**Estimating Trends in Indian Food Grain Production based on Panel Regression Model**

**Rajarathinam, A1., and Suba, S.S2.**

1,2Department of Statistics, Manonmaniam Sundaranar University, Tirunelveli – 627 012, Tamil Nadu State, India

rajarathinam@msuniv.ac.in1 and sssuba@gmail.com2

**Abstract**

The present investigation was carried out to study the food grain production trends of different states in India based on the Panel Regression Model for the period 2001-02 to 2020-2021. The results reveal that state-to-state food grain production was highly significant. The highest food grain production was registered in Uttar Pradesh, followed by Punjab and Madhya Pradesh. The very lowest was recorded in Kerala and Himachal Pradesh. The findings reveal that the highly significant fixed effect model was suitable for studying the trends in Indian food grain production, and this model explains the 82 % of variations in food grain production. Overall, an increase in food grain production is noted.

**Keywords:** Panel Regression Model, Least-Squares Dummy Variable, Fixed-Effect Model, Random-Effect Model, Wald Test, Hausman Test.

**Introduction**

Over the last few decades, regression modeling has traditionally been employed in agricultural production prediction and classification. For agricultural planning purposes, decision-makers need simple and reliable estimation techniques for crop production prediction. Multiple regressions, Discriminant analysis, factor analysis, principal component analysis, cluster analysis, and logistic regression analysis are the most commonly used statistical techniques for predicting and classifying agricultural-related production. In agricultural production time series data, the problems of multicollinearity, autocorrelation, and extreme values are unavoidable. In such complex situations, regression models may not provide accurate predictions. Regression models must fulfill regression assumptions such as autocorrelation and multiple colinearities between the independent variables, which causes the estimated regression models to be unfit and the estimated parameter values obtained based on these models to be inefficient. In most agricultural practices, crop production is influenced by various interrelated factors, such as autocorrelation, and it isn't easy to describe their relationships using conventional methods (Zaefizadah *et al.,* 2011).

In this study, the panel data regression model is used to combat the complicated relations and strong autocorrelation in the crop production data.

Panel data is a combination of cross-sectional and time series data. Therefore, using a regression suited to panel data can distinguish between fixed and random effects. Fixed effects are independent of random disturbances, e.g., observations independent of time. Random effects are effects that include random disturbances. Panel data is more informative since it contains more information, but it has to be modeled correctly by considering fixed vs. random effects.

Panel data helps us control the heterogeneity of cross-section units, such as individuals, states, firms, countries, etc., over time. Panel data estimation considers all cross-section units as heterogeneous. It helps us to get an unbiased estimate. There are time-invariant and state-invariant variables that we observe or not. Compared to pure cross-sections and time series, panel data estimation can better identify and measure independent variables' effects on dependent variables that we cannot measure using time series and cross-section data. In addition to this, "Panel data give more informative data, more variability, less colinearity among the variables, more degree of freedom and more efficiency." It is also a better estimation method to study the duration of economic states and the "dynamics of change" over time. Finally, it is a reasonable estimation method to 'construct and test complicated behavioral models' (Baltagi, (2001)).

Based on the above discussion, the present study is aimed to study the trends in food grain production in different states in India over the period 2001-02 to 2020-2021 based on the panel regression model.

**Materials and Methods**

**Materials:** The present investigation was carried out to study India's food grain production trends based on the Panel Regression model. The cross-sectional time series data on food grain production for 2001-02 to 2020-2021 (twenty years) have been collected from the [Reserve Bank of India - Handbook of Statistics on Indian Economy (rbi.org.in)](https://m.rbi.org.in/scripts/AnnualPublications.aspx?head=Handbook%20of%20Statistics%20on%20Indian%20Economy). The food grain production in eighteen states of India viz. Andhra Pradesh, Karnataka, Kerala, Tamil Nadu, Gujarat, Chhattisgarh, Bihar, Haryana, Himachal Pradesh, Jharkhand, Madhya Pradesh, Maharashtra, Odisha, Punjab, Rajasthan, Uttar Pradesh, Uttarakhand, and West Bengal have been considered.

