Received: xxx **Accepted:** xxx **Published:** xxx.
DOI: 10.30473/coam.2026.79995.1484
Volume xxx, Issue xxx, p.p. 1-48: xxx
Research Article

Open Access

Control and Optimization in Applied Mathematics - COAM

A Fuzzy Additive DEA Framework for Conditional Short-Run Scale-Direction Assessment with Quasi-Fixed Inputs and Minimum Service-Level Constraints

Marzieh Hossein Nakhai, Mohsen Rostamy Malkhalifeh ✉, Farhad Hosseinzadeh Lotfi

Department of Mathematics, SR.C., Islamic Azad University, Tehran, Iran

✉ **Correspondence:**
Mohsen Rostamy Malkhalifeh
E-mail:
rostamy@srbiau.ac.ir

Abstract. A fuzzy additive data envelopment analysis (DEA) framework is developed to evaluate the scale direction in the short run when part of the capacity is quasi-fixed and a minimum service level is required. A common-realization triangular possibility method is formulated as two sign-dependent linear programs, and the scale parameter is studied over the optimal face to reduce dependence on an arbitrary solver-selected optimal solution. The framework is applied to 2024 National Transit Database (NTD) data for 284 directly operated motor-bus agencies. The additive model, the capacity-restricted model, the deterministic capacity- and service-constrained model, and the baseline fuzzy model identify 8, 16, 34, and 28 agencies on the frontier, respectively. The baseline fuzzy frontier includes 15 agencies with an increasing direction, 11 with no improvement, and 2 with a decreasing direction, according to the specified LP rule. Fleet size and the number of revenue hours are strongly correlated (Pearson $r = 0.9909$). A 21-point service-floor sensitivity analysis, an independent LP implementation, alternative treatments of uncertainty, and stratified-sample models show that the results depend strongly on the assumed technology and on the uncertainty assumptions. In particular, scaling the service floor with the scale parameter changes the classification of the baseline frontier to 3, 5, and 20 agencies in the increasing, zero-improvement, and decreasing categories, respectively. The mathematical results establish feasibility and boundedness, the exactness of the two-sign-branch decomposition, and the identifiability of the effective fuzzy spread $\varepsilon_f = \delta(1-\eta)$. The scale-direction classifications describe the specified global constrained program and should not be interpreted as estimates of local production elasticity.

Keywords. Fuzzy data envelopment analysis, Additive optimization, Conditional scale direction, Quasi-fixed input, Minimum service level, Possibility theory.

MSC. 90C05; 90C70; 91B38.

How to Cite
Hossein Nakhai, M., Rostamy Malkhalifeh, M., Hosseinzadeh Lotfi, F. (2027). "A fuzzy additive DEA framework for conditional short-run scale-direction assessment with quasi-fixed inputs and minimum service-level constraints". *Control and Optimization in Applied Mathematics*, X(x), 1-48. https://doi.org/10.30473/coam.2026.79995.1484

https://mathco.journals.pnu.ac.ir

©2026 by the authors. Lisensee PNU, Tehran, Iran. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 International (CC BY4.0) (http:/creativecommons.org/licenses/by/4.0)

Copyright proofs

1 Introduction

Returns to scale (RTS) are a feature of production technology. They describe how a change in the quantity of resources used changes the level of output. In data envelopment analysis (DEA), the analysis of RTS supports the evaluation of technical efficiency, because it makes it possible to distinguish units operating at an efficient scale from units located in regions of increasing or decreasing returns to scale. This issue is particularly important in services, where decisions about expanding capacity, allocating resources, or changing service levels cannot rely on technical efficiency scores alone.

Standard RTS models in DEA usually depend on assumptions about the production environment. Input and output data are entered as exact numbers, and all production factors are assumed to be adjustable within the defined technology. Neither assumption holds in many real situations. Some resources cannot be changed quickly, and rules or service constraints can limit the set of acceptable reference points. Banker and Morey [3] showed that external or non-discretionary factors should be separated from inputs that managers can control. This separation provides a basis for efficiency analysis when fixed and quasi-fixed factors exist.

Operational or regulatory restrictions add a further layer to RTS analysis. Ouellette, Quesnel, and Vigeant [20] showed that restrictions of this kind can change the shape of the attainable technology; consequently, applying standard RTS rules directly is not always appropriate. Two concepts are central here. A quasi-fixed input limits how much of a resource can be changed during the period studied, and a service constraint removes some reference combinations from the production possibility set.

Imprecise data can also change the shape of the DEA frontier. In a conventional model, each input and output is represented by an exact number, whereas real data can be affected by grouping, reporting errors, rounding, or other sources of error. Lertworasirikul et al. [17] proposed a DEA model based on possibility theory, in which constraints with fuzzy data are treated as fuzzy events and converted into crisp constraints at a chosen possibility level. Khodabakhshi, Gholami, and Kheirollahi [16] extended this idea to the RTS problem and proposed a method for determining RTS when the data are imprecise.

These issues arise when the performance of public transport systems is evaluated. A bus agency takes in funds and vehicles and converts them into transportation services and mobility for passengers. Not all of its inputs can be changed over the same time horizon. Operating costs and their allocation can change fairly easily in the short term, whereas changing the size of the fleet takes longer: the agency must purchase vehicles, secure funding, wait for delivery, maintain the fleet, and plan the replacement of old vehicles. In addition, it is not always acceptable for an agency to reduce its service. Comparing a bus agency with others that perform better because they offer fewer service hours may not provide a useful benchmark.

Corrected proofs

DEA has been widely used to assess the performance of public transport agencies. Chu, Fielding, and Lamar [7] used the method to compare transit agencies with their peers. Odeck and Alkadi [19] examined the efficiency and scale status of Norwegian bus companies. Yu [25] found that the efficient scale of a bus company and its mode of operation can differ across types of bus service. These studies show that the size of a company and its mode of operation can affect the results of the evaluation.

The individual components considered in this study are already known separately. They include the estimation of RTS under uncertainty, the treatment of inputs that cannot be changed easily, and technologies that depend on conditions. Emami and Nasirzade Tabrizi [11] proposed an additive approach to RTS estimation in a regulated environment with quasi-fixed inputs, and Emami, Rostamy Malkhalifeh, and Safdari [12] studied RTS in the presence of undesirable outputs. The contribution of the present work is a formulation that links these features under uncertainty, two sign-dependent branches that together represent the same fuzzy program, and a diagnostic of the scale direction. No novelty is claimed for the fleet restriction or the minimum service level considered individually.

The discretionary input and the two outputs are divided by their sample ranges before the slacks are aggregated. The resulting range-adjusted measure (RAM) objective proposed by Cooper, Park, and Pastor [8] is evaluated on the conditional technology. Although range scaling removes differences due to units of measurement in the slack aggregate, the global scale parameter ξ is not invariant under translation: scaling the actual levels of the variables is different from scaling their differences. Consequently, frontier membership, score discrimination, and direction labels based on ξ are treated as normalization-dependent issues and are examined in Section 4.3.

