



Improving Agriculture Statistics at sub-state level from NSS data

Issues and Challenges

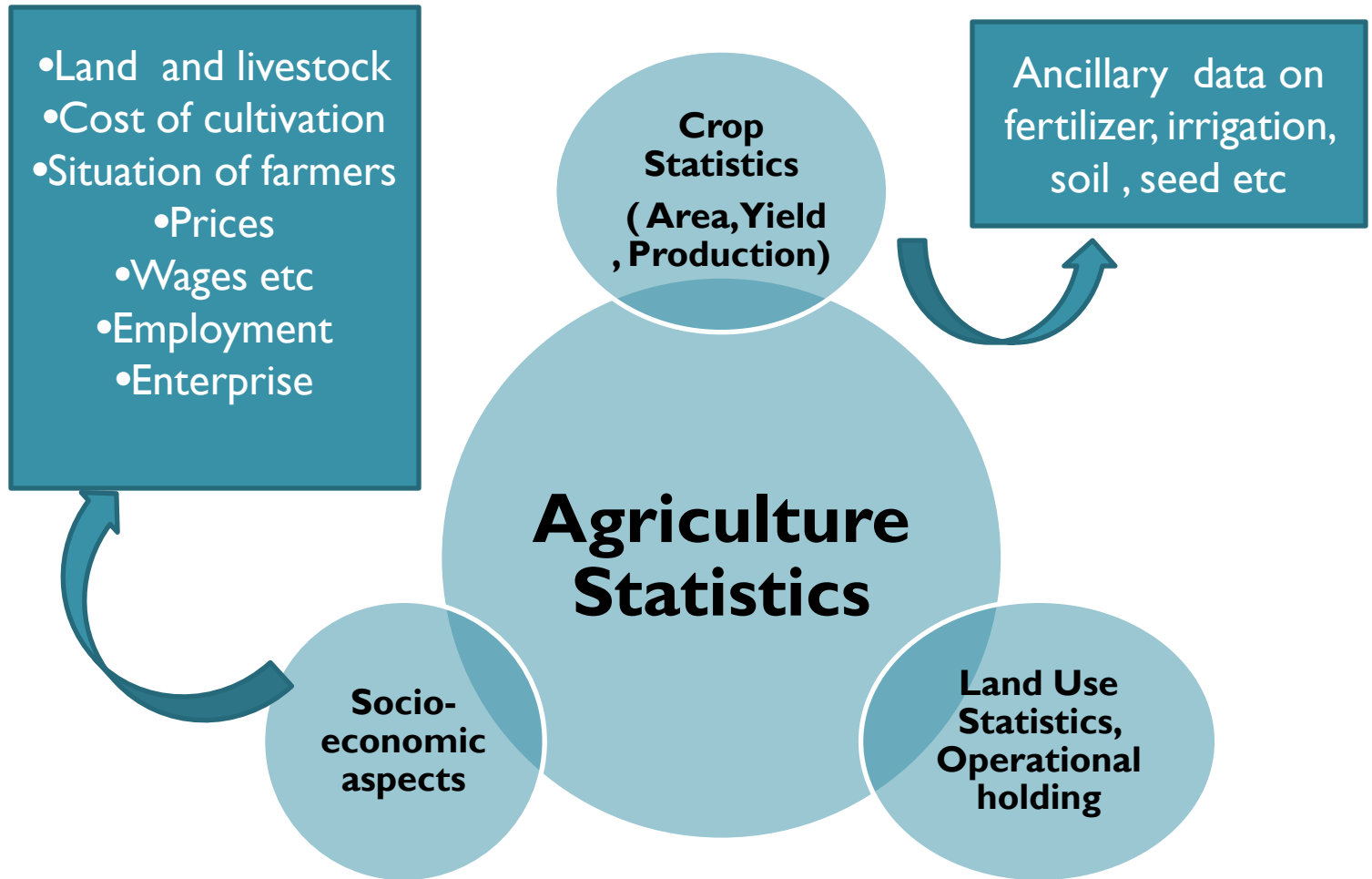
**A Presentation
by
Directorate of Economics and Statistics, Odisha
24th COCSSO, 19-20th January, 2017**

