilon_{it}. (5)$$

Here, q_i represents the quality of management.

If this variable is significant, then the model (4) cannot be estimated using OLS. However, since it is unobservable, it can be accounted for as a unique individual effect of management. In model (5), it cannot be identified, so it cannot be estimated. In model (3), depending on the assumptions about the characteristics of the parameter \(u _i \), two models can be considered.

Fixed effects model: In equation (3), \(u _i \) is assumed to be an unknown parameter (or a model with fixed effects).

Random effects model: In equation (3), \(u _i \) is assumed to be random. Here, \(\ mu \) is a parameter common to all units at all times. \(\ epsilon_{it} \) is the error term (this model is recommended for the reader to study independently).

The choice of model is resolved individually in each case.

Fixed effects regression models: The fixed effects regression estimation method allows accounting for the influence of characteristics that differ across observed objects but are constant over time, which are not included in the regression equation.

In a fixed effects regression model, n different constant terms are used, one for each object. These constants can be seen as a set of binary (or indicator) variables that account for the influence of all unobserved variables that differ across observation objects but remain constant over time.

Fixed effects regression model:

If this variable important if , then model (4). with EKKU of assessment done by increasing it won't be. But q_i unobservable that it was for him management the unique quality of the individual effect account get can β_{k+1} (5) in the model non-identifiable that it was for him evaluation opportunity there is it's not . (3) in the model α_i of the parameter feature relatively to assumptions dependent respectively two the model to see can

Immutable effective model: (3) in Eq α_i unknown parameter assumption will be done (or immutable effect doer to the factor have the model that is).

Random effective model: (3) in Eq $\alpha_i = \mu + u_i$ assuming that will be done . Here μ is a parameter all units for everyone time general . u_i - error (this model to the reader independent learning for recommendation will be done).
The model choose issue each one in case separately will be solved .

Immutable efficient regression models

Immutable efficient regression evaluation method regression to Eq not included , being monitored objects according to difference doer , but time according to immutable has been of characters effect account get enable giver is a method .
Immutable to the effect have has been regression in the model each one object for from one *n* different immutable numbers (constants) are used . These are constants different wait objects for different , but time during immutable which is or this level regression from Eq down left all variables effect level account will receive binary (or indicator) variables collection in the form of imagination to be done can
Immutable to the effect have has been regression model .

$$y_{it} = \beta_0 + \beta_1 x_{it} + \beta_2 z_i + u_{it} \quad (6)$$

this where : y_{it} – dependent variable ;
x_{it} - observable dependent didn't happen variable ;
z_i - separately in objects differently values acceptance doer , but time during unchanging , unobservable variable (eg object to the situation effect doer cultural , social or political status).
Here the main issue is unobservable , fixed *z* variable when participating *x* and *y* of effect pointer β_i coefficients from assessment consists of z_i different objects for different to values have was , but time during variable expression that it was for $\alpha_i = \beta_0 + \beta_1 \cdot z_i$ the total *n* different in objects differently to values have has been view as an expression can Then the expression (6) . as follows to write can :

$$y_{it} = \beta_1 x_{it} + \alpha_i + u_{it} \qquad (7)$$

This is an expression immutable to the effect have has been regression model represents Here $\alpha_1, \alpha_2, \ldots, \alpha_n$ assessment need was unknown are considered invariant -constants . This of constants The interpretation of equation (7). right side while being studied come comes out Regression the line in assessment received deviation constant b_1 all objects for one different Constants to objects depends respectively will change .
In expression (7). α_i lar " *i* – of the object special to himself special as a feature view can , $\alpha_1, \alpha_2, \ldots, \alpha_n$ and some cases immutable effect of contributing (effective) factors exactly

he is considered himself . of these constants o' change attention not received of variables change with depends to be possible , (for example , in equation (6). z_i).

Tracking objects for special was , above constants each one observation object designation for used binary variables using modeling can Binary of variables using , unchanged effect doer factor regression model make up for , together at *i* =1 equal to to be and another cases to zero equal to to be D_{1i} - binary variable , when *i* =2 together equal to divisor and another cases to zero equal to to be D_{2i} - variable and etc. But all n are binary variables common with a free term one of the moment in itself regression to Eq input multicollinearity come to exit take will come . That's why for optional respectively binary of variables one down to leave need (eg first group observations for D_{1i}) . Then to equation (6) . equivalent was immutable efficient regression equation as follows writing can :

$$y_{it} = \beta_0 + \beta_1 x_{it} + \gamma_2 D_{2i} + \gamma_3 D_{3i} + \ldots + \gamma_n D_{ni} + u_{it} \quad (8)$$

this on the ground $\beta_0, \beta_1, \gamma_2, \ldots, \gamma_n$ - evaluation need was unknown coefficients .