**Methods:** Panel data contain observations of multiple phenomena collected over different time periods for the same group of individuals, units, or entities. In short, econometric panel data are multidimensional data collected over a given period.

A simple panel data regression model is specified as



where are the estimated residuals from the panel regression analysis. Here, Y is the dependent variable, X is the independent or explanatory variable,  the intercept, and slope, i stands for the ith cross-sectional unit and t for the tth month, and X is assumed to be non-stochastic and the error term to follow classical assumptions, namely, .In this study, i, the number of cross-sections is 18 (i=1, 2, 3, 4,…,18), and t=1, 2, 3,…, 20.Detailed discussions of panel data models were given in Hsiao (2003), Greene (2008), and Gujarathi (2017).

**Unit Root Test:** Unit roots for the panel data can be tested using either the Leuin-Llin-Chu, (2002) test or the Hadri (2000)LM stationarity test. The null hypothesis is that panels contain unit roots, and the alternative hypothesis is that panels are stationary. In the results, if the p-value is less than 0.05, then one can reject the null hypothesis and accept the alternative hypothesis. Similarly, the unit root for the first difference can also be tested using a similar method.

**Constant Coefficients Model: The** Constant Coefficients Model (CCM) assumes that all coefficients (intercept and slope) remain unchanged across cross-sectional units and over time. In other words, the CCM ignores panel data's space and time dimensions. Put differently, under the CCM, and the cross-sectional units are assumed to be homogeneous such that the values of intercept and slope coefficients are the same irrespective of the cross-sectional unit being considered. Accepting this homogeneity assumption (also called pooling assumption), the CCM uses the panel (or pooled) data set and applies the Ordinary Least Squares (OLS) method to estimate unknown parameters of the model. Thus, the CCM is nothing but the straightforward application of OLS to a given panel or pooled data to obtain estimates for unknown parameters of the model (Bhaumik, (2017)).

**Individual Specific-Effect Model:** Here, it is assumed that there is unobserved heterogeneity across individuals and captured by . The main question is whether the individual-specific effects  are correlated with the regressor; if they are correlated, a fixed effects model exists. If these factors are not correlated, a random effects model exists.

**Fixed-Effect OR Least-Square Dummy Regression Model:** Fixed effect regression model indicates that each unit has its intercept. There will be heterogeneity among the unit due to individual intercepts. Here in the fixed effect model, the unit intercepts are time-invariant (do not vary over time) even if they might differ among cross-section units. However, the fixed effect model believes that the coefficients of the independent variables do not vary across cross-section units or over time.These fixed effects models can be implemented with the dummy variable technique. Therefore, the fixed effects model can be written as



where =1 if the observation is from Karnataka State and is 0 otherwise, =1 if the observation is from Kerala and is 0 otherwise, and =1 if the observation is from Tamil Nadu and is 0 otherwise, etc... Here,  the intercept of Andhra Pradesh and α2, α3, and α4 are different intercept coefficients that indicate how much the intercepts of Karnataka, Kerala, and Tamil Nadu differ from that of Andhra Pradesh state. Since the dummies are used to estimate the fixed effects, the model is also known as the least-squares dummy variable (LSDV) model; hence, one can conclude that the restricted panel regression model is invalid and that the LSDV model is valid (Bhaumik, (2017)).

**Random-Effect (RE) Model:** Random effects model is called the error component model (ECM). In this model, the cross-section units will have a random intercept instead of a fixed intercept. The rationale behind the random effects model is that, unlike the fixed effect model, the variation across entities is assumed to be random and uncorrelated with the predictor or independent variables included in the model, the crucial distinction between the fixed and random effects is whether the unobserved individual effects embody elements that are correlated with regressors in the model, not whether these effects are stochastic or not (Green, 2008)**.** The RE model assumes that individual-specific effects  are random, and one should include  them in the error term. Each cross-section has the same slope parameters and a composite error term. So model (1) becomes Random-Effect Model (REM) :



Let.