Decentralized Planning for Agriculture Development

- **Diverse agriculture scenario**
 - Fragmentation of operational holdings
 - Diverse agro-climatic condition
 - Diverse socio economic condition
 - Scarce resources (land, water etc)
- **73rd and 74th amendment: Local Governance**
- **11th plan: Thrust on decentralized planning and district plan process**
- **12th plan: Implementation of major programmes for inclusive agriculture growth**
- **Demand for data at lower level(Local level statistics)**

•Need: more productivity
•Calls for location specific plans
•Planning and implementation to be decentralized

Domain of Agriculture Statistics



Domain of Agriculture Statistics

Crop Statistics

- EARAS for Odisha
- Data from DAFP, Odisha
- **Availability of data at district level**

Land Use Statistics and Operational holding

- DES, Odisha/DAFP, Odisha
- Agriculture Census
- **Availability of data at district/ block level**

Socio-Economic Aspects

- NSS
- **Availability of data at state level**
- Price Statistics
- Agriculture labour wage
- **Availability of data at district level**

Agriculture Statistics and NSS

NSS surveys

- Land Holdings & Livestock Holdings
- AIDIS
- Situation Assessment of Agriculture Households
- Employment Survey (employment by broad industry division)

Agriculture Statistics and NSS

- Available of data at **State level**
- Due to low sample size, **no estimation at sub-state level**
- Pooling of central and state sample data for **district level estimates**

Initiatives for getting sub-state level estimate

- Technical committee report of Shri S. C. Chaudhury, 1983 (central and state samples may be combined as a weighted average with the number of primary sampling units at stratum level)
- Minhas and Sardana (1990) paper (the same as that recommended by technical group headed by S. C. Chaudhury in 1983)
- National Statistical Commission(NSC) ,2010

NSC Recommendation

- **Testing poolability:**
- To test whether both data (State and Central sample data) are coming from same population
- **Procedure for pooling If centre and state samples are identical**

Assumption and tools for testing

- Observations are independently identically distributed(iid)
- Classical non parametric tests like **Run test, K-S test, Median test** may be used for testing similarity of population represented by centre and state sample

Bottlenecks in using classical tests in Survey data

iid	Survey data
Based on occurrence	Based on selection
Each observation is independent (like coin tossing).	Sample data are often clustered due to the use of multistage cluster samples. Observations within the same cluster are correlated.
Governed by same distribution function even in non parametric case	Concept of distribution function not used
Infinite population (whole popn not possible)	Finite population (whole popn is possible if sample is large)
Sampling distn of commonly used test statistics are available	Obtaining sampling distribution of traditional test statistic is very difficult

Generally, It is not advisable to apply statistical tests to NSS data. Therefore use of non- parametric tests like Run test, Median test etc is not appropriate. (In Super population approach, such assumption can be applied)

Poolability test for NSS data: a simple technique

- **Confidence interval**
 - Estimates of given characteristics of two data sets should lie within the confidence interval of each other)
- If it fails, both estimates are considered significantly different (pooled estimate will be misleading)
- Discussion in para 3.2.2 of NSC report (pooling) is not clear

