A Convex Optimization Approach to Constrained LPG Allocation

Utkarsh Bharadwaj Neerav Bhuyan Saptarshi Datta
Indian Statistical Institute

April 2026

Abstract

This project formulates a data-driven convex optimization model to manage national liquefied petroleum gas (LPG) rationing during severe supply chain disruptions. Utilizing state-level historical consumption data (2023 to 2025), we optimally allocate a constrained fuel supply across domestic and commercial sectors to minimize the combined societal penalties of commercial economic loss and household biomass pollution. The model introduces proportional normalization to eliminate population-size bias, employs chance-constrained programming to guarantee a stochastic minimum survival baseline, and enforces real-world policy constraints (equity floors and survival minimums) to ensure politically viable allocation.

1 Problem

1.1 Research Problem (RP)

The 2026 geopolitical conflict in West Asia has severely disrupted India's LPG imports, resulting in a national fuel shortage. To manage the crisis, the government must ration available LPG supplies. When commercial entities (e.g., restaurants) face shortages, they suffer measurable economic losses. When domestic households face shortages, they are forced to substitute LPG with biomass or firewood, which significantly increases indoor air pollution and associated health burdens. The challenge is to optimally allocate a strictly limited national LPG supply across domestic and commercial sectors across various states to minimize the combined socioeconomic and public health damages.

In addition to the immediate supply shortage, the allocation problem is complicated by the heterogeneity across states in terms of population density, economic structure, and access to alternative cooking fuels. States with higher dependence on LPG for household consumption face disproportionately larger health risks when supply is curtailed, while states with significant commercial activity face measurable economic losses.

2 Data Collection and Parameter Synthesis

A significant component of this study involved transforming qualitative societal impacts into empirically rigorous parameters for the optimization model. To achieve geographical coverage, data was aggregated for nine highly-populated states alongside a tenth synthesized "Others" category representing the national remainder.

2.1 Baseline Demands (D_i^d and D_i^c) and Supply Constraint (S)

Baseline domestic and commercial demand requirements (D_i^d and D_i^c) were calculated by averaging month-wise empirical consumption data from January 2023 to December 2025. To simulate the 2026 geopolitical shock, the total supply parameter (S) was constrained to 75% of normal operational availability. Based on Petroleum Planning & Analysis Cell (PPAC) data, the baseline monthly physical supply average was established at 2727.33 Thousand Metric Tonnes (TMT), yielding a crisis supply constraint of:

$$S = 2045.5 \text{ TMT}$$

2.2 Vulnerability Weights (w_i)

To prioritize states where an LPG shortage directly forces populations into severe respiratory distress, demographic vulnerability weights (w_i) were synthesized using the 2019 to 2021 National Family Health Survey (NFHS-5). Using the state-level indicator for the percentage of households utilizing clean cooking fuel (C_i), vulnerability was initially quantified as the proportional reliance on solid biomass.

However, to prevent the algorithm from treating highly developed states as completely expendable, a mathematical vulnerability floor of 0.20 was introduced:

$$w_i = \max\left(0.20, \frac{100 - C_i}{100}\right)$$

Critique of Clean Fuel Definition (Resilience vs. Reversion): It must be acknowledged that the NFHS-5 definition of "clean fuel" (C_i) encompasses LPG, Piped Natural Gas (PNG), and electricity. During a strict LPG supply shock, absolute clean fuel access in a highly developed state like Delhi would theoretically plummet. However, we maintain that our formulation of w_i remains robust. In this model, w_i serves as a proxy for a state's structural propensity to *revert to solid biomass.* Highly developed states possess the infrastructural and economic resilience (e.g., uninterrupted electricity for induction cooking, active PNG networks) to substitute missing LPG without generating indoor air pollution. Conversely, states with a high baseline reliance on biomass lack these safety nets, making them mathematically valid targets for optimization priority.

Table 1: Derivation of State Vulnerability Weights (w_i) from NFHS-5 Data

State	**Clean Fuel Access** (C_i)	**Raw Weight** $\left(\frac{100-C_i}{100}\right)$	**Final Adjusted** w_i
Uttar Pradesh	32.7%	0.673	0.673
Madhya Pradesh	40.1%	0.599	0.599
West Bengal	40.6%	0.594	0.594
Rajasthan	41.5%	0.585	0.585
Others (National)	58.6%	0.414	0.414
Karnataka	68.2%	0.318	0.318
Gujarat	68.3%	0.317	0.317
Maharashtra	71.9%	0.281	0.281
Tamil Nadu	84.8%	0.152	**0.200***
Delhi	97.4%	0.026	**0.200***

**Adjusted upwards to meet the 0.20 mathematical vulnerability floor.*

2.3 Penalty Weights (c_h and c_e)

Translating disparate socioeconomic damages into a unified mathematical objective space requires establishing a comparable economic denomination. For this model, both health and economic damages were standardized to **Millions of INR per TMT of shortage**.