Here , and are normally distributed with zero means and constant variances  , respectively.

Hence: , and ; therefore, .

Rho is the interclass correlation of the error or the fraction of the variance in the error term due to individual-specific effects. These variable approaches 1 if individual effects dominate the idiosyncratic error (Bhaumik,(2017)).

**Hausman test:** The Hausman test (Hasman, 1978) is the standard procedure used in empirical panel data analysis to distinguish between fixed and random effects. In the Hausman test, the null hypothesis signifies no significant difference in the estimator of the fixed effect model and random effect model. If we reject the null hypothesis, the fixed effect model will be appropriate. Rejecting the null hypothesis shows that there might be a correlation between the error term and the dependent variable. The test statistic can be calculated given as follows:

****

Here is the vector of random and fixed effects parameter estimates, respectively. Under the null hypothesis, this statistic has asymptotically the chi-squared distribution with the number of degrees of freedom equal to the rank of the matrix :



**Wald Test:** The Wald test (Wald, 1943) can determine which model variables make significant contributions. The **Wald test** (also called the Wald chi-squared test) is a way to determine if explanatory variables in a model are significant, meaning that they add something to the model; variables that add nothing can be deleted without affecting the model in a meaningful way. The test can be used for many models, including those with binary or continuous variables. The null hypothesis for the test is *some parameter* = *some value*.

**Breusch-Pagan Lagrange Multiplier Test**: The Breusch-Pagan-Godfrey test (Breusch and Pagan, (1980)) is a Lagrange multiplier test of the null hypothesis of no heteroskedasticity, i.e., constant variance among residuals.

**Ho:**The test null hypothesis states that there is constant variance among residuals.

**Results and Discussion**

The results obtained in this paper based on applying different statistical tools related to panel regression models are discussed in subsequent sections.

**Summary Statistics:** The descriptive statistics results presented in Table 1 and Fig.1., reveal that state-wise food grain production is normally distributed as indicated by Jarque-Bera statistic's p-values except for the states of Himachal Pradesh and Odisha. The highest food grain production was registered in Uttar Pradesh, followed by Punjab and Madhya Pradesh. On the contrary, the lowest production is reported in Kerala and Himachal Pradesh. Fig.2 depicts that the highest food grain production is registered during 2018-19, and the year-wise production shows an increasing trend.