In the proposed technology, fleet capacity is treated as a quasi-fixed input, and the reference combination must deliver at least a minimum level of service hours. Vehicle Revenue Hours are not used here as a measure of any legal regulation; instead, they define a minimum service requirement within the production technology. This distinction ensures that the observed results are not mistaken for the effect of a specific regulation.

The fuzzy component models the discretionary input and the two outputs by symmetric triangular fuzzy numbers. Their possibility-level constraints are handled through two sign-dependent programming branches, because the observed value of the evaluated agency appears in both the reference and the comparison terms, which leads to the signed coefficient $\lambda_o - \xi$ (Section 3.10). Optimal-face bounds $[\xi_{\min}, \xi_{\max}]$ are used to prevent the direction label from depending on an arbitrarily chosen LP solution (Section 3.7). These bounds serve as a diagnostic for the constrained program and, without additional production assumptions, do not identify a unique local economic elasticity (Proposition 5).

The specific methodological contribution is a conditional range-adjusted additive DEA framework. This framework combines a quasi-fixed fleet ceiling, a fixed minimum-service

Corrected proofs

requirement, and shared-realization triangular imprecision. It uses sign-dependent linear programs and optimal-face ξ intervals to allow a transparent, program-specific assessment of the scale direction. These features distinguish it from established additive DEA models and regulated RTS models, as shown in Table 1.

The empirical illustration uses 2024 National Transit Database data for 284 directly operated U.S. motor-bus agencies. Four nested model specifications are presented to isolate the changes in the comparisons that arise from model design. The growth of the frontier after the restrictions are imposed is interpreted not as evidence of better performance but as reduced discrimination resulting from a smaller reference set.

The remainder of this paper is organized as follows. Section 2 reviews the related literature and positions the proposed framework. Section 3 describes the data, the conditional technology, the deterministic and fuzzy scale programs, and the validation procedures. Section 4 reports the empirical results and sensitivity analyses, which are discussed in Section 5. Section 6 presents the conclusions, limitations, and directions for future research. Agency-level results are reported in Appendices A and B.

2 Literature Review

2.1 Data Envelopment Analysis and Returns to Scale

Data envelopment analysis is a nonparametric method for constructing a production frontier from the observed performance of decision-making units. The Charnes–Cooper–Rhodes (CCR) model [6] was developed under the assumption of constant returns to scale, whereas Banker, Charnes, and Cooper [2] introduced a convexity constraint that made it possible to evaluate technology under variable returns to scale (VRS). This development helped to separate technical inefficiency from scale-related effects and became the foundation for a substantial body of subsequent RTS research.

Under VRS technology, an efficient unit may operate in a region of increasing, constant, or decreasing returns to scale. Increasing returns to scale (IRS) apply when a proportional expansion in the scale of activity is associated with a more-than-proportional increase in output. Under decreasing returns to scale (DRS), further expansion in scale is associated with a less-than-proportional increase in output. Constant returns to scale (CRS) describe a situation in which the unit lies on a region of the frontier where proportional scaling preserves the overall input–output relationship [22].

The piecewise-linear structure of the DEA frontier can make the determination of RTS difficult when multiple solutions exist. In such circumstances, relying on a single dual solution or supporting hyperplane may cause the classification to depend on the particular solution re-

Corrected proofs

turned by the numerical solver. For this reason, methods that examine intervals of scale-related parameters are important for producing more reliable classifications of frontier units [4, **?**].

Additive models differ structurally from radial models. In these models, input excesses and output shortfalls enter the objective function simultaneously. A difficulty is that directly summing slacks measured in different units can make the result sensitive to the numerical magnitudes of the variables. Cooper, Park, and Pastor [8] proposed the range-adjusted measure to reduce this dependence on units of measurement. The logic of that approach provides the basis for scaling the discretionary input and outputs in the present model.

Khodabakhshi et al. [16] proposed an additive model based on the parameter ξ for determining returns to scale. In that framework, the input and output vectors of an efficient unit are scaled by ξ, and the behavior of the additive objective around the frontier unit provides the basis for identifying conventional increasing, constant, or decreasing returns to scale (IRS, CRS, or DRS). The relevance of this approach to the present study is that it can also be extended to environments with imprecise data.

2.2 Nondiscretionary and Quasi-Fixed Inputs

The assumption that managers have complete control over all inputs does not hold in many real-world environments. Banker and Morey [3] showed that exogenous inputs and outputs cannot be reduced or increased in an efficiency program in the same way as discretionary variables. Under such conditions, the reference set must be defined so that it respects constraints beyond the control of the unit being evaluated.

Economically, a quasi-fixed input is not identical to a completely nondiscretionary input. Its level may change in the long run, but it cannot be adjusted rapidly within the current decision horizon. This distinction matters for RTS analysis because proportional scaling of all inputs may imply changes that cannot actually be achieved in the short run.

Sueyoshi and Sekitani [24], in their study of returns to scale in dynamic DEA, distinguished between variable and quasi-fixed inputs. Blavet, Ouellette, and Vigeant [5], in an application to Québec hospitals, developed the α-RTS approach for settings with quasi-fixed inputs. These studies emphasize a common point: the interpretation of scale must be consistent with the time horizon over which resources can be adjusted.

In an urban bus system, Vehicles Operated in Maximum Service (VOMS) represents a capacity measure that cannot easily be changed in the short run. Adding or removing a substantial number of vehicles requires capital and operational decisions that differ from adjustments to current operating expenditures. Treating VOMS as a quasi-fixed input therefore brings the RTS estimate closer to the short-run operating conditions faced by the system.

Corrected proofs

2.3 Constrained Technologies and Service Requirements

A restriction on production technology is not merely another form of nondiscretionary input. A quasi-fixed input limits the capacity or resource available to the unit, whereas an operational constraint may exclude some technologically feasible points from the comparison set.

Ouellette et al. [20] examined this issue in regulated environments and showed that a regulatory constraint can alter how returns to scale are inferred from a DEA model. The importance of this result is that the frontier used for RTS evaluation should be consistent with the constraints actually faced by the unit.

In public services, a restriction does not always take the form of an explicit legal regulation. Requirements relating to coverage, accessibility, quality, or preservation of a specified service level can also restrict acceptable reference combinations. If such conditions are ignored, a unit may be compared with a reference that has more favorable efficiency ratios partly because it provides a lower level of service.

In the present model, Vehicle Revenue Hours (VRH) are used to define this restriction. The reference combination must provide at least a specified fraction of the service level observed for the unit under evaluation. The parameter ρ controls the stringency of this condition and is set equal to one in the baseline specification (Section 3.4). This variable is interpreted strictly as a measure of service level and not as a direct indicator of a legal regulation.

2.4 Fuzzy Data Envelopment Analysis

Conventional DEA models operate with exact values. This assumption becomes problematic when data cannot be represented fully by a single crisp value because of reporting practices, rounding, aggregation, or other sources of imprecision. Fuzzy DEA allows this imprecision to be incorporated directly into the production technology [14].