2.3.1 Health Penalty Weight (c_h)

The health penalty quantifies the public health damage incurred when a domestic household is forced to substitute clean LPG with solid biomass or firewood due to supply constraints. This was calculated by translating epidemiological metrics into financial costs.

- **Establishing the Health Burden (DALYs):** The Global Burden of Disease (GBD) study explicitly tracks "Household Air Pollution from Solid Fuels" as a primary mortality and morbidity risk factor. Epidemiological data indicates that a household fully transitioning to biomass loses approximately 0.05 Disability-Adjusted Life Years (DALYs) annually. For our monthly optimization model, this translates to an average burden of **0.004 DALYs per month** per household.

- **Monetization via VSL:** To convert this epidemiological burden into a financial penalty, we utilized the Value of a Statistical Life (VSL) for India. VSL measures a society's willingness to pay to reduce the risk of mortality. Using conservative economic estimates, the statistical value of a single DALY in India equates to approximately **₹**13.3 Lakhs (1.33 Million INR).

- **Final Derivation:** By multiplying the monthly DALY loss by its economic value, the societal damage of denying one household LPG for a month is roughly **₹**5,320. Scaling this figure up to match our physical supply units (Thousand Metric Tonnes), the final health penalty coefficient was derived as:

$$c_h = 532 \text{ Millions INR per TMT} \tag{1}$$

2.3.2 Economic Penalty Weight (c_e)

The economic penalty quantifies the direct financial revenue lost when the commercial sector faces an LPG shortage, effectively operating as an exercise in macroeconomic attribution.

- **Sector Isolation and GVA:** Heavy industries do not generally rely on 19kg commercial LPG cylinders; the primary driver of this demand is the "Accommodation and Food Services" sector (e.g., hotels, restaurants, and street food vendors). According to the Ministry of Statistics and Programme Implementation (MoSPI), the annual Gross Value Added (GVA) for this specific sector is approximately 2,000,000 Million INR. Dividing this yields a monthly GVA of **166,666 Million INR**.
- **Total Baseline Demand:** Aggregating the historical commercial demand (D_i^c) across all analyzed geographical entities gives a total national commercial LPG requirement of **275.73 TMT per month**.
- **Energy Inelasticity Assumption:** To calculate the penalty, we assumed strict energy inelasticity. This posits that commercial cooking fuel acts as an absolute bottleneck; if a restaurant lacks LPG, its core revenue-generating operations halt entirely.
- **Final Derivation:** The economic value generated by a single TMT of fuel was calculated by dividing the total monthly economic output of the sector by the total monthly commercial fuel demand:

$$c_e = \frac{\text{Total Monthly Sector GVA}}{\text{Total Monthly Comm. Demand}} = \frac{166,666}{275.73} \approx 604 \text{ Millions INR per TMT} \tag{2}$$

3 Mathematical Formulation

The allocation problem is modeled as a strictly convex Quadratic Programming (QP) problem. The formulation is strictly convex under the condition that all vulnerability weights $w_i > 0$ and baseline demands $D_i^d, D_i^c > 0$.

3.1 Stochastic Minimum Survival Baseline (θ_i)

Emergency fuel demand during severe crises often exhibits heavily right-skewed, "fat-tailed" behavior due to unpredictable logistical compounding and localized hoarding. Therefore, modeling survival demand with a symmetric Normal distribution theoretically underestimates the probability of extreme shortage events.

To correct for this, the minimum required survival allocation ($\tilde{\theta}_i$) was mathematically modeled as a strictly positive **Log-Normal** distributed random variable: $\tilde{\theta}_i \sim$

Lognormal($\mu_{\log}, \sigma^2_{\log}$). The expected baseline (μ_i) assumes a requirement of 10 kg per vulnerable household, with a Coefficient of Variation (CV) of 0.15. The parameters for the underlying normal distribution were derived as:

$$\sigma^2_{\log} = \ln(1 + \text{CV}^2), \quad \mu_{\log} = \ln(\mu_i) - \frac{1}{2}\sigma^2_{\log}$$