Table 1: State-wise Food Grain Production Details

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Sr.No. | Name of State | Sum | Mean | Max. | Mini. | S.D | Jarque-Bera Prob. |
| 1 | Andhra Pradesh | 303003.00 | 15150.15 | 20421.00 | 10365.40 | 3572.47 | 0.4094 |
| 2 | Bihar | 253527.60 | 12676.38 | 17036.90 | 7704.40 | 2737.38 | 0.6372 |
| 3 | Chhattisgarh | 129993.00 | 6499.65 | 9324.10 | 3274.70 | 1366.25 | 0.8944 |
| 4 | Gujarat | 139058.80 | 6952.94 | 9179.60 | 3566.30 | 1446.41 | 0.7472 |
| 5 | Haryana | 314500.40 | 15725.02 | 18274.20 | 12328.90 | 1886.79 | 0.4766 |
| 6 | Himachal Pradesh | 29545.90 | 1477.30 | 1740.60 | 1017.20 | 170.97 | 0.0408 |
| 7 | Jharkhand | 75343.20 | 3767.16 | 6001.30 | 1876.60 | 1301.33 | 0.5478 |
| 8 | Karnataka | 220384.00 | 11019.20 | 14187.00 | 6562.10 | 2081.62 | 0.4980 |
| 9 | Kerala | 11739.30 | 586.97 | 718.90 | 439.00 | 70.08 | 0.9303 |
| 10 | Madhya Pradesh | 425622.00 | 21281.10 | 33523.10 | 10748.80 | 8469.48 | 0.3126 |
| 11 | Maharashtra | 246466.90 | 12323.35 | 16069.20 | 8754.40 | 2034.96 | 0.6186 |
| 12 | Odisha | 151779.00 | 7588.95 | 9459.00 | 3573.70 | 1280.89 | 0.0013 |
| 13 | Punjab | 551900.70 | 27595.04 | 31691.90 | 23491.20 | 2260.11 | 0.7910 |
| 14 | Rajasthan | 342927.50 | 17146.38 | 24313.10 | 7536.00 | 4165.03 | 0.6823 |
| 15 | Tamil Nadu | 159365.50 | 7968.28 | 11478.50 | 4141.60 | 2423.01 | 0.5525 |
| 16 | Uttar Pradesh | 931735.50 | 46586.78 | 58313.30 | 37836.30 | 5906.19 | 0.5973 |
| 17 | Uttarakhand | 35634.70 | 1781.74 | 2003.50 | 1559.10 | 107.42 | 0.9087 |
| 18 | West Bengal | 333035.80 | 16651.79 | 20106.70 | 14466.90 | 1250.53 | 0.0892 |

 **Fig.1. State-wise Food Grain production**

**Fig.2. Year-wise Food Grain Production**

The results in Table 2 reveal that the ANOVA F-test and Welch F-test statistic values are significant, indicating that the production is statistically significant in all the states.

 **Table 2: Analysis of Variance test for equality of production means**

|  |  |  |  |
| --- | --- | --- | --- |
| Method | df | Value | Probability |
| ANOVA F-test | (17, 342) | 251.0402 | 0.0000 |
| Welch F-test\* | (17, 124.252) | 744.1693 | 0.0000 |
| \*Test allows for unequal cell variances |  |

|  |
| --- |
| **Analysis of Variance** |
| **Source of Variation** | **df** | **Sum of Sq.** | **Mean Sq.** |
| Between | 17 | 4.18E+10 | 2.46E+09 |
| Within | 342 | 3.35E+09 | 9785232. |
| Total | 359 | 4.51E+10 | 1.26E+08 |

**Unit Root Test:** Before estimating Panel Data Regression Model, it is necessary to determine the stationarity of the variable under study. The Unit root test result presented in Table 3 reveals that since the Levin, Lin & Chu t statistics values are significant at 1 % level of significance since the p-value is very low and hence the study variable, PROD, is stationary at the level and therefore the variable is I(0).

**Table 3: Characteristics of the unit root test**

|  |  |  |
| --- | --- | --- |
|  **Method** | **Individual Effect** | **Individual effects, linear trends** |
| **Statistic** | **Prob.\*\*** | **Statistic** | **Prob.\*\*** |
| Levin, Lin & Chu t\* |  3.05481 |  0.0011 | 9.12116 | 0.0000 |

\*\* Probabilities are computed assuming asymptotic normality.

**Constant Coefficient Model(Panel OLS).** The CCM es method is employed considering the food grain production (PRODN) as the dependent variable and X, time, as the independent variables; the results are presented in Table 4. The result reveals that the intercepts and slopes are positive and highly significant at the 1% significance level. The positive slope indicates that food grain production is an increasing trend. The model is highly significant at the1% level of significance with an incredibly low R2 value of 15%, which is very low. Additionally, the estimated Durbin-Watson value of 0.043090 is relatively low, suggesting autocorrelation in the data.