Lertworasirikul et al. [17] proposed a possibility-theory-based approach to fuzzy DEA. In their method, constraints containing fuzzy data are interpreted as fuzzy events and transformed into deterministic programs by specifying a possibility level. The advantage of this structure is that it retains the solvability of the model with mathematical-programming tools while explicitly representing imprecision.

Khodabakhshi et al. [16] extended the use of imprecise data to the RTS problem. Their model showed that an additive RTS framework can also be defined with fuzzy data, provided that the fuzzy relationships are correctly transformed into deterministic equivalents. Subsequent studies have examined other aspects of returns to scale in fuzzy environments. Payan, Hajihosseini, and Nosrat [21] investigated one-sided returns to scale in crisp and fuzzy settings, while

Corrected proofs

Amirteimoori, Allahviranloo, and Fakhr Mousavi [1] proposed a fully fuzzy range-adjusted framework for two-stage processes.

The present model does not fuzzify all variables. Imprecision is introduced only for the discretionary input and the two outputs, while the quasi-fixed capacity and the minimum-service variable remain crisp. This structure allows the effect of production-data imprecision to be examined separately from the capacity and service-level restrictions.

2.5 DEA Applications in Public Transportation

The activity of a public-transport operator cannot be reduced to a single productivity ratio. Cost, fleet size, service supply, passenger volume, and passenger distance each capture a different aspect of performance. DEA permits the simultaneous evaluation of multiple inputs and outputs without requiring the specification of a parametric production function [9].

Chu et al. [7] used DEA to measure public-transport performance and compare agencies with similar units. Odeck and Alkadi [19] examined the efficiency of subsidized Norwegian bus companies and also analyzed scale status. Their results showed increasing returns for a subset of companies, particularly smaller operators. Yu [25], in a study of multimode bus systems, likewise showed that RTS status is not identical across different service structures.

The importance of these studies for the present research lies in the distinction between the apparent size of an organization and its position on the production technology. Two operators with similar fleet sizes may occupy different regions of the frontier because of differences in service levels, demand structure, and resource use. Scale analysis should reflect these production differences in the model structure.

2.6 Research Gap and Positioning

The focused literature assessment used combinations of the terms “additive returns to scale”, “regulated DEA”, “quasi-fixed inputs”, “fuzzy DEA”, “possibility approach”, “range-adjusted measure”, and “complementary slackness” for studies available through October 2026. Records and full-text descriptions were checked against journal and publisher pages, indexed bibliographic pages, and original references, with particular attention to Khodabakhshi et al. [16], Ouellette et al. [20], Blavet et al. [5], Emami and Nasirzade Tabrizi [11], and Emami et al. [12]. This targeted comparison is not treated as an exhaustive systematic review; the contribution is defined positively by mathematical and technological differences rather than by an unverifiable claim that no related model exists.

Corrected proofs

In general, these issues have been studied separately. RTS models with quasi-fixed inputs do not typically place fuzzy imprecision at the center of the analysis, and fuzzy RTS models do not necessarily incorporate a service-level-constrained technology and quasi-fixed capacity simultaneously. The question therefore remains how the RTS classification changes when a unit faces imprecise data, operates with quasi-fixed capacity in the short run, and is compared against a reference set that must preserve a specified level of service.

In this study, these three features are incorporated simultaneously into a conditional additive model. The model structure makes it possible to observe the effect of each component on the frontier separately: the standard additive model provides the benchmark; introducing VOMS captures the effect of quasi-fixed capacity; the VRH constraint defines a service-oriented technology; and the fuzzy formulation reveals the sensitivity of the frontier and the RTS classification to imprecision in input and output data. Table 1 positions the proposed specification relative to the most closely related models.

Table 1: Positioning of the proposed specification relative to closely related models.

Study	Closest subject	Distinguishing formulation	Relevance
Khodabakhshi et al. (2010) [16]	Additive ξ-based RTS with imprecise data	No combined fleet ceiling and service floor	Foundational fuzzy additive RTS
Ouellette et al. (2012) [20]	RTS under regulatory constraints	Not the present shared-realization fuzzy LP	Conditional technology precedent
Blavet et al. (2023) [5]	Quasi-fixed inputs and α-RTS	Nonconvex technology	Alternative short-run scale definition
Amirteimoori et al. (2025) [1]	Fully fuzzy range-adjusted two-stage model	Strong complementary slackness	Alternative treatment of dual/multiple optima
Emami and Nasirzade Tabrizi (2016) [11]	Regulated additive RTS; quasi-fixed inputs	Crisp precursor	Closest constraint/RTS precursor
Emami et al. (2014) [12]	Regulated RTS with undesirable outputs	Not the joint triangular fuzzy formulation	Relevant regulated RTS antecedent
Present formulation	Conditional convex reference sets plus shared triangular η-cuts	Two sign branches; optimal-face ξ intervals	Joint implementation; no novelty claim for individual constraints

3 Model Framework

The methodological objective of this study is to estimate conditional short-run returns to scale when part of the production capacity is quasi-fixed over the horizon considered, the reference

Corrected proofs

set must preserve a specified minimum level of service, and some input and output values are imprecise. To this end, the additive DEA framework based on the scale parameter ξ is extended to a conditional technology, after which its fuzzy version is formulated using a possibility approach.

To separate the effects of quasi-fixed capacity, the service constraint, and data imprecision, the model is implemented in four sequential specifications: the standard additive model; the model with a quasi-fixed input; the deterministic model with a quasi-fixed input and a service constraint; and the proposed fuzzy model. These specifications make it possible to distinguish the effect of restricting quasi-fixed capacity, the effect of imposing the minimum-service requirement, and the effect of fuzzy imprecision in the production data. Their results are compared in Section 4.2.

3.1 Data and Sample Construction

The empirical observations were extracted from the 2024 reporting year of the TS2.1 Service Data and Operating Expenses Time Series by Mode dataset of the U.S. National Transit Database (NTD). The unit of analysis was a directly operated urban motor-bus system. Observations were retained only if they met the following filters: reporting year 2024, Motor Bus mode, Directly Operated service, Active agency status, Full Reporter class, Urban module, and Operating status [13].

Applying these criteria yielded 288 observations. Four observations with zero Passenger Miles Traveled were excluded so that the production variables used in the model contained only positive values. The final sample therefore consisted of 284 decision-making units.

Data reconciliation. The twelve operational tabs contain agency–mode–service records for 2015–2024. Filtering the 2024 columns for operating cost, Vehicles Operated in Maximum Service (VOMS), Vehicle Revenue Hours (VRH), Unlinked Passenger Trips (UPT), and Passenger Miles Traveled (PMT), and then joining by NTD ID, returned 288 eligible observations. Four records reported zero PMT (NTD IDs 40044, 40056, 40102, and 60090); these were excluded from the strictly positive input–output analysis without treating a reported zero as a confirmed error. The final 284-agency extract reproduces Table 2 to its reported precision.

