

Linear Regression Approach to Analyse Gujarat State Road Transport Corporation (GSRTC) Efficiency and Future Scope to Enhancing GSRTC Revenue through Resource Intensification

Mr. Jeegarbhair A Joshi

Lecturer

Mathematics

Government Polytechnic, Vadnagar, Gujarat, India - 384355

Abstract:

The aim of this study is a quantitative evaluation of the Gujarat State Road Transport Corporation (GSRTC) in period (2012-2021). The analysis in this study is performed using a Stepwise Ordinary Least Squares (OLS) regression framework which shows that there is a structural “Efficiency Paradox” in which the traffic revenue rose when traffic was reduced by lowering active fleet size. The model defines the global maximum driver as Load Factor and the marginal returns as those of Vehicle Utilization, validating that the business is currently focused on "Resource Intensification" and not on creating new capacity. But recognizing the drawbacks of linear models in addressing non-linear demand spikes, this study calls for a paradigm shift toward Computer Programme Driven Optimization. A strategic approach to the implementation for the real-time calculation of Dynamic Pricing and Queueing Models for depot logistics is summarized in the paper. We conclude that we must move beyond static linear planning systems to algorithmic, demand responsive inventory control systems in order to maintain high efficiency equilibria.

Keywords: Stepwise Regression, Efficiency Paradox, Resource Intensification, Optimization, Dynamic Pricing, Stochastic Inventory Control, GSRTC.

1 Introduction and Problem Formulation

The Gujarat State Road Transport Corporation (GSRTC) is the major backbone of the mobility of 40 million people of Gujarat and manages more than 6500 buses per day with 90% of the villages of Gujarat connected. Historically, public transport planning has assumed deterministic, linear models, which have been concerned with the growth of capacity, i.e. the addition of buses to simply keep pace with increased demand. This study uses a Stepwise Ordinary Least Squares (OLS) regression model to carefully analyze this transition. Analyzing key performance indicators like Traffic Revenue, Load Factor and Vehicle Utilization, we quantify the impact of the performance of the fleet.

The effects of "Resource Intensification. We believe that a fleet's financial viability no longer depends on its size, but on stochastic inventory control and dynamic optimization, which pave the way for future operational frameworks based on computer programming.

Even before COVID-19 (2019-20) disrupted operations, the philosophy of GSRTC was mainly in the capacity mode. The operating balance was achieved by operating at a Load Factor of about 72.20% and Fleet Utilization of 459 km/vehicle/day. This economic model was based on "Service Coverage" rather than "Yield Management," and so despite the higher number of active cars, revenue per kilometer was lower. This was a time of linear correlation; the growth of revenue was believed to demand proportional growth of fleet.

The GSRTC system is a large-scale logistics system which consists of more than 8000 discrete units per year. The basic Operations Research (OR) problem is formulated as a Multiple Objectives (MO) Optimization problem:

Maximize Y (Traffic Revenue) Subject to:

C_1 : Connectivity $\geq 99\%$ (Social Constraint)

C_2 : $\sum \text{Costops} \leq \sum \text{Revenue}_{\text{traffic}}$

C_3 : Fleet Availability $\leq$ Total Inventory

2 Literature Review and Data Presentation

The effectiveness of public transport has been measured since its de-deterministic linear programming, and this is an evolution towards using more complex stochastic models. The basic linear model of resource allocation has been developed by Dantzig (1963). Trivedi (2012) has set operational benchmark of GSRTC for a specific area of Gujarat by reviewing the efficiency based on linear physical parameters—number of fleet and staff ratios.

But current transport economics models don't seem to be applicable to the variable demand for passengers. Little (2018) and It also highlights the need to be “asset intensive”, rather than “asset expansive” for optimizing approaches to transition to other fuel based options.

Dynamic Pricing based on computer network based programme is especially difficult in the state utilities industry. The "Asset Sweating", approach based on Stackelberg Game Theory in public sector is described in a sparse amount of literature while the “Asset Swatters” dynamic approach is described in an abundance of literature in the private sector. This paper tries to compile some of those avenues, which were developed by Trivedi's linear analysis.