**Table 4: Characteristics of the fitted panel least-squares method**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | Coefficient | Std. Error | t-Statistic | Prob.   |
| C | 5384.438 | 1092.674 | 4.927761 | 0.0000 |
| X | 41.81541 | 5.246203 | 7.970603 | 0.0000 |
| Root MSE | 10315.66 |     R-squared | 0.150714 |
| Mean dependent var | 12932.12 |     Adjusted R-squared | 0.148342 |
| S.D. dependent var | 11209.18 |     S.E. of regression | 10344.43 |
| Akaike info criterion | 21.33182 |     Sum squared resid | 3.83E+10 |
| Schwarz criterion | 21.35341 |     Log-likelihood | -3837.728 |
| Hannan-Quinn criteria. | 21.34041 |     F-statistic | 63.53052 |
| Durbin-Watson stat | 0.043090 |     Prob(F-statistic) | 0.000000 |

The estimated model assumes that the slope coefficients of time variables X are all identical in all four states. Therefore, despite its simplicity, the CCM may distort the relationship between the dependent variable— food grain production (PRODN)**—**and time, the independent variable X, across the four districts.

**Fixed-Effect OR Least-Square Dummy Variable Regression Model:** The result presented in Table 5 reveals that the fixed effect model explains the 82 % of variations in the dependent variable. The model is highly significant at 1 % level of significance. All the dummy variables were also highly significant at 1 % significance level. The root means square error value is 2609.55, with the S.E. of regression being 2681.27.

Based on the statistical significance at the 1% level of significance of the estimated coefficients and the substantial increase in the R2 value to 95% (significant at the 1% level of significance), one can conclude that the fixed effects model or the LSDSV regression model performs better than the panel least-squares regression model (CCM).

**Table 5: Characteristics of the fixed effects or LSDSV regression model**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Coefficient** | **Std. Error** | **t-Statistic** | **Prob.** |
| C(1) | 12278.95 | 652.44 | 18.82 | 0.0000 |
| C(2) | 273.45 | 24.51 | 11.15786 | 0.0000 |
| C(3) | -9599.90 | 979.37 | -9.802151 | 0.0000 |
| C(4) | -25501.08 | 1296.10 | -19.67522 | 0.0000 |
| C(5) | -23588.71 | 1697.37 | -13.89718 | 0.0000 |
| C(6) | -30073.00 | 2136.06 | -14.07871 | 0.0000 |
| C(7) | -35995.23 | 2593.25 | -13.88038 | 0.0000 |
| C(8) | -35287.45 | 3060.65 | -11.52941 | 0.0000 |
| C(9) | -37707.76 | 3534.22 | -10.66934 | 0.0000 |
| C(10) | -57424.43 | 4011.77 | -14.31400 | 0.0000 |
| C(11) | -60603.51 | 4492.03 | -13.49133 | 0.0000 |
| C(12) | -48558.52 | 4974.23 | -9.762025 | 0.0000 |
| C(13) | -62985.22 | 5457.84 | -11.54033 | 0.0000 |
| C(14) | -73188.56 | 5942.52 | -12.31609 | 0.0000 |
| C(15) | -58651.42 | 6428.02 | -9.124332 | 0.0000 |
| C(16) | -74569.03 | 6914.19 | -10.78493 | 0.0000 |
| C(17) | -50597.57 | 7400.87 | -6.836703 | 0.0000 |
| C(18) | -100871.60 | 7887.99 | -12.78799 | 0.0000 |
| C(19) | -91470.45 | 8375.46 | -10.92125 | 0.0000 |
| Root MSE | 2609.550 |     R-squared | 0.945651 |
| Mean dependent var | 12932.12 |     Adjusted R-squared | 0.942782 |
| S.D. dependent var | 11209.18 |     S.E. of regression | 2681.265 |
| Akaike info criterion | 18.67730 |     Sum squared resid | 2.45E+09 |
| Schwarz criterion | 18.88240 |     Log-likelihood | -3342.914 |
| Hannan-Quinn criteria. | 18.75885 |     F-statistic | 329.6257 |
| Durbin-Watson stat | 0.667659 |     Prob(F-statistic) | 0.000000 |