3.2 Model Variables

The production technology includes one discretionary input, one quasi-fixed input, two desirable outputs, and one variable defining the service constraint. Each observed decision-making unit (DMU) j, $j = 1, \ldots, n$, is described by the following variables:

Corrected proofs

$$x_j = \text{Total Operating Expenses}, \tag{1}$$

$$k_j = \text{Vehicles Operated in Maximum Service}, \tag{2}$$

$$y_{1j} = \text{Unlinked Passenger Trips}, \tag{3}$$

$$y_{2j} = \text{Passenger Miles Traveled}, \tag{4}$$

$$g_j = \text{Vehicle Revenue Hours}. \tag{5}$$

Total Operating Expenses were treated as the discretionary input. Vehicles Operated in Maximum Service (VOMS) were entered as the quasi-fixed input. The two production outputs are Unlinked Passenger Trips and Passenger Miles Traveled, and Vehicle Revenue Hours were used to define the minimum service level. Operating expenditure is more adjustable in the short run than fleet capacity. By contrast, a substantial change in VOMS depends on capital decisions, vehicle acquisition or retirement, and operational constraints; therefore, this variable was not included in the proportional scaling of the input–output vector. VRH is likewise not a conventional input or output of the efficiency model; rather, it defines the boundary of the acceptable reference set in terms of service level.

Variable selection was also examined to avoid entering highly similar indicators simultaneously. The correlation between Total Operating Expenses and Vehicle Operations Operating Expenses was approximately 0.997; therefore, only Total Operating Expenses was retained in the model. The correlation between Vehicle Revenue Hours and Vehicle Revenue Miles was approximately 0.965, and VRH was selected to define the service constraint because it more directly reflects the duration of service supply [10].

3.3 Scaling of Additive Variables

In an additive model, the sum of slacks can be affected by the units of measurement of the variables. To reduce this dependence, the discretionary input and outputs were scaled by their observed ranges. The range of the discretionary input was defined as

$$R_x = \max_j x_j - \min_j x_j, \tag{6}$$

and the range of each output as

$$R_{y_r} = \max_j y_{rj} - \min_j y_{rj}, \qquad r = 1, 2. \tag{7}$$

The values used in the model are

$$\overline{x}_j = \frac{x_j}{R_x} \tag{8}$$

and

$$\overline{y}_{rj} = \frac{y_{rj}}{R_{y_r}}, \qquad r = 1, 2. \tag{9}$$

Corrected proofs

Range normalization divides each variable's nonnegative slack by its observed sample range. With one discretionary input, two outputs, and the factor $1/(m+s) = 1/3$, the resulting additive efficiency objective is exactly the range-adjusted measure (RAM) objective of Cooper, Park, and Pastor [8], evaluated over the present conditional reference technology. The ξ-dependent scale program modifies the ray constraints, not this slack objective. A positive affine transformation preserves the ordering of the raw observations within each individual variable, but neither the efficiency rankings under alternative weights nor the ξ-based scale directions are invariant to arbitrary changes of weights or origins.

3.4 Conditional Production Technology

If the service indicator is exactly proportional to fleet capacity, $g_j = c\,k_j$ for a common $c > 0$, then at $\rho = 1$ the inequalities $\sum_j \lambda_j k_j \le k_o$ and $\sum_{j=1}^{n} \lambda_j g_j \ge \rho\, g_o$ force $\sum_j \lambda_j k_j = k_o$. Thus, the joint restrictions become an exact fleet-utilization equality in this limiting case. A change in efficient-unit counts cannot by itself distinguish genuine service preservation from capacity-utilization selection; an assessment of their empirical correlation and binding status is necessary (Section 4.1 and Section 4.8).

For the unit under evaluation, o, the conditional production possibility set was defined as

$$T_o(\rho) = \left\{ (\overline{x}, \overline{y}) : \begin{array}{ll} \displaystyle\sum_{j=1}^{n} \lambda_j \overline{x}_j \le \overline{x}, & \\ \displaystyle\sum_{j=1}^{n} \lambda_j \overline{y}_{rj} \ge \overline{y}_r, & r = 1, 2, \\ \displaystyle\sum_{j=1}^{n} \lambda_j k_j \le k_o, & \\ \displaystyle\sum_{j=1}^{n} \lambda_j g_j \ge \rho\, g_o, & \\ \displaystyle\sum_{j=1}^{n} \lambda_j = 1, & \\ \lambda_j \ge 0 & \end{array} \right\}. \tag{10}$$

The constraint

$$\sum_{j=1}^{n} \lambda_j k_j \le k_o \tag{11}$$

prevents the unit under evaluation from being benchmarked against a reference combination that would require fleet capacity greater than its quasi-fixed capacity. The service constraint was defined as

Corrected proofs

$$\sum_{j=1}^{n} \lambda_j g_j \geq \rho\, g_o. \tag{12}$$

The parameter ρ determines the stringency of the service-preservation requirement. In the baseline specification,

$$\rho = 1$$

was imposed; hence, the reference combination must provide at least the observed VRH of the evaluated agency. The primary service-floor scan used 21 equally spaced values $\rho = 0.80, 0.81, \ldots, 1.00$; the endpoint was further examined at $\rho = 0.995$, 0.999, 0.9995, 0.9999, 0.99999, and 0.999999 to resolve the interaction between a tightening service constraint and the numerical efficiency threshold (Section 4.8).

3.5 Deterministic Additive Efficiency Model

For each unit o, the additive inefficiency is obtained by solving the following program:

$$\max \; D_o = \frac{1}{3}\left(s_x^- + s_1^+ + s_2^+\right) \tag{13}$$

subject to

$$\begin{aligned}
&\sum_{j=1}^{n} \lambda_j \overline{x}_j + s_x^- = \overline{x}_o, \\
&\sum_{j=1}^{n} \lambda_j \overline{y}_{1j} - s_1^+ = \overline{y}_{1o}, \\
&\sum_{j=1}^{n} \lambda_j \overline{y}_{2j} - s_2^+ = \overline{y}_{2o}, \\
&\sum_{j=1}^{n} \lambda_j k_j \leq k_o, \\
&\sum_{j=1}^{n} \lambda_j g_j \geq \rho\, g_o, \\
&\sum_{j=1}^{n} \lambda_j = 1, \\
&\lambda_j,\; s_x^-,\; s_1^+,\; s_2^+ \geq 0.
\end{aligned} \tag{14}$$

In these relationships, s_x^- denotes the input excess and s_r^+ denotes the shortfall in output r relative to the reference combination. Because all three slacks are placed on comparable scales after range scaling, their average is used in the objective function. The efficiency score was calculated as

Corrected proofs

$$E_o = 1 - D_o^{*}. \tag{15}$$

Units whose optimal inefficiency value was no greater than 10^{-7}, were identified as efficient. The stability of this threshold was examined by repeating the analysis for values between 10^{-10} and 10^{-6} (Section 4.10).