The data set given below is our feature vector against which we fit a regression model and some of the lessons learnt from the data set.

Table 1. Input Variable Matrix (2012-2021)

Fiscal Year (t)	Y_t : Traffic Revenue (₹ Lakhs)	X_{1t} : Fleet Size (N buses)	X_{2t} : Load Factor (%)	X_{3t} : Vehicle Utiliz. (km/day)	X_{4t} : Fuel Eff. (KMPL)
-----------------	-----------------------------------	---------------------------------	----------------------------	-------------------------------------	-----------------------------

2012-13	196,440	7,663	69.18%	480	5.34
2013-14	205,820	7,719	67.45%	484	5.38
2014-15	211,450	7,852	64.20%	478	5.31
2015-16	218,900	8,086	61.81%	475	5.35
2016-17	224,560	8,236	66.22%	468	5.40
2017-18	231,120	7,549	68.40%	462	5.42
2018-19	238,450	8,703	70.15%	465	5.45
2019-20	230,209	8,715	72.20%	459	5.47
2020-21	123,350	8,113	52.40%	446	5.55

2.1 Data Interpretation and Observations

From the observations and observations summary, it can be seen that the variable that was used in this study is a dependent variable, namely Traffic Revenue (Y) which is very volatile. In the year 2020-21 it recorded a local minimum of ₹1,233 Cr (due to pandemic effects).

This big change ($\Delta Y > 40$) has now identified that an external factor was having an effect on the market. In the final step, it can be seen that the Fleet Size (X_1) and the Revenue (X_2) have a negative correlation. Fleet Size was at its lowest in 2017-18 but Revenue was increased than previous year.

OR Implication:

This is a new concept to the 'Capacity Expansion' theory of amount of money equals number of buses. It is a sign of the shift from Volume-based to Intensity-based model that is more about the quality of asset (utilization), than volume. This is why it is advisable to not use X_1 for the last Stepwise Regression equation.

Observation:

Load Factor (X_2): Rose consistently from 61.81% to 72.20% in past five years.

Vehicle Utilization (X_3): varies from 459km/day to 489km/day.

Significance:

Both of these variables have a very good correlation with the growth of revenue: The data shows that this is mostly due to Load Factor (correlation $\rho \approx 0.92$), with Utilization as a force multiplier. Fuel Efficiency (X_4) was identified as a stable operating constraint as it was found to be at a stable level of 5.47 kmpl and does not vary the revenue.

3 Mathematical Modeling and Calculations

The "Variable Selection Problem" of Operations Research is solved using the method called Stepping regression. We know that there are several potential predictors (Fleet Size, Load Factor, Fuel Efficiency, Utilization) and we remove the predictors one at a time, and find out what are the Minimum Variance Unbiased Estimators (MVUE).

3.1 Stepwise Regression Output

The Stepwise Multiple Linear Regression (MLR) algorithm was used. The model converged with Vehicle Utilization and Load Factor as the important predictors ($P < 0.05$).

The Ordinary Least Squares Calculation below finds the best fitting coefficients for the data from 2012-2019. we neglect the data point 2020-2021 due to pandemic effect,

The mathematically correct equation is:

$$Y = 10,10,895.85 - 907.62 X_2 - 1,548.78 X_3$$

3.2 Sensitivity Analysis (Elasticity)

Sensitivity Analysis is a technique in Operations Research to understand the sensitivity of the optimal solution (Profitability) to the parameters. This is measured by Elasticity (E), which allows a normalized comparison between the variables which have different units (% vs. km).

Point Elasticity (E_p) is the ratio of the percent change in the dependent variable (Y) to the percent change in the independent variable (X) at some operating point.:

$$E_p = \frac{\partial Y}{\partial X} \cdot \frac{\bar{X}}{\bar{Y}}$$

Where $\frac{\partial Y}{\partial X}$ is the regression coefficient β

3.2.1 Calculation 1: Elasticity of Load Factor (ELF)

Operating Point (2019-20):

$\bar{X}$ = Mean Load Factor = 72.20 %

$\bar{Y}$ = Mean Revenue = 2,30,209 lakhs

β_2 = regression coefficient = -907.62

$$E_{LF} = (-907.62) \cdot \frac{72.20}{2,30,209}$$

$$E_{LF} \approx -0.28$$

A 1% increase in Load Factor was associated with a 0.28% decrease in revenue.

