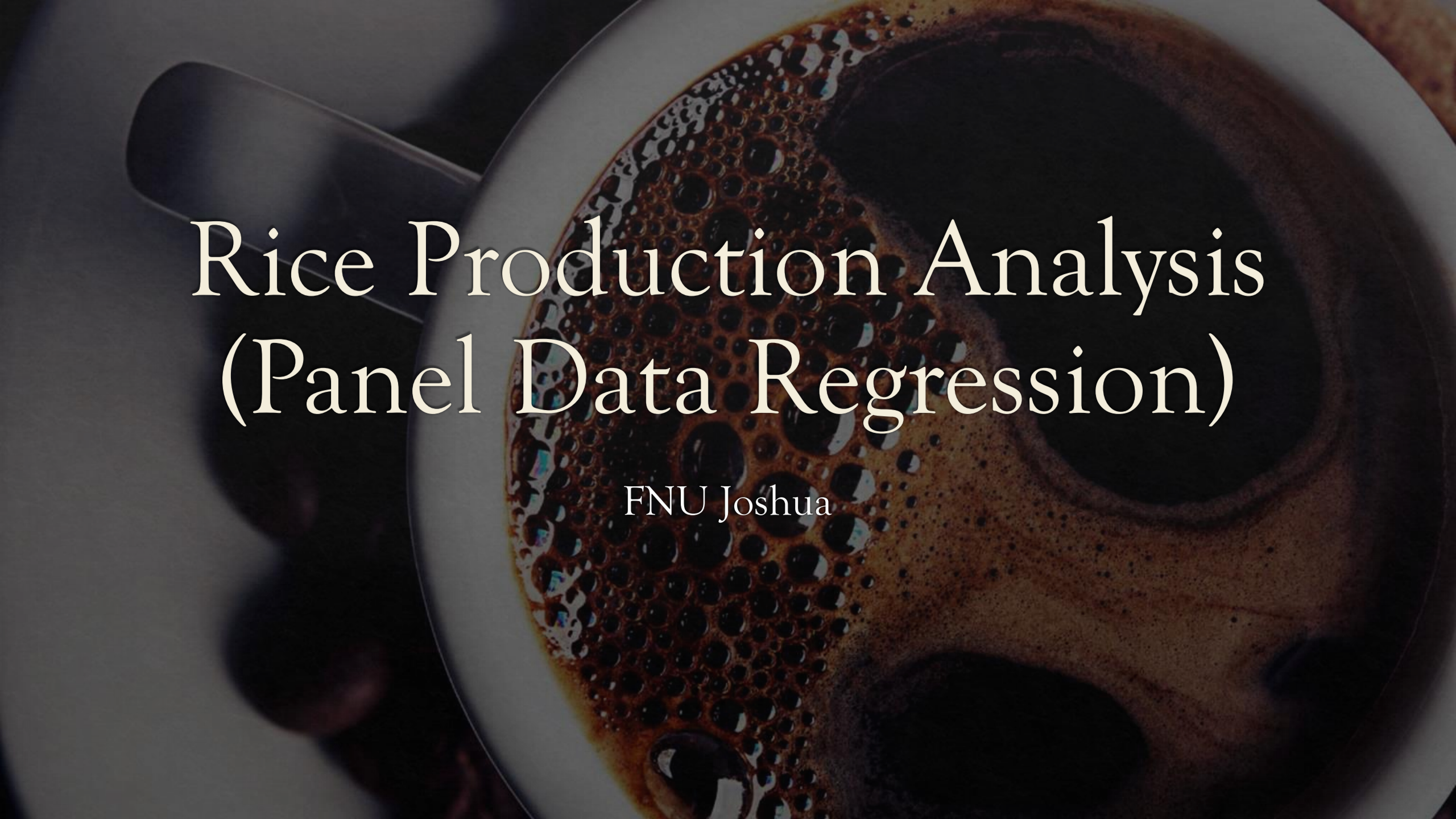

Rice Production Analysis (Panel Data Regression)

FNU Joshua

RESEARCH BACKGROUND

01

WOULD LIKE TO **STUDY CONTROL VARIABLES** THAT ARE CRUCIAL IN THE PRODUCTION OF **RICE**.