3.6 Conditional Short-Run Scale Program

For a differentiable conditional production function $y = f(x; k)$, the local single-factor elasticity is $\left(\frac{x}{y}\right)\left(\frac{\partial f}{\partial x}\right)$ at fixed k. The finite ξ of program (17), by contrast, is selected globally over a piecewise-linear conditional reference set, with k_o and the service floor fixed. The two objects need not coincide. A policy-sensitive alternative replaces the fixed service floor $\sum_j \lambda_j g_j \geq \rho\, g_o$ by $\sum_j \lambda_j g_j \geq \rho\, \xi\, g_o$. The latter remains linear in the scale program but changes the economic question by requiring service provision to expand with the scale parameter (Section 4.8).

Returns to scale were determined only for units that lay on the efficiency frontier under the relevant model. For an efficient unit o, the parameter ξ was introduced into the discretionary-input and output vector:

$$\max_{\xi,\,\lambda,\,s} \; D_o^{\xi}(\xi) = \frac{s_x^{-} + s_1^{+} + s_2^{+}}{3} \tag{16}$$

subject to

$$\begin{aligned}
&\sum_{j=1}^{n} \lambda_j \overline{x}_j + s_x^{-} = \xi\, \overline{x}_o, \\
&\sum_{j=1}^{n} \lambda_j \overline{y}_{rj} - s_r^{+} = \xi\, \overline{y}_{ro}, \quad r = 1, 2, \\
&\sum_{j=1}^{n} \lambda_j k_j \leq k_o, \\
&\sum_{j=1}^{n} \lambda_j g_j \geq \rho\, g_o, \\
&\sum_{j=1}^{n} \lambda_j = 1, \\
&\lambda_j,\; s_x^{-},\; s_r^{+},\; \xi \geq 0.
\end{aligned} \tag{17}$$

In (17), ξ proportionally scales the discretionary operating expenditure and the two passenger outputs; the quasi-fixed fleet capacity and the service floor remain fixed. The optimum of the scale program is denoted by $D_o^{\xi *}$, whereas D_o^{*} is reserved for the efficiency program (13). The direction labels below refer exclusively to the global constrained scale program and do not identify a local production elasticity.

Corrected proofs

3.7 Control for Multiple Optimal Solutions

For each sign branch b (defined in Section 3.10), maximize the scale objective D_o^{ξ} to obtain $D_{o,b}^{\xi*}$, and let $D_o^{\xi*} = \max_b D_{o,b}^{\xi*}$. Retain the branches satisfying $D_{o,b}^{\xi*} \geq D_o^{\xi*} - \tau_{\text{obj}}$. Within each retained branch, obtain both extremes of ξ over the near-optimal face $D_o^{\xi} \geq D_o^{\xi*} - \tau_{\text{obj}}$. The global $\xi_{\min}$ and $\xi_{\max}$ are the minimum and maximum, respectively, across all retained branches.

A linear program may return one of several optimal solutions. If different values of ξ exist within the set of optimal or near-optimal solutions, relying on the value of ξ reported by a single solver run can make the RTS classification dependent on the solver's numerical selection.

After the maximum scale-program value $D_o^{\xi*}$ has been determined, two secondary LPs use the shared near-optimality condition

$$D_o^{\xi} \geq D_o^{\xi*} - \tau_{\text{obj}}. \tag{18}$$

In the first program, ξ was minimized:

$$\xi_o^{\min} = \min \xi, \tag{19}$$

and in the second program, its maximum value was obtained:

$$\xi_o^{\max} = \max \xi. \tag{20}$$

The global optimum is computed across all feasible sign branches:

$$D_o^{\xi*} = \max_{b \in \{+,-\}} D_{o,b}^{\xi*}. \tag{21}$$

The corresponding optimal-face interval is defined by the extrema (19) and (20):

$$\left[\xi_o^{\min}, \xi_o^{\max}\right] = \left[\min_{b \in B^*} \xi_{o,b}^{\min}, \max_{b \in B^*} \xi_{o,b}^{\max}\right], \tag{22}$$

where B^* is the set of retained branches. The numerical thresholds, fixed for the reported baseline analysis, are

$$\tau_{\text{eff}} = 10^{-7}, \qquad \tau_{\text{obj}} = \tau_D = 10^{-8}, \qquad \tau_{\xi} = 10^{-6}. \tag{23}$$

The single classification rule is as follows. If $D_o^{\xi*} \leq \tau_D$, assign zero-improvement (ZI), irrespective of the ξ interval. Otherwise, assign increasing-direction (ID) only if $\xi_{\min} > 1 + \tau_{\xi}$, or decreasing-direction (DD) only if $\xi_{\max} < 1 - \tau_{\xi}$. All other positive-improvement cases are unresolved. Here, τ_{obj} specifies the admissible objective-face margin, τ_D the scale-objective decision threshold, τ_{ξ} the distance from unity, and τ_{eff} the separate frontier-membership threshold applied to D_o^* from (13); none of them is the fuzzy spread ε_f. The interval $[1, 1]$ of a ZI unit is not counted as a positive-improvement interval crossing unity.

This rule separates the two distinct optimization objectives and prevents near-unity numerical variation from being mistaken for an economically meaningful local scale effect.

Corrected proofs

3.8 Fuzzy Representation of Production Data

Imprecision was introduced for the three scaled variables consisting of the discretionary input and the two outputs. For

$$z_j \in \{\overline{x}_j,\ \overline{y}_{1j},\ \overline{y}_{2j}\}, \tag{24}$$

the symmetric triangular fuzzy number was defined as

$$\widetilde{z}_j = \big[(1-\delta)z_j,\ z_j,\ (1+\delta)z_j\big]. \tag{25}$$

The parameter δ controls the relative width of the triangular fuzzy number. The sensitivity scenarios use $\delta \in \{0.025, 0.05, 0.10\}$, with $\delta = 0.05$ in the baseline specification. These scenario widths are assumptions rather than empirical estimates of NTD reporting error; VOMS and VRH remain crisp so that their operational restrictions are distinct from the fuzzy production data.

3.9 Possibility Level and Deterministic Equivalent

A possibility level η transforms each symmetric triangular fuzzy observation into an interval. The effective spread is $\varepsilon_f = \delta(1-\eta)$, and the lower and upper endpoint multipliers are $a = 1-\varepsilon_f$ and $b = 1+\varepsilon_f$, respectively, as specified in (26) and (27):

$$a = 1-\delta(1-\eta), \tag{26}$$

$$b = 1+\delta(1-\eta). \tag{27}$$

At possibility level η, the triangular interval for each observed value z is $[az,\ bz]$. The three possibility levels are $\eta \in \{0.50, 0.75, 1.00\}$, with $\eta = 0.75$ in the baseline specification. When $\eta = 1$, the effective spread vanishes ($\varepsilon_f = 0$), so $a = b = 1$ and the model reduces to the deterministic formulation.