The cross-section fixed effects (as deviations from common intercept) in the context of the fixed effect model are calculated and presented in Table 6. The fixed effects are positive in Andhra Pradesh, Karnataka, Kerala, Tamil Nadu, Gujarat, Chhattisgarh, Bihar, Haryana, and Madhya Pradesh. The fixed effect is negative in Himachal Pradesh, Jharkhand, Maharashtra, Odisha, Punjab, Rajasthan, Uttar Pradesh, Uttarakhand, and West Bengal have been considered. In Andhra Pradesh, the fixed effect is 48704.08, the highest in other states.

**Table 6: Cross-Section Fixed Effects Values**

|  |  |  |
| --- | --- | --- |
| **Sr.No.** | **CROSSID** | **Effect** |
| 1 |  1 | 48704.08 |
| 2 |  2 | 39104.18 |
| 3 |  3 | 23203.00 |
| 4 |  4 | 25115.36 |
| 5 |  5 | 18631.08 |
| 6 |  6 | 12708.84 |
| 7 |  7 | 13416.63 |
| 8 |  8 | 10996.32 |
| 9 |  9 | -8720.35 |
| 10 |  10 | -11899.43 |
| 11 |  11 | 145.56 |
| 12 |  12 | -14281.14 |
| 13 |  13 | -24484.48 |
| 14 |  14 | -9947.343 |
| 15 |  15 | -25864.95 |
| 16 |  16 | -1893.496 |
| 17 |  17 | -52167.48 |
| 18 |  18 | -42766.37 |

The diagrammatic representation of fixed effects in all eighteen states is depicted in Fig,3. Based on this result, the Fixed effect model is better than CCM.

**Fig.3. Fixed effect in different states**

The Redundant Fixed Effect test has been carried out to confirm the presence of Fixed Effect, and the results are presented in Table 7. The test results reveal that the Cross-section F and Chi-square statistics values are significant at 1 % level of significance, indicating that the presence of fixed effects is different from one state to another.

**Table 7: Results of Redundant Fixed Effects Test**

|  |  |  |  |
| --- | --- | --- | --- |
| Effects Test | Statistic   | d.f.  | Prob.  |
| Cross-section F | 293.390789 | (17,341) | 0.0000 |
| Cross-section Chi-square | 989.629016 | 17 | 0.0000 |
| PRODN=C(1)+C(2)\*X |  |  |
|  | Coefficient | Std. Error | t-Statistic | Prob.   |
| C(1) | 5384.438 | 1092.674 | 4.927761 | 0.0000 |
| C(2) | 41.81541 | 5.246203 | 7.970603 | 0.0000 |
| Root MSE | 10315.66 |     R-squared | 0.150714 |
| Mean dependent var | 12932.12 |     Adjusted R-squared | 0.148342 |
| S.D. dependent var | 11209.18 |     S.E. of regression | 10344.43 |
| Akaike info criterion | 21.33182 |     Sum squared resid | 3.83E+10 |
| Schwarz criterion | 21.35341 |     Log-likelihood | -3837.728 |
| Hannan-Quinn criteria. | 21.34041 |     F-statistic | 63.53052 |
| Durbin-Watson stat | 0.043090 |     Prob(F-statistic) | 0.000000 |

**Wald Test:** The Wald test has been carried out to compare the fixed effect model with CCM. The null hypothesis of the Wald test is H0=C(3)=C(4)=C(5)…=C(18)=0, i.e., all three dummy variable values are zero (there is no fixed effect). The result presented in Table 8 reveals that, since the F and Chi-square statistics values are significant at 1 % level of significance, the null hypothesis H0=C(3)=C(4)=C(5)=…=C(18)=0 is rejected which indicates that the values of the dummy variables are not equal to zero which confirms fixed effects or LSDV regression model is an appropriate model in comparisons to CCM.