02

IT IS USEFUL FOR AGRIBUSINESS RICE ENTREPRENEUR TO RECEIVE SOME INSIGHT FROM THE QUANTITATIVE RESEARCH ABOUT THEIR PRODUCT.

MODEL EQUATION: $PROD_{it} = \beta_1 + \beta_2 * AREA_{it} + \beta_3 * FERT_{it} + \beta_4 * LABOR_{it} + u_i + e_{it}$

i	: indicates each different observations
t	: indicates each different time periods
YEAR	: Year = 1993 to 1994
PROD	: Rice production (tons)
AREA	: Area planted to rice (hectares)
LABOR	: Hired + family labor (person days)
FERT	: Fertilizer applied (kilograms)

a, b, c, d

ISSUES AND SUITABLE METHODS

1. Unobserved Heterogeneity: There might be some factors that were not captured in the data which led to model bias. [i.e. $E(B_2) \neq \beta_2$, instead $E(B_2) = \beta_2 + \beta_3 * E((COV(X,Z)/VAR(X))]$
   - ✓ Use Fixed Effect Model to factor out unobserved heterogeneity from the equation. ($PROD_{IT} = \beta_2 * AREA_{IT} + \beta_3 * FERT_{IT} + \beta_4 * LABOR_{IT} + E_{IT}$)

2. Time Variant Variable: Consider whether the variables that are being observed is consistent over time or not. (i.e. $X_{21} \neq \ldots\ldots \neq X_{2T}$, then error not consistent. $VAR(V_{IT}) = \Psi_{IT}\ (VAR_U + VAR_{IT})$
   - ✓ Use random effect model to take care of the heterogeneity caused by time variation. ($PROD_{IT} = \beta_1 + \beta_2 * AREA_{IT} + \beta_3 * FERT_{IT} + \beta_4 * LABOR_{IT} + E_{IT}$)
   - ✓ Cluster Standard Error and Feasible GLS

e, g

Results & Findings

1. Run a Lagrange Multiplier test, the results shows Pooled OLS Model to be appropriate than Random Effect Model.
   - Implications: None of the variable in the data sets are time variant.

2. Run a Fischer test, the result shows that Fixed Effect Model is a more appropriate model than the Pooled OLS Model.
   - Implications: We can conclude from the two test that the case of Unobserved Heterogeneity is present

3. Run a Hausman test, (basically the same implication as above), Fixed Effect is more appropriate than the Random Effect Model.

Conclusion:

Hence, we know that there are some unobserved variable in the error components (u_i) that cannot be ignored. One of the possible cause of this is land quality, weather conditions, etc.

h

Final Model (Fixed Effect Model)

Within Estimator

```{r}
plm3 <- plm(prod ~ area + fert + labor, data = rice, model = "within")
summary(plm3)
```

```
Oneway (individual) effect Within Model

Call:
plm(formula = prod ~ area + fert + labor, data = rice, model = "within")

Balanced Panel: n = 44, T = 2, N = 88

Residuals:
    Min.  1st Qu.   Median  3rd Qu.     Max.
-3.17932 -0.39825  0.00000  0.39825  3.17932

Coefficients:
        Estimate Std. Error t-value Pr(>|t|)
area   4.5312237  1.8920594  2.3949  0.02128 *
fert  -0.0103367  0.0045189 -2.2874  0.02740 *
labor  0.0135664  0.0108567  1.2496  0.21854
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    131.9
Residual Sum of Squares: 103.13
R-Squared:      0.21809
Adj. R-Squared: -0.65919
F-statistic: 3.81179 on 3 and 41 DF, p-value: 0.016897
```

Fixed Effect model:

We can see that the R-Squared is low (21.809%). This supports our claim that there are some statistically significant variable in the error component that is not captured in the model.

In order to improve this model, we need to collect and formulate another variable metric to be considered into the model.

h

Thank you!
Any questions?