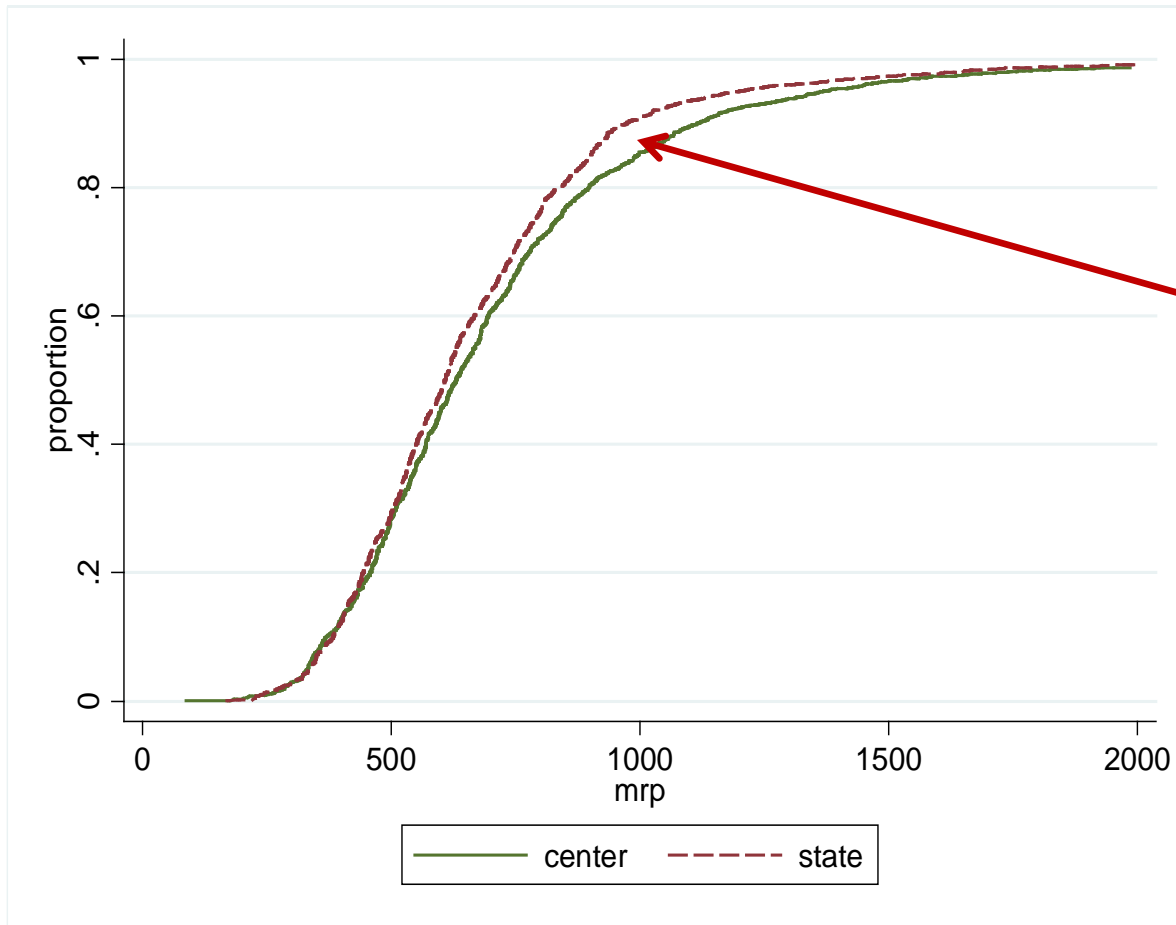
Empirical Study on Consumer Expenditure of 66th Round NSS(Rural)

- MPCE based on MRP(rural): variable for testing. Result of MPCE estimate, SE and CI

Agency	Sample Size	MPCE (Rs)	SE	95% CI
Central	2975	715	12.6	690-740
State	2976	672	8.11	655-690

- MPCE central sample does not lie within 95% CI of State sample and vice versa
- Hence both data sets should not be pooled

Graphs of Cumulative distributions based on central and state sample(66th round NSS)



Systematic
difference
in centre
and state
sample
data

Use of NSC recommendation (run test)

- Null hypothesis is not rejected. It means distributions of MPCE are similar i.e., central and state sample represent identical distribution.
- Why it happens?

Solution for Complex Survey data

- Major challenge for applying **run** test is to create iid data from state and centre sample and use this to estimate runs
- **Reshampling/simulation technique**

Way forward

- Standardize of reshampling / simulation technique
- Standardize of tests
- NSC may kindly consider using these tools like reshampling / simulation technique for testing of survey data
- Capacity building of DES
- Pooling of NSS data to fill data gaps at district level

- 
- **Thank You All**