**Table 8: Characteristics of the Wald test**

|  |  |  |  |
| --- | --- | --- | --- |
| Test Statistic | Value | df | Probability |
| F-statistic |  292.7218 | (16, 341) |  0.0000 |
| Chi-square |  4683.549 |  16 |  0.0000 |

**Random-Effect Model**: Finally, the random-effect model is estimated, and the results are presented in Table 9. The result reveals that the model is highly significant at 1 % level of significance with a low R2 value of 17 % with S.E. of regression 2845.515, RMSE, 2837.60. As in the case of the fixed effect model, the random-effect model's intercept and slope are highly significant at 1 % significance. The rho value is 0.9385, indicating that the cross-section's individual effects are 0.9%.

**Table 9: Characteristics of the fitted random effects model**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variable** | **Coefficient** | **Std. Error** | **t-Statistic** | **Prob.** |
| C | -14867.17 | 3952.539 | -3.761424 | 0.0002 |
| X | 154.0127 | 17.08422 | 9.014909 | 0.0000 |
| **Effects Specification** |
|  |  |  | **S.D.** | **Rho** |
| Cross-section random | 10472.86 | 0.9385 |
| Idiosyncratic random | 2681.265 | 0.0615 |
| **Weighted Statistics** |
| Root MSE | 2837.599 |     R-squared | 0.167746 |
| Mean dependent var | 739.1265 |     Adjusted R-squared | 0.165422 |
| S.D. dependent var | 3114.779 |     S.E. of regression | 2845.515 |
| Sum squared resid | 2.90E+09 |     F-statistic | 72.15735 |
| Durbin-Watson stat | 0.565478 |     Prob(F-statistic) | 0.000000 |
| **Unweighted Statistics** |
| R-squared | -0.934325 |     Mean dependent var | 12932.12 |
| Sum squared resid | 8.73E+10 |     Durbin-Watson stat | 0.018787 |

The cross-section random effects in the context of the random effect model are calculated and presented in Table 10. The random effects are positive in Andhra Pradesh, Karnataka, Kerala, Tamil Nadu, Gujarat, Chhattisgarh, Bihar, Haryana, Madhya Pradesh, and Punjab. On the other hand, the random effects are negative in Himachal Pradesh, Jharkhand, Maharashtra, Odisha, Rajasthan, Uttar Pradesh, Uttarakhand, and West Bengal have been considered.

**Table 10: Cross-Section Random Effects Values**

|  |  |  |
| --- | --- | --- |
| **Sr.No.** | **CROSSID** | **Effect** |
| 1 |  1 |  28307.42 |
| 2 |  2 |  21119.77 |
| 3 |  3 |  7651.421 |
| 4 |  4 |  11938.43 |
| 5 |  5 |  7856.217 |
| 6 |  6 |  4334.215 |
| 7 |  7 |  7420.576 |
| 8 |  8 |  7389.066 |
| 9 |  9 | -9882.309 |
| 10 |  10 | -10670.12 |
| 11 |  11 |  3716.420 |
| 12 |  12 | -8282.265 |
| 13 |  13 | -16071.39 |
| 14 |  14 |  799.1541 |
| 15 |  15 | -12685.57 |
| 16 |  16 |  13588.47 |
| 17 |  17 | -34140.40 |
| 18 |  18 | -22389.11 |

The diagrammatic representation of random effect in all eighteen states is depicted in Fig,4. Based on this result, the presence of random effects in all four different districts is confirmed.

**Fig.4. Random effect in different districts**

**Breusch-Pagan Lagrange-Multiplier Test (Heteroskedasticity Test):** It is well known that heteroskedasticity in the disturbances of an otherwise properly specified linear model leads to consistent but inefficient parameter estimates and inconsistent covariance matrix estimates. As a result, faulty inferences will be drawn when testing statistical hypotheses in the presence of heteroskedasticity14.

