

Stochastic Optimization of Food Subsidy Allocation Using NSSO Household Consumption Data

Author:

Ridhwan Choudhari (BSD-CC-2417)

Rajneesh Yadav (BSD-CC-2415)

Supervised By:

Prof. Kaushik Jana

In Partial Fulfillment of the requirements for the Degree of Bachelor Of
Statistical Data Science

Contents

1	Problem	2
2	PLAN	3
2.1	Model Development	3
2.2	Methodology	3
3	DATA	4
3.1	Data Source and Study Design	4
3.2	Data Collection	4
3.3	Data Description	5
3.4	Data Cleaning and Processing	5
4	Analysis	6
4.1	Optimization Model	6
4.2	Optimization Method: SLSQP	6
4.3	Plotting And Inferences	7
5	Conclusion	8
5.1	Key Insights:	8
5.2	Implications for Policy	8
5.3	New Ideas	8
6	Acknowledgments	9
6.1	Individual Contributions	9
7	REFERENCES	9

1 Problem

India, despite being the world’s largest producer of several food commodities, continues to grapple with widespread food insecurity, especially among lower-income households. The Global Hunger Index 2023 highlights India’s concerning levels of undernourishment, drawing attention to the gap between food production and equitable food access. Similarly, news coverage around the National Family Health Survey (NFHS-5) findings points to significant disparities in access to sufficient and nutritious food across different socio-economic groups.

While government interventions such as the Public Distribution System (PDS) and food subsidies aim to alleviate hunger, there is growing concern about whether these resources are optimally allocated to those who need it most.

The issue is further complicated by random fluctuations in local food prices due to supply chain disruptions, seasonal variations, and market dynamics, household-level heterogeneity in food expenditure needs, and budgetary constraints faced by policymakers. A static, one-size-fits-all subsidy mechanism may lead to inefficiencies — under-provision for some vulnerable households and over-provision for others.

Motivated by these findings, and drawing upon personal observations of inequities in food access in local communities during periods of price volatility (such as during COVID-19 lockdowns), this project seeks to explore how data-driven optimization techniques can improve food subsidy allocation.

Research Question (RQ):

How can a fixed food subsidy budget be optimally allocated across households to minimize the expected weighted per capita shortfall in food consumption, under random food prices and varying household vulnerability?

2 PLAN

This project formulates the food subsidy allocation challenge as a stochastic optimization problem. The goal is to contribute towards more effective and equitable food security policies through statistically sound optimization methods.

2.1 Model Development

Key components of the model are as follows:

- **Household Stratification:** Households are categorized into 12 fractile groups for both rural and urban sectors. Estimated per capita food expenditure is obtained by applying average food shares to MPCE values. The model uses average values for each fractile group, including MPCE, food share (percentage of MPCE spent on food), and average household size.
- **Stochastic Price Simulation:** To capture price variability, food prices are simulated 1,000 times per group using a normal distribution (mean = 1.0, standard deviation = 0.1), ensuring all prices remain positive.
- **Shortfall Function:** For each group and simulation, a shortfall is calculated as the amount by which effective per capita food consumption falls short of a predefined threshold (1891 for rural and 2078 for urban). The model uses a power function controlled by parameter γ (set to 1) to capture the severity of shortfalls.
- **Weighted Aggregation:** Initial shortfalls without subsidies are used to generate weights for each group, emphasizing higher-need populations in the optimization objective. We have used dynamic weighting, based on initial shortfall.
- **Solver Setup:** The optimization is performed using the Sequential Least Squares Programming (SLSQP) algorithm, ensuring non-negative subsidies and compliance with the total budget constraint.

2.2 Methodology

1. Descriptive analysis of MPCE across income fractiles
2. Stochastic modeling to simulate expected shortfalls under price volatility
3. Optimization using SLSQP under a fixed budget
4. Comparative graphs and visualizations

3 DATA

3.1 Data Source and Study Design

This study uses secondary data from the Household Consumption Expenditure Survey (HCES) 2022–23, conducted by the National Statistical Office (NSO) under the Ministry of Statistics and Programme Implementation (MoSPI), Government of India.