(7) in Eq coefficients with (8) in Eq coefficients between relationship determination for their right sides equalization need (3) in Eq the first object regression line i for $\beta_0 + \beta_1 x_{it}$ expression to write can , then $\alpha_1 = \beta_0$ will be Second and from him next of objects regression line i for expression $\beta_0 + \beta_1 x_{it} + \gamma_i$ appearance takes and suitable respectively for $i \geq 2$ $\alpha_i = \beta_0 + \gamma_i$.

So by doing immutable efficient regression model make up (3) and (4) for equations shaped method apply can In equation (3) , the model values are u or this of objects feature dependent has been *n* different constants through written In equation (4), the model is one in general constant and $n-1$ is binary variables through written

Both method also *x* variable front (angle) coefficients all objects for one different values acceptance does Knowing to put (3) in Eq to himself special constants and (4) in Eq binary variables source one is unobservable different objects for different values acceptance doer but time according to immutable has been z_i is a variable (character).

The model one how many *x* to variables expand

If the effect on *y* doer , *x* with correlated and time according to variable has been another observation factors there is if so , them attention not received because of , to come coming out from errors case to be for them regression to Eq input necessary

Immutable efficient regression model as follows to express can :

$$y_{it} = \beta_1 x_{1,i1} + \ldots + \beta_k x_{k,it} + \alpha_i + u_{it} \quad (9)$$

this on the ground $i = 1, \ldots, n;$ $t = 1, \ldots, T$. $x_{1,i1} - i$ – the value of independent variables included in the first stage of the object at time *t* , $x_{k,it} - i$ - the value of the independent variables entered in the $\alpha_1, \alpha_2, \ldots, \alpha_n$ *k - step of the object* at time *t ;* – constant quantities depending on the object of observation.

The model can be equivalently written using common constants, *x* - independent variables and *n* -1 binary variables:

$$y_{it} = \beta_0 + \beta_1 x_{1,i1} + \ldots + \beta_k x_{k,it} + \gamma_2 D_{2i} + \gamma_3 D_{3i} + \ldots + \gamma_n D_{ni} + u_{it} \quad (10)$$

this on the ground $D_{2i} - i = 2$ when together equal to has been and another cases to zero equal to has been binary variable and etc.

So so , panel data different different questions analysis in doing research of objects dynamics and individual characteristics account to get possibility will give . They are research for strong tool more precisely and reliable to the results to reach possibility will give . From panel data use historical trends more precisely analysis do , cause and effect relationships determination and more precisely forecast to do enable will give .

References

1. Muminova MA Econometrics: Study guide.–T.: "Economics-Finance", 2022.
2. Habibullayev IH, Utanov B. Fundamentals of Econometrics: A Study Guide. - T.: "Economy-Finance", 2018, -172p.
3. Habibullayev I. Econometrics: Textbook . -T.: "Economy-Finance", 20 20 , - 240 p.
4. Habibullayev I., Jumayev A. Econometrics: a tutorial for practical training. -T.: "Economy-Finance", 2020, -176 p.
5. Eliseeva I.I., Kurisheva S.V. i dr. Econometrics: A textbook . - M.: Izdatelstvo Yurayt , 2018 . -2 88 s .
6. Muminova Makhbuba Abduvafo y evna . (2022). Model of correlation-regression analysis of social-economic processes. International journal of social science and interdisciplinary research issn: 2277-3630 impact-factor: 7.429, 11(11), 441-451. izvlecheno iz https :// www. gejournal . net / index . php / ijssir / article / view /1227
7. Mullabayev Bakhtiyarjan Bulturbayevich , Muminova Makhbuba Abduvafoyevna . " Development of methodology for managing the activities of vertically integrated industrial enterprises". International journal of research in commerce, it, engineering and social sciences isn : 2349-7793 impact factor: 6.876., volume: 16 issue: 06 in June 2022.
8. Mullabayev Bakhtiyarjan Bulturbayevich , Muminova Makhbuba Abduvafoyevna . IMPROVEMENT OF MODERN ENERGY INDUSTRIES IN THE CONDITIONS OF DIGITAL ECONOMY. ASIA PACIFIC JOURNAL OF MARKETING & MANAGEMENT REVIEW ISSN: 2319-2836 IMPACT FACTOR: 7.603 Vol 12, Issue 05, 2023.