We utilized chance-constrained programming to enforce a strict 95% probabilistic guarantee of survival ($\alpha = 0.05$). The threshold was extracted using the Log-Normal quantile function (F^{-1}):

$$\theta_i = F^{-1}(1 - \alpha; \mu_{\log}, \sigma_{\log})$$

Table 2: Derivation of Log-Normal Minimum Survival Baseline (θ_i) by State

State	**Vulnerable Households**	**Expected Min. (μ_i in TMT)**	**Log-Normal Threshold (θ_i in TMT)**
Others (National)	15,000,000	150.00	189.59
Gujarat	2,200,000	22.00	27.81
Tamil Nadu	1,950,000	19.50	24.65
Madhya Pradesh	1,800,000	18.00	22.75
West Bengal	1,700,000	17.00	21.49
Rajasthan	1,600,000	16.00	20.22
Maharashtra	1,450,000	14.50	18.33
Karnataka	800,000	8.00	10.11
Uttar Pradesh	750,000	7.50	9.48
Delhi	150,000	1.50	1.90

Calculated using μ_i = 10 kg/household and a 95% probabilistic guarantee on the Log-Normal curve.

3.2 Objective Function (Proportional Normalization)

To eliminate the "Big State Bias" (where the model disproportionately wiped out supplies to smaller states to shave minor absolute volumes off massive populations), the objective function was refined to incorporate proportional normalization, penalizing the percentage shortfall rather than the raw tonnage:

$$\min_x f(x) = \sum_{i=1}^{n} w_i \left[c_h D_i^d \left(\frac{D_i^d - x_i^d}{D_i^d} \right)^2 + c_e D_i^c \left(\frac{D_i^c - x_i^c}{D_i^c} \right)^2 \right] \tag{3}$$

For proof of strict convexity of the objective function, refer to Appendix A.

3.3 Constraints

To ensure the mathematical optimum is politically and practically viable, the optimization is subjected to the following physical, survival, and real-world equity constraints:

$$\sum_i (x_i^d + x_i^c) = S_{\text{usable}} \qquad \text{(Total Usable Supply)} \tag{4}$$

$$S_{\text{usable}} = S - 0.05S \qquad \text{(5\% National Reserve Policy)} \tag{5}$$

$$x_i^d \geq 0.40 \cdot D_i^d \qquad \text{(Domestic Baseline Equity Guarantee)} \tag{6}$$

$$x_i^c \geq 0.25 \cdot D_i^c \qquad \text{(Commercial Survival Floor)} \tag{7}$$

$$x_i^c \leq 0.70 \cdot D_i^c \qquad \text{(Government Commercial Policy Cap)} \tag{8}$$

$$x_i^d \leq D_i^d \qquad \text{(No Domestic Over-allocation)} \tag{9}$$

$$x_i^d \geq \theta_i \qquad \text{(Chance-Constrained Survival Threshold)} \tag{10}$$

For the matrix formulation in R, refer to Appendix B.1.

4 Results and Analysis

The quadratic programming formulation was executed using the `quadprog` model in R. The 75% national supply constraint, compounded by the 5% reserve requirement ($S_{\text{usable}} = 1943.225$ TMT), forced the algorithm into intense resource competition.

Table 3: Optimal LPG Allocation with Real-World Constraints (March 2026)

State	Domestic Alloc. (TMT)	% of Normal	Comm. Alloc. (TMT)	% of Normal
Delhi	19.10	40.00%	4.62	41.89%
Tamil Nadu	79.57	40.00%	9.24	41.89%
Maharashtra	148.99	53.03%	24.25	58.63%
Gujarat	63.47	58.36%	12.22	63.31%
Karnataka	93.01	58.50%	12.25	63.47%
Others (National)	642.97	68.12%	59.83	70.00%
Rajasthan	114.98	77.44%	9.65	70.00%
West Bengal	175.27	77.78%	11.58	70.00%
Madhya Pradesh	113.70	77.97%	9.65	70.00%
Uttar Pradesh	315.73	80.39%	23.16	70.00%

For the R code of solver execution, refer to Appendix B.2.

4.1 Analytical Findings

The optimization results demonstrate the successful fusion of pure mathematical efficiency with real-world political viability. The inclusion of equity constraints drastically reshaped the allocation landscape, yielding several key analytical insights:

- **Prioritization of High-Risk Demographics:** The algorithm successfully targeted states with the highest risk of solid biomass reversion. Highly vulnerable states (e.g., Uttar Pradesh, Madhya Pradesh, West Bengal, and Rajasthan) retained approximately 77% to 80% of their normal domestic demand, minimizing the projected DALYs lost to indoor air pollution.