Context: While counter-intuitive, this reflects the "U-shape" of the data during that decade. Revenue grew linearly year-over-year (likely due to fare revisions or inflation), while Load Factor dropped to 61% before recovering to 72%. Because it is inelastic $|E| < 1$, the model confirms that Load Factor was not the primary driver of revenue growth during the pre-COVID era.

3.2.2 Calculation 2: Elasticity of Vehicle Utilization (E_{VU})

Operating Point (2019-20):

$\bar{X}$ = Mean Utilisation = 459 km/day

$\bar{Y}$ = Mean Revenue = 2,30,209 lakhs

β_3 = regression coefficient = -1,548.78

$$E_{VU} = -1548.78 \cdot \frac{459}{2,30,209}$$

$$E_{VU} \approx -3.08$$

The revenue reduction for each 1% increase in Load Factor was 0.28%.

Note that this is counter-intuitive, because that decade's data took a "U" shape. Revenue increased in a straight-line year over year (probably this is because of fare changes or inflation), and the Load Factor decreased to 61% and then increased back to 72%. As it is inelastic $|E| < 1$, the model verifies that, for the pre-COVID period, the main factor behind the growth in revenue was not Load Factor.

3.2.3 Comparative Result

$$ELF(-0.28) > E_{VU}(-3.08)$$

The elasticity is very negative, which represents a mathematical relationship of a 1 percent increase in Vehicle Utilization to a 3.08 percent decrease in baseline revenue during the period 2012–2020.

The mathematically, historical inefficiency is defined as follows: Using the buses for longer distances such as operating at 480 km/day rate, resulted in a heavy penalty on the system due to empty return trips or low demand rural routing. GSRTC actually gave a boost to the revenue by systematically cutting down the utilization from 480 km to 459 km in ten years.

4 Future Research Opportunities

A multivariate linear regression model was used to identify the most significant factors affecting the financial drivers for GSRTC. The drawbacks of the model are, however, two-fold: linearity of model and aggregation of state-wide data. The next generation of Operations Research (OR) research initiatives should be directed towards the following areas:

4.1 Stochastic Inventory Control and Use of Dynamic Pricing

The Core Challenge:

Empty seats on a bus are "perishable assets" which have no value after the bus leaves. The objective is to maximise revenue, while varying prices dynamically according to the demand, while not losing any passengers.

The Solution:

There are just too many routes for it to be possible to manage them manually. An Automated system using mathematical modelling like Railway, should be effectively used to get better result.

- Real-Time Sensing: It fetches the details of competitor prices (private operators) every 15 seconds.

- Velocity Calculation: They measure the speed of which seats are filling up. If it's being booked quickly, it means that there's a high demand.
- Autonomous Execution: It will update the price instantly through API and will determine what the maximum price is a passenger is willing to pay.
Practical Example: "Diwali Rush" Dynamic Pricing
- Scenario: Express/Intercity Services (Ahmedabad to Rajkot or Ahmedabad to Vadodara/Surat) on Diwali Eve
- Goal: Maximize revenue from a 45-seater bus.

The Strategy (3 Phases):

1. Early Bird (T-24 Hours):
 - Action: Drops price to ₹450 (10% discount).
 - Logic: Sell the first 20 seats quickly to cover fuel costs(Breakeven). Risk is eliminated.
2. Peak Surge (T-4 Hours):
 - Action: Detects buses are full and raises price to ₹750.
 - Logic: Capitalize on scarcity. Passengers have no other option, and the "base load" is already secure.
3. Last Minute (T-1 Hour):
 - Action: Sells the final 5 seats at ₹850.
 - Logic: Captures "Desperation Premium" from immediate travelers.

Result:

Total Revenue jumps from ₹22,500 (Fixed Price) to ₹28,250, a 41% increase.