The HCES 2022–23 was based on a probabilistic, stratified multistage sampling design, ensuring national and state-level representativeness for both rural and urban sectors.

multistage stratified sampling design

- **Primary Sampling Units (PSUs):** Villages (rural) and urban blocks (urban) were the first stage units.
- **Ultimate Sampling Units (USUs):** Households within each PSU
- **Selection Method:** Households were selected using Simple Random Sampling Without Replacement (SRSWOR).
- **Stratification** based on land ownership (rural) and car ownership (urban) to ensure economic diversity.

Additionally:

Data were collected through Computer-Assisted Personal Interviewing (CAPI) and each selected household was visited three times across three months, using different questionnaires in each visit to avoid response bias

3.2 Data Collection

Table 1 : Sample Design and Size

Category	Details
Total Households	2,61,746
Rural	1,95,014 households (8,723 villages)
Urban	1,06,732 households (6,115 blocks)
Geographic Coverage	All states and UTs (except remote Andaman & Nicobar villages)

3.3 Data Description

The data provides household-level and per capita-level consumption expenditure information across a wide array of items, categorized under food and non-food heads. For this study focused on food subsidy modeling, variables related to food consumption, household size, household sector (rural/urban), and price data were extracted and used. The primary dependent variable in the optimization model was derived from per capita food expenditure, and prices were estimated as unit values (expenditure divided by quantity). Subsidy weights and shortfall gaps were derived, prices were simulated.

3.4 Data Cleaning and Processing

The dataset used in this model is already cleaned and aggregated. Therefore, there was no need for imputation or handling of missing household data. The focus is entirely on group-level modeling.

4 Analysis

4.1 Optimization Model

Objective Function:

$$\min \mathbb{E}_p \left[\sum_i w_i \cdot \max \left(0, \theta_i - \frac{e_i + x_i}{p_i h_i} \right) \right]$$

Subject to:

$$\sum_i x_i \leq B, \quad x_i \geq 0 \quad \forall i$$

Variable Descriptions:

- x_i : Subsidy allocated to MPCE group i
- p_i : Random price of food (simulated from a normal distribution for group i)
- h_i : Average household size in group i
- e_i : Current per capita food expenditure in household i
- θ_i : Nutritional threshold (1891 for rural, 2078 for urban)
- w_i : Weight of group i , proportional to initial shortfall and population share
- B : Total subsidy budget available

4.2 Optimization Method: SLSQP

We solve the model using the Sequential Least Squares Programming (SLSQP) algorithm. This is a widely used method for constrained nonlinear optimization.

SLSQP is well-suited for our problem because:

- It handles both equality and inequality constraints.
- It works for smooth nonlinear objective functions.
- It can handle bound constraints on variables ($x_i \geq 0$).

We implemented the solution using Python's `scipy.optimize.minimize()` function with the `method='SLSQP'` op

4.3 Plotting And Inferences

Subsidy Allocation Across Fractile Groups:

Shows the percentage of the budget allocated to each group. Lower MPCE groups receive more due to weighted allocation after seeing shortfall.

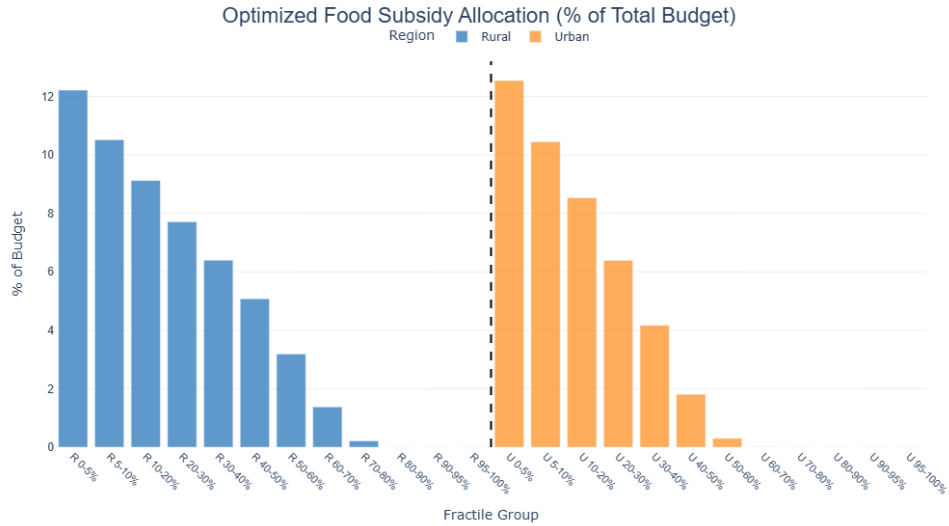


Figure 1:

Shortfall Before vs. After Subsidy:

Compares the average shortfall in food expenditure (as a % of the threshold) before and after the subsidy. The plots confirm that targeted subsidies significantly improve food security for the most vulnerable shows how optimal subsidy allocation can help reduce shortfall.