3.10 Sign-Dependent Fuzzy Formulation

In the RTS model, unit o appears both in the reference sum and in the scaled vector. The following expression was defined to isolate its contribution:

$$q_o = \lambda_o - \xi. \tag{28}$$

The appropriate lower- and upper-bound coefficients for the term associated with unit o depend on the sign of q_o. Because the sign of this expression is not known before the model is solved, a

Corrected proofs

single linear equivalent that ignores the sign can generate incorrect fuzzy bounds. The problem was therefore solved in two branches. In the case

$$q_o \geq 0,$$

the input constraint becomes

$$a\Big(\sum_{j=1}^{n} \lambda_j \overline{x}_j - \xi\, \overline{x}_o\Big) + s_x^- \leq 0, \tag{29}$$

and the constraint for each output becomes

$$b\Big(\sum_{j=1}^{n} \lambda_j \overline{y}_{rj} - \xi\, \overline{y}_{ro}\Big) - s_r^+ \geq 0, \qquad r = 1, 2. \tag{30}$$

This branch is solved together with the condition

$$\lambda_o - \xi \geq 0. \tag{31}$$

In the second case,

$$q_o \leq 0,$$

the input constraint takes the form

$$a \sum_{j \neq o} \lambda_j \overline{x}_j + b\,(\lambda_o - \xi)\,\overline{x}_o + s_x^- \leq 0, \tag{32}$$

and, for each output,

$$b \sum_{j \neq o} \lambda_j \overline{y}_{rj} + a\,(\lambda_o - \xi)\,\overline{y}_{ro} - s_r^+ \geq 0, \qquad r = 1, 2, \tag{33}$$

subject to the condition

$$\lambda_o - \xi \leq 0. \tag{34}$$

The two sign branches share the fleet, minimum-service, convexity, and nonnegativity restrictions. Their feasible sets form an exact union under the common-realization convention (Proposition 3), including the shared boundary $\lambda_o = \xi$. Maximization is carried out in each branch, and the larger optimum is selected. The ξ interval is then constructed over every branch achieving the global optimum within the declared objective tolerance (Section 3.7). The main theoretical properties of the formulation are stated in the following propositions.

Proposition 1 (Feasibility and boundedness)**.** Assume that n is finite, that the observed discretionary inputs and outputs are strictly positive, that the range denominators are positive, that k_o and g_o are positive, that the fleet and service observations are nonnegative and nonzero, that $0 \leq \rho \leq 1$, and that $0 \leq \varepsilon_f < 1$. Then both the crisp scale program (17) and each closed sign branch of its fuzzy counterpart are feasible and attain finite optima.

Corrected proofs

Proof. The point $\lambda_o = \xi = 1$, $\lambda_j = 0$ for $j \neq o$, with zero slacks, satisfies the convexity, capacity, and service inequalities with $q_o = 0$; it belongs to both branches. Under convexity, all reference sums of nonnegative observations are bounded. In the crisp model, the positive-output equalities imply $\xi\, \overline{y}_{ro} \leq \sum_{j=1}^{n} \lambda_j \overline{y}_{rj}$ and hence bound ξ; all slacks are then bounded by the same input and output levels. In the positive-sign branch, $q_o \geq 0$ implies $0 \leq \xi \leq \lambda_o \leq 1$. In the negative-sign branch, each fuzzy output inequality (33), with nonnegative output slack, implies $a\, \xi\, \overline{y}_{ro} \leq b \sum_{j \neq o} \lambda_j \overline{y}_{rj} + a\, \lambda_o \overline{y}_{ro} - s_r^+$; since $a > 0$, this also bounds ξ. The input and output inequalities bound every nonnegative slack, and the feasible sets are closed. The linear objectives thus attain finite maxima. □

Proposition 2 (Effective-spread identifiability)**.** With the previously defined coefficients, let $\varepsilon_f = \delta(1 - \eta)$, $a = 1 - \varepsilon_f$, and $b = 1 + \varepsilon_f$. If two pairs (δ, η) have the same ε_f, their transformed LP coefficient matrices, feasible sets, and objectives coincide; hence, so do their optimal values and complete sets of optimizers.

Proof. Every triangular η-cut endpoint is $z(1 \pm \varepsilon_f)$, and the remaining constraints are identical. The nine parameter combinations in Table 9 therefore correspond to only five distinct programs. □

Proposition 3 (Exact sign partition)**.** Suppose that each appearance of the evaluated observation uses the same triangular fuzzy realization. Write $q_o = \lambda_o - \xi$. At any η-level, multiplication by q_o uses the lower endpoint if $q_o \geq 0$ and the upper endpoint if $q_o \leq 0$ for the input constraint, with the reverse ordering for the output inequalities. Thus, the feasible set is precisely the union of the two closed polyhedra specified by (29)–(34). At $q_o = 0$, the evaluated observation contributes zero in either representation, so the branches coincide on their common boundary. It follows that maximization over the union equals the larger of the two branch optima.

Proof. The two weak inequalities $q_o \geq 0$ and $q_o \leq 0$ cover all real q_o, and each branch applies the exact interval endpoint rule for the sign in that half-space. At $q_o = 0$, both products vanish. No point is omitted or introduced by the partition. □

Proposition 4 (Monotonicity of the optimistic formulation)**.** For $0 \leq \varepsilon_1 \leq \varepsilon_2 < 1$, the feasible region of either sign branch at ε_1 is contained in that of the same branch at ε_2, provided that the crisp capacity, service, and convexity restrictions are unchanged. Consequently, the maximum slack objective is nondecreasing in ε, the scores $E = 1 - D$ cannot increase, and the set of exactly efficient units cannot expand as ε increases.

Proof. For $q_o \geq 0$, the input expression multiplying a is nonnegative and the output expression multiplying b is nonnegative; decreasing a and increasing b relax their respective inequalities. For $q_o \leq 0$, the input left-hand side $a \sum_{j \neq o} \lambda_j \overline{x}_j + b\, q_o \overline{x}_o + s_x^-$ decreases with ε, while the output left-hand side $b \sum_{j \neq o} \lambda_j \overline{y}_{rj} + a\, q_o \overline{y}_{ro} - s_r^+$ increases. Hence, all points feasible at ε_1

Corrected proofs

remain feasible at ε_2, and the maximum slack objective is nondecreasing in ε. These statements are structural, not independent empirical validation. □

Proposition 5 (Limits of local RTS identification)**.** Without additional assumptions on the production function and a specified local directional perturbation, the sign of an optimal finite $\xi - 1$ in (17) does not identify the sign of a local conditional output elasticity minus one.