- **Activation of Protective Equity Floors:** The 40% domestic equity guarantee proved to be a strictly binding constraint ($x_i^d = 0.40D_i^d$) for highly developed states like Delhi and Tamil Nadu. Unconstrained, the purely convex objective function would have nearly zeroed out their domestic supplies; however, the floor successfully prevented a complete systemic collapse, enforcing a baseline sharing of the national shortage.

- **Binding Commercial Policy Caps:** For highly vulnerable states, the model attempted to aggressively defund their commercial sectors to reallocate fuel to struggling households. However, the 70% policy cap acted as a hard mathematical ceiling. The model funded the commercial sectors of UP, MP, WB, Rajasthan, and the "Others" category exactly up to this 70% limit to prevent catastrophic economic scarring, reflecting the heavy weight of the c_e penalty.

- **The "Squeezed Middle" Phenomenon:** Mid-tier states with moderate baseline clean fuel access (e.g., Maharashtra, Gujarat, and Karnataka) did not trigger the protective equity floors, nor did they hit the 70% commercial caps. Consequently, their allocations (settling around 53% to 58% for domestic and 58% to 63% for commercial) represent the pure gradient equilibrium of the objective function, where the marginal economic penalty perfectly balanced the marginal health penalty.

- **Validation of Strategic Reserves:** The algorithm successfully operated within the artificially tightened S_{usable} bound. It successfully distributed the fuel while holding back exactly 102.275 TMT (5% of the total constrained supply) as an unallocated emergency reserve, proving the model can accommodate real-world physical supply chain buffers without mathematical failure.

5 Conclusion and Limitations

This framework demonstrates that convex optimization can effectively navigate complex socioeconomic trade-offs during acute supply disruptions. By integrating epidemiological indicators (DALYs) with macroeconomic metrics (GVAs), the model provides a systematic and transparent approach to resource allocation.

However, the framework operates under several necessary assumptions that limit its real-world precision:

- **Energy Demand Inelasticity:** The calculation of the economic penalty (c_e) assumes perfect energy inelasticity (i.e., cutting LPG completely shuts down a commercial unit's revenue generation). In reality, many restaurants possess the capital to temporarily switch to electric induction or procure alternative fuels, meaning the actual economic damage curve may be softer than modeled.

- **Static Penalty Weights:** The model treats health damages and economic bankruptcies as linear, static coefficients for a single month. In reality, these damages compound non-linearly over time; a restaurant might survive a 1-month shortage on reserves but go permanently bankrupt in month two.

- **Assumption of Uniform Distribution:** The state-level parameters treat entire states as monolithic entities. The model assumes the state government will perfectly and equitably distribute its allocated TMT to its most vulnerable residents, ignoring severe intra-state logistical inequalities (e.g., rural vs. urban distribution pipelines).

References

[1] International Institute for Population Sciences (IIPS) and ICF. (2021). *National Family Health Survey (NFHS-5), India, 2019 to 2021: Main Report* (Table 2.6 and 2.7). Mumbai: IIPS. Available at: https://dhsprogram.com/pubs/pdf/FR375/FR375.pdf

[2] Ministry of Statistics and Programme Implementation (MoSPI), Government of India. (2024). *National Accounts Statistics: Gross Value Added (GVA) at Basic Prices by Economic Activity (Accommodation and Food Services)*. New Delhi: MoSPI.

[3] Institute for Health Metrics and Evaluation (IHME). (2020). *Global Burden of Disease Study 2019 (GBD 2019) Risk Factor Results: Household Air Pollution from Solid Fuels*. Seattle, WA: IHME, University of Washington.

[4] Petroleum Planning & Analysis Cell (PPAC), Ministry of Petroleum and Natural Gas, Government of India. (2025). *Industry Performance Review: LPG Production, Imports, and Consumption Historical Data*. New Delhi. Available at: https://ppac.gov.in/

[5] Viscusi, W. K., & Masterman, C. J. (2017). *Income Elasticities and Global Values of a Statistical Life*. Journal of Benefit-Cost Analysis, 8(2), 226 to 250. (Utilized for baseline VSL scaling methodology for India).

[6] Turlach, B. A., & Weingessel, A. (2019). *quadprog: Functions to solve Quadratic Programming Problems*. R package version 1.5-8. Comprehensive R Archive Network (CRAN).