The result presented in Table 11 indicates that the Breusch-Pagan LM, Pesaran scaled LM, and Pesaran CD tests statistic values are highly significant at 1 % level of significance since both statistics p-values are equal to 0.0000, indicating that the null hypothesis of the test, "H0:There is constant variance among residuals" is rejected. Hence, the above random effect model has the problem of heteroscedasticity.

**Table 11: Characteristics of the residual cross-section dependence test**

|  |  |  |  |
| --- | --- | --- | --- |
| **Test** | **Statistic** | **d.f.** | **Prob.** |
| Breusch-Pagan LM | 566.0751 | 153 | 0.0000 |
| Pesaran scaled LM | 23.61393 |  | 0.0000 |
| Pesaran CD | 3.092296 |  | 0.0020 |

The Hausman test result presented in Table 12 reveals that the Chi-Sq.statistics value of 0.0000 with 1 degree of freedom is highly significant at 1 % significance value; the null hypothesis "H0:Random Effect Model" is rejected. So, among the three models, viz. CCM, Fixed effect, and Random effect model, the fixed effect model emerged as an appropriate model.

**Table 12: Characteristics of the Hausman Test**

|  |  |  |  |
| --- | --- | --- | --- |
| Test Summary | **Chi-Sq. Statistic** | **Chi-Sq.D.F.** | **Prob.** |
| Cross-Section Random | 46.2043 | 1 | 0.0000 |

The following fig 5. depicts and confirms that the coefficients of intercept and slope are lying in the 99 % Confidence Interval (CI)

**Conclusion**

The present investigation was carried out to study the food grain production trends in different states in India based on Panel Regression Model for the period 2001-02 to 2020-2021. The result reveals that state-to-state food grain production is highly significant. The highest food grain production was registered in Uttar Pradesh, followed by Punjab and Madhya Pradesh. The very lowest was recorded in Kerala and Himachal Pradesh. The fixed effect model was found to study the trend, and this model explains the 82 % of variations in food grain production. Increases in food grain production have been observed.

**REFERENCES**

1. Baltagi, B.H.,(2021). Econometric Analysis of Panel Data.6th Edition, Wiley, New York, NY, [https://link.springer.com/book/10.1007%2F978-3-030-53953-5](https://link.springer.com/book/10.1007/978-3-030-53953-5).
2. Bhamik, S.K., (2017). *Principles of Econometrics*, 2nd edition, Oxford University Press.
3. Breusch, T., and A. Pagan (1980).The Lagrange Multiplier Test and its Application to Model Specification in Econometrics, Review of Economic Studies, 47, 239–253.
4. Greene, W.H. (2008). Econometric Analysis, 2008, Pearson Prentice Hall, Upper Saddle River, NJ.
5. Gujarati, D.N., D.C.Porter, G.Sangeetha,(2017). *Basic Econometrics*, 5th edition. McGraw Hill Education, New York, NY.
6. Hadri, K., 2000.Testing for Units Roots in Heterogeneous Panel Data*.* Econometrics Journal, 3: 148-161.
7. Hausman, J. A. (1978).Specification Tests in Econometrics, *Econometrica*, 46, 1251–1272.
8. Hsiao, C., (2003). *Analysis of Panel Data*. Cambridge University Press, Cambridge.
9. Levin, A., C.F. Lin, and C.S.J.Chu,(2002).Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties. Journal of Econometrics, 108:1-24.
10. Wald, A (1943).Test of statistical hypotheses concerning several parameters when the number of observations is large, Transactions of the American Mathematical Society, 54,426-482, http://www.jstor.org/stable/1990256
11. Zaefizadah M., Khayatnezhad M.,and Ghlaomin M. (2011). Comparison of Multiple Linear Regression (MLR) and Artificial Neural Network (ANN) in Predicting the Yield Using its Components in the Hulless Barley, J. Agriculture, and Environmental Sciences, 10(1), 60-64.