Proof. Consider a finite set of distinct positive discretionary input observations (x_j, y_j) and a differentiable positive production function f passing through these observations near the evaluated input x_o. Choose a smooth bump function ψ supported in a neighborhood of x_o that contains no other x_j, with $\psi(x_o) = 1$, and define $f_t(x) = f(x) + t\,(x - x_o)\,\psi(x)$. All observed points remain unchanged, since $f_t(x_j) = f(x_j)$ for every j, whereas $f_t'(x_o) = f'(x_o) + t$. For sufficiently small $|t|$, positivity and any strict local monotonicity of f are preserved. Hence, the same finite dataset, and consequently the same linear program and optimal ξ, are consistent with different local elasticities $x_o f_t'(x_o)/f_t(x_o)$. A unique local elasticity therefore cannot be inferred from ξ without extra identifying restrictions. The terminology "scale direction" is retained for the derived global LP categories. □

The counterexample is intentionally formulated in a reduced one-input, one-output setting. It assumes that every observed pair lies on the perturbed differentiable production function and establishes the failure of identification from finite observations in that reduced setting; it does not assert equivalence between this constructed technology and the complete two-output, fleet- and service-constrained production set.

Fuzzy dependence check. The branch construction assumes that the observed input and output of the evaluated agency form one shared uncertain realization wherever they occur in the same inequality. A separate-occurrence interval calculation instead assigns independent endpoint choices to the reference and evaluation terms; at $\varepsilon_f = 0.0125$, this comparison produces no exactly efficient unit among the 284 agencies. Its different self-comparison property precludes treating those scores as interchangeable with the shared-realization model.

Notation summary. ξ is the nonnegative proportional ray parameter of the global scale program; $\delta \geq 0$ is the assumed relative half-width of each symmetric triangular fuzzy observation; $\eta \in [0, 1]$ is the possibility-cut level; $\varepsilon_f = \delta(1 - \eta)$ is the only effective spread entering the deterministic coefficients $a = 1 - \varepsilon_f$ and $b = 1 + \varepsilon_f$; and $\rho \in [0, 1]$ defines the minimum-service fraction. None of τ_{eff}, τ_D, τ_{obj}, or τ_ξ is a fuzzy width.

3.11 Baseline Specifications and Sensitivity Analysis

The analysis compares four nested model specifications: (i) the unrestricted additive model; (ii) the additive model with a quasi-fixed fleet-capacity ceiling; (iii) its deterministic extension

Corrected proofs

with a minimum-service floor; and (iv) the shared-realization triangular fuzzy extension of the same conditional technology.

The deterministic baseline uses $\rho = 1$, $\delta = 0$, and $\eta = 1$; the baseline fuzzy specification uses $\rho = 1$, $\delta = 0.05$, and $\eta = 0.75$. The sensitivity calculations cover all nine combinations of $\delta \in \{0.025, 0.05, 0.10\}$ and $\eta \in \{0.50, 0.75, 1.00\}$, corresponding to five distinct effective spreads (Section 4.7).

The primary service-floor analysis comprises the 21-point grid $\rho = 0.80, 0.81, \ldots, 1.00$, complemented by six finer endpoint values (Section 4.8). The representative values $\rho = 0.90$, 0.95, and 1.00 indicate alternative service-stringency scenarios, not observed legal thresholds.

3.12 Computational Implementation

Computations used Python, NumPy 2.3.5, and SciPy 1.17.0, with the bundled HiGHS 1.8.0 optimization engine [15]. The primal and dual feasibility tolerances were both 10^{-9}. The model-specific absolute thresholds were $\tau_{\text{eff}} = 10^{-7}$, $\tau_{\text{obj}} = \tau_D = 10^{-8}$, and $\tau_\xi = 10^{-6}$ (Equation (23)). The original-variable sample ranges were computed once for each analytical reference set. The effective spread ε_f and the service parameter ρ index the respective scenario analyses; the full coefficient matrices and decision rules follow Sections 3.5–3.10.

For numerical validation, the conditional crisp and fuzzy coefficient matrices were also independently assembled and solved with the HiGHS interior-point algorithm. The reported objective differences, frontier membership, and scale directions are evaluated at explicitly stated numerical thresholds. The full agency-level record retains the input and output values and the corresponding scores; the software-version specifications permit independent computational replication.

3.13 Validation and Numerical Checks

The limiting case $\eta = 1$ yields $\varepsilon_f = 0$ and $a = b = 1$, so the fuzzy LP algebraically reduces to its deterministic counterpart. Equality of the numerical output at this limit is an implementation-consistency check, not a validation of the possibility model.

The eight frontier units of the unrestricted additive model were checked against a Banker–Charnes–Cooper (BCC) [2] supporting-hyperplane calculation, with matching direction labels under the original convention (Section 4.9). This benchmark exercise verifies only the unrestricted comparison; it does not supply an independent verification of the conditional or fuzzy programs. The efficiency classification threshold was also varied from 10^{-10} to 10^{-6}. Across

Corrected proofs

this range, the number of efficient units in both the deterministic model and the baseline fuzzy specification remained unchanged (Section 4.10).

Potentially extreme observations were also examined separately. Removing the unit with the largest ξ, and then removing the three observations identified by a stringent three-times-interquartile-range rule applied to logarithmic ratios, did not change the RTS classification of the other frontier units common to the samples. The purpose of this test was to assess the extent to which the frontier depended on extreme observations (Section 4.11).

Independent numerical implementation. A second, independently assembled inequality matrix was solved with the HiGHS interior-point algorithm for all 284 agencies, separately for $\varepsilon_f = 0$ and $\varepsilon_f = 0.0125$. All problems terminated optimally. The maximum absolute differences in D were 4.75×10^{-16} and 5.38×10^{-16}, respectively; no frontier-membership decisions differed at 10^{-7}. These checks validate the implementation of the stated finite LPs, not the economic interpretation of ξ as a derivative-based scale elasticity.

3.14 Worked Three-Agency Numerical Example

Consider three agencies A, B, and C with one discretionary input $x = (1, 1.5, 2)$, two equal outputs $y_1 = y_2 = (1, 1.7, 2.6)$, a quasi-fixed capacity $k = (2, 1.9, 2.1)$, and service $g = (1, 1.4, 1.5)$. Evaluate A with $\rho = 1$. The observed ranges are $R_x = 1$ and $R_y^1 = R_y^2 = 1.6$. This is an illustrative constructed example, not an observation drawn from the NTD sample.

In model (17), the reference weights $\lambda = (0, \frac{1}{2}, \frac{1}{2})$ and $\xi = 1.75$ are feasible with $s_x^- = 0$ and $s_1^+ = s_2^+ = 0.25$ (in range-adjusted units). The fleet ceiling is binding: $0.5(1.9) + 0.5(2.1) = 2$. Service provision equals $0.5(1.4) + 0.5(1.5) = 1.45 \geq 1$; the weighted input equals 1.75, and each weighted output equals 2.15. Consequently, $D = (0 + 0.25 + 0.25)/3 = 1/6$. Direct LP optimization confirms the scale-program optimum $D_o^{\xi*} = 1/6$ at $\xi^* = 1.75$.

Under the alternative proportional service-floor policy $\sum_j \lambda_j g_j \geq \rho\,\xi\, g_A$, the same reference combination violates the floor because $1.45 < 1.75$. Solving the alternative scale program gives $D_o^{\xi*} = 0$ with $\lambda_A = \xi = 1$. Therefore, fixed and proportional service floors define different feasible conditional technologies; the example does not establish a general local-elasticity theorem.