Appendix A: Proof of Strict Convexity of Objective Function

To mathematically guarantee that the optimization model converges to a global minimum rather than trapping itself in a local minimum, we must prove that the objective function $f(x)$ is strictly convex. A twice-differentiable multivariable function is strictly convex if and only if its Hessian matrix ($\nabla^2 f$) is strictly positive definite.
Let us simplify the objective function by isolating the constants for a single geographical entity i. By expanding the denominators, we define the positive coefficients A_i and B_i:

$$A_i = \frac{w_i c_h}{D_i^d}, \quad B_i = \frac{w_i c_e}{D_i^c}$$

The objective function can thus be rewritten as a sum of independent quadratic terms:

$$f(x) = \sum_{i=1}^{n} \left[A_i (D_i^d - x_i^d)^2 + B_i (D_i^c - x_i^c)^2 \right]$$

We compute the second-order partial derivatives. For the domestic variables x_i^d:

$$\frac{\partial f}{\partial x_i^d} = -2A_i(D_i^d - x_i^d) \quad \Longrightarrow \quad \frac{\partial^2 f}{\partial (x_i^d)^2} = 2A_i$$

For the commercial variables x_i^c:

$$\frac{\partial f}{\partial x_i^c} = -2B_i(D_i^c - x_i^c) \quad \Longrightarrow \quad \frac{\partial^2 f}{\partial (x_i^c)^2} = 2B_i$$

Because the decision variables are completely independent within the objective function (i.e., there are no interaction terms such as $x_i^d \cdot x_i^c$), all mixed partial derivatives are exactly zero ($\frac{\partial^2 f}{\partial x_i^d \partial x_i^c} = 0$). Consequently, the Hessian matrix H is a $2n \times 2n$ diagonal matrix:

$$H = \text{diag}(2A_1, \dots, 2A_n, 2B_1, \dots, 2B_n)$$

For a diagonal matrix to be strictly positive definite, all values on its main diagonal must be strictly greater than zero. We observe our foundational parameters:

- $w_i \geq 0.20 > 0$ (due to the implemented mathematical vulnerability floor).
- $c_h = 532 > 0$ and $c_e = 604 > 0$.
- $D_i^d > 0$ and $D_i^c > 0$ (baseline demands are intrinsically strictly positive).

Since all constituent parameters are strictly positive, the coefficients $2A_i > 0$ and $2B_i > 0$. Therefore, H is strictly positive definite, proving that the objective function $f(x)$ is strictly convex.

Appendix B: R Code Formulation

```
# Dmat and dvec (Proportional Normalization applied)
diag_elements = c((2*w_i * c_h) / D_id, (2*w_i * c_e) / D_ic)
Dmat <- diag(diag_elements)
Dmat <- Dmat + diag(1e-8, 2 * n)

dvec <- c( rep(2 * c_h, n) * w_i, rep(2 * c_e, n) * w_i )

# Amat and bvec: The Constraints
# Constraint 1: Sum of all allocations == S_usable
A_sum <- rep(1, 2 * n)
b_sum <- S_usable

# Constraint 2: x_id >= theta_i (Stochastic Survival)
A_theta <- rbind(diag(n), matrix(0, nrow = n, ncol = n))
b_theta <- theta_i

# Constraint 3: x_id <= D_id => -x_id >= -D_id
# (No Domestic Over-allocation)
A_did <- rbind(-diag(n), matrix(0, nrow = n, ncol = n))
b_did <- -D_id

# Constraint 4: x_ic >= 0.25 * D_ic
# (25% Commercial Survival Floor)
A_xic_floor <- rbind(matrix(0, nrow = n, ncol = n), diag(n))
b_xic_floor <- 0.25 * D_ic

# Constraint 5: x_id >= 0.40 * D_id
# (40% Domestic Equity Guarantee)
A_xid_floor <- rbind(diag(n), matrix(0, nrow = n, ncol = n))
b_xid_floor <- 0.40 * D_id

# Constraint 6: x_ic <= 0.70 * D_ic => -x_ic >= -0.70 * D_ic
# (Commercial Cap)
A_dic <- rbind(matrix(0, nrow = n, ncol = n), -diag(n))
b_dic <- -0.70 * D_ic

# Combine all constraints
Amat = cbind(A_sum,A_theta,A_did,A_xic_floor,A_xid_floor,A_dic)
bvec = c(b_sum,b_theta,b_did,b_xic_floor,b_xid_floor,b_dic)
```

Appendix B:.1 Solver Execution

```
result <- solve.QP(Dmat, dvec, Amat, bvec, meq = 1)

# Extract optimal allocations
optimal_x <- result$solution
baseline_demand <- baseline_demand %>%
  mutate(
    Optimal_x_id = round(optimal_x[1:n], 2),
    Optimal_x_ic = round(optimal_x[(n+1):(2*n)], 2),
    Domestic_Pct_Met = round((Optimal_x_id / D_id) * 100, 2),
    Commercial_Pct_Met = round((Optimal_x_ic / D_ic) * 100, 2)
  )
```