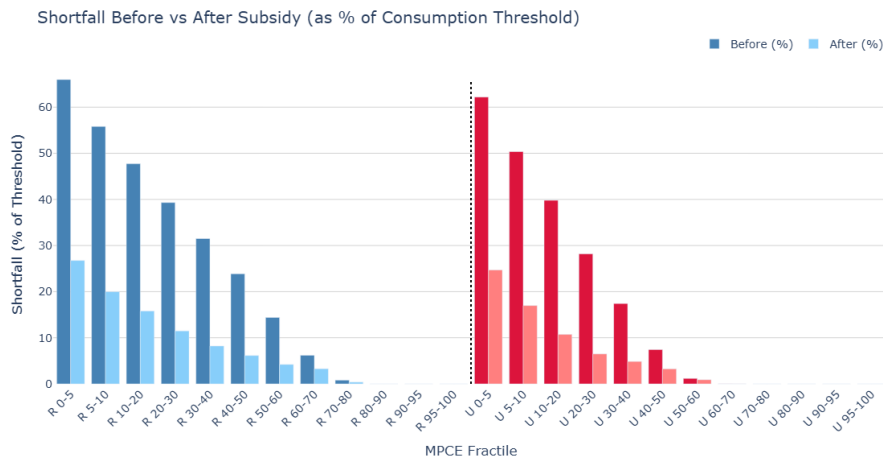


Figure 2:

5 Conclusion

5.1 Key Insights:

- **Targeting:** The majority of the budget is directed at the bottom 30–40% MPCE groups in both rural and urban areas — those with the greatest food shortfalls. Hence, the model acts progressively — higher relative allocations are made to poorer groups, reflecting a need-based policy design.
- **Limitations:**
 1. Some shortfall remains in middle-income urban groups due to budget limits.
 2. We have assumed the budget is 1.5 times initial shortfall while allocating optimal subsidies but have taken this and account while plotting shortfall before v/s after by taking it 0.7 times and hence making practical and realistic sense.

5.2 Implications for Policy

The framework can be adapted for state-level or district-level targeting, depending on available data. By using actual consumption and price variation data, this approach is grounded in empirical realism, making it well-suited for government planning bodies like the Ministry of Consumer Affairs or NITI Aayog.

5.3 New Ideas

- Incorporating individual-level data or additional vulnerability indicators (e.g., child malnutrition, gender) could refine targeting.
- The model can be extended to incorporate temporal variations by accounting for inflation and evolving household needs over time.

6 Acknowledgments

We gratefully acknowledge the invaluable guidance and support of Prof. Kaushik Jana, whose teaching and critical insights throughout the course significantly shaped the development of this project. His emphasis on bridging theoretical frameworks with applied policy challenges provided the foundation for our approach to modeling food subsidy allocation.

We also extend our appreciation to the teaching assistant Mr. Subhajit Pramanick for their consistent support, constructive feedback, and clarifications, which were instrumental during both the formulation and implementation stages of this work.

We would like to thank the National Statistical Office (NSO), Ministry of Statistics and Programme Implementation (MoSPI), Government of India, for making the Household Consumption Expenditure Survey (HCES) 2022–23 dataset publicly available. This high-quality dataset enabled a meaningful exploration of food subsidy targeting in India and served as the empirical backbone of our analysis.

6.1 Individual Contributions

Ridhwan Choudhari: Plotting models, making inferences, editing the Latex report.

Rajneesh Yadav: Writing and simulating code, editing Latex report.

7 REFERENCES

References

- [1] [Household Consumption Expenditure Survey \(HCES\) 2022–23.](#)
- [2] [Factsheet of Household Consumption Expenditure Survey \(HCES\) 2022-23.](#)
- [3] [NSS 68th round data.](#)
- [4] [Boyd, S., & Vandenberghe, L. \(2004\). *Convex Optimization*. Cambridge University Press.](#)