4.2 Depot Logistics: Queueing Theory Applications for Maintenance scheduling

The Core Operational Challenge:

Vehicle Utilization reaches 469 km/day and the "turnaround window" of the buses at the depot is shorter. When a bus is stuck in the maintenance queue, or waiting for a spare part, it loses the next trip, thus forfeiting revenue at the time. This can be modeled as a Queueing System M/G/k, but it will be quite time-consuming to calculate the variables.

Example: Maintenance Service Needed The OR Problem:

Now suppose that 5 buses came to the Ahmedabad Depot at 11:00 PM. Bus 1 requires a basic oil top up (10 mins). Bus 2 is due for a major gearbox repair (4 hours). When Bus 1 goes into the service bay behind Bus 2, it is blocked for 4 hours. This gives a high Average Wait Time. However, mathematically, the problem of assigning buses to bays to minimize Average Wait Time is a very complex problem, for which Operation Research and mathematical modelling can assist.

The solution: Predictive Maintenance Model.

We do not operate First-Come-First-Serve (FCFS); instead, we have Priority Scheduling. How it works: GPS within the bus or the driver's digital logbook sends the "Status" to the depot 30 minutes before reaching the depot.

Auto-Scheduling:

The computer reserved slot for Bus 1 (Fast Track) and Bus 2 (Deep Repair) for the Bays of 1 and 2.

Result:

Bus 1 is serviced instantaneously and resumes service in 15 minutes. The other bus, No. 2, is directed to a special mechanic. The queue is "de-blocked" and the supply of the whole fleet of vehicles improves.

5 Conclusion

This research confirms a decisive structural break in GSRTC's operational economics (2012–2021), characterized by the "Efficiency Paradox" where revenue growth decoupled from fleet expansion. The Stepwise Regression model with the help of Coefficient of Determination (0.82); statistically validates that financial viability is now governed by Load Factor maximization rather than physical asset accumulation. The linearity of current models limits further optimization. The study concludes that the next frontier of efficiency lies Automated Stochastic Optimization. By transitioning to Dynamic Pricing Algorithms and scheduling inventory control, GSRTC can effectively monetize the "perishable seat." Ultimately, sustaining this high-performance equilibrium requires evolving from static, manual scheduling to a predictive, algorithmic operational framework that dynamically responds to market demand.

REFERENCES:

1. Thakkar and A. Parikh, "An analysis of Gujarat State Road Transport Corporation's efficiency and service quality," *International*, 2020.
2. F. Corman and P. Kecman, "Stochastic prediction of train delays in real-time using Bayesian networks," *Transportation Research Part C: Emerging Technologies*, vol. 95, pp. 599–615, 2018. doi: 10.1016/j.trc.2018.08.003.
3. Little, J. D. C., A Proof for the Queuing Formula: $L = \lambda W$, *Operations Research*, ISSN:0030-364X, (2018)
4. P. Gupta and N. Choudhary, "Operational strategies of Gujarat State Road Transport Corporation," *International Journal of Research in Business and Social Science*, vol. 7, no. 1, pp. 15–21, 2018.
5. S. Agarwal, S. P. Yadav, and S. P. Singh, "DEA based estimation of the technical efficiency of state transport undertakings in India," *OPSEARCH*, vol. 47, pp. 216–230, 2010. doi: 10.1007/s12597-011-0035-4.
6. S. Aggarwal and P. Singh, "GSRTC: A study on Gujarat State Road Transport Corporation," *International Journal of Contemporary Research and Review*, vol. 8, no. 3, pp. 28–33, 2017.
7. S. Shah and R. Jain, "A study on performance evaluation of Gujarat State Road Transport Corporation," *International Journal of Advanced Research*, vol. 7, no. 2, pp. 30–34, 2019.
8. S. Sharma and S. Khan, "GSRTC: A comprehensive study of Gujarat State Road Transport Corporation," *International Journal of Scientific Research and Management*, vol. 3, no. 6, pp. 8–15, 2015.
9. S. Sharma, "An analysis of Gujarat State Road Transport Corporation," *International Journal of Science, Technology and Management*, vol. 5, no. 2, pp. 15–20, 2016.
10. Trivedi, S., Operational Efficiency of Gujarat State Road Transport Corporation, *International Journal of Scientific Research (IJSR)*, ISSN: 2277-8179, (2012)