For the common-realization fuzzy model with effective spread $\varepsilon_f = 0.0125$, the two sign-dependent linear programs yield $D_o^{\xi*} = 0.2047582$ at $\xi^* = 1.7067901$ for agency A under the fixed service floor. At $\varepsilon_f = 0$, the optimal value returns to $1/6$. These calculations illustrate the nested possibility formulation without estimating uncertainty for any observed transit agency.

The example isolates the two operational constraints: the chosen reference mix exhausts the fixed fleet ceiling ($k = 2$), yet exceeds the fixed service minimum ($g = 1.45 \geq 1$); scaling that same minimum with ξ would instead require $g \geq 1.75$, which the mix violates. Consequently,

Corrected proofs

holding capacity constant does not make the service-floor policy immaterial. The different LP optima reflect distinct feasible comparison technologies, not different observed production data.

Synthetic known-technology experiment. For each curvature $p \in \{0.65, 1, 1.35\}$, 20 independent replicates of 16 artificial agencies were generated using a NumPy random generator with seed 20261026. In each replicate, x was formed by sorting the fixed endpoints 1 and 25 together with 14 independent Uniform$(1.01, 24.99)$ draws. Both outputs were set to $y^1 = x^p$ and $y^2 = 1.3\,x^p$, with $k_j = g_j = 1$ for every artificial agency; consequently, the capacity and service constraints were feasible at the self-comparison, and the output ranges were recalculated within each replicate. The same conditional crisp LP and baseline numerical classification rule were applied to all 320 agency-replicate observations at each p. At $p = 0.65$, 300 observations received DD and 20 ZI (the endpoint case); at $p = 1$, all 320 received ZI; at $p = 1.35$, 280 were inefficient, and 20 ID plus 20 ZI were classified among the frontier observations. These are transparent finite-sample outcomes under the stated generating process rather than proof that a global ξ direction identifies a derivative-based elasticity (Proposition 5).

4 Results

4.1 Descriptive Characteristics of the Sample

After the selection criteria were applied and four observations with zero Passenger Miles Traveled were removed, the analysis was conducted on 284 decision-making units. The ranges of the operating variables show substantial heterogeneity in the scale of activity. Total Operating Expenses ranged from $108,845 to $3,471,179,450, with a median of $15,492,838. Vehicles Operated in Maximum Service ranged from 2 to 3,264 vehicles. Table 2 reports the descriptive statistics of all variables used in the model.

Table 2: Descriptive statistics of the variables used in the model.

Variable	Role in the model	Mean	Std. dev.	Median	Minimum	Maximum
Total Operating Expenses	Discretionary input	86,339,287	271,384,760	15,492,838	108,845	3,471,179,450
VOMS	Quasi-fixed input	116.56	290.17	35.50	2	3,264
Vehicle Revenue Hours	Minimum service variable	412,789	1,044,227	109,915	1,001	11,831,472
Unlinked Passenger Trips	Output	10,721,005	44,859,753	1,780,830	2,242	647,859,171
Passenger Miles Traveled	Output	36,528,423	112,658,201	6,877,116	17,263	1,196,432,322

Corrected proofs

The gap between the mean and the median indicates that the distribution of operating scale, particularly for cost, UPT, and PMT, is skewed. Such dispersion is important for RTS analysis because the observed technology includes small, medium-sized, and very large units.

Across the retained agencies, VOMS and VRH have a Pearson correlation of 0.9909. Additional pairwise coefficients are given in Table 3. This strong dependence explains why the capacity ceiling and the service floor must be interpreted jointly; linear correlation does not establish that either individual inequality is globally redundant.

Table 3: Pearson correlations among the main production and constraint variables ($N = 284$).

Variable	Operating cost	VOMS	VRH	UPT	PMT
Operating cost	1.000	0.978	0.983	0.975	0.955
VOMS	0.978	1.000	0.991	0.933	0.976
VRH	0.983	0.991	1.000	0.942	0.971
UPT	0.975	0.933	0.942	1.000	0.910
PMT	0.955	0.976	0.971	0.910	1.000

4.2 Comparison of Alternative Model Specifications

Changing the structure of the technology substantially altered the number of units on the efficiency frontier. In the standard additive model, 8 of the 284 units, or 2.82% of the sample, were efficient. The global scale-program labels were one increasing-direction (ID), two zero-improvement (ZI), and five decreasing-direction (DD).

Introducing VOMS as a quasi-fixed input increased the number of efficient units to 16. This represented 5.63% of the sample, and the scale-direction composition consisted of 3 ID, 3 ZI, and 10 DD units.

Imposing the quasi-fixed-capacity restriction and the full minimum-service requirement simultaneously produced a different frontier. In the conditional deterministic model, 34 units, or 11.97% of the sample, were identified as efficient. Their scale-direction composition consisted of 14 ID, 16 ZI, and 4 DD units.

Under the baseline fuzzy-model specification with parameters

$$\rho = 1, \qquad \delta = 0.05, \qquad \eta = 0.75,$$

the number of efficient units decreased to 28. Of these, 15 were ID, 11 ZI, and 2 DD. Table 4 compares all four specifications and their associated ID/ZI/DD distributions.

In Tables 4, 8, 9, 11, A.1, A.2, and B.1, ID denotes increasing-direction, ZI denotes zero-improvement, and DD denotes decreasing-direction of the specified global LP diagnos-

Corrected proofs

Table 4: Frontier counts and model-specific scale-direction labels (ID/ZI/DD) for four DEA specifications; all models use the same 284 bus agencies.

Model specification	Efficient units	% of sample	ID	ZI	DD
Standard additive model	8	2.82	1	2	5
Additive model with VOMS quasi-fixed input	16	5.63	3	3	10
Conditional deterministic model with quasi-fixed input and minimum service	34	11.97	14	16	4
Proposed conditional fuzzy model	28	9.86	15	11	2

tic. These labels replace the conventional IRS/CRS/DRS terms in the numerical results and do not represent independently estimated local output elasticities.

The increase in efficient units from 8 to 16 after the introduction of VOMS indicates that some comparisons allowed by the standard model are no longer admissible once quasi-fixed capacity is taken into account. Adding the full-service constraint restricted the reference technology further and increased the number of frontier units to 34. Introducing fuzzy imprecision in the next stage reduced the frontier from 34 to 28 units.

4.3 Comparison of Deterministic and Fuzzy Scores

The efficiency scores under the deterministic and fuzzy technologies were highly similar. The mean efficiency score was 0.992987 in the deterministic model and 0.992256 in the fuzzy model. The medians were 0.996163 and 0.995653, respectively. The minimum score decreased from 0.910526 in the deterministic model to 0.902000 in the fuzzy model. Table 5 reports these distributions.

Table 5: Distribution of efficiency scores in the deterministic and fuzzy models.

Statistic	Deterministic model	Fuzzy model
Mean	0.9930	0.9923
Median	0.9962	0.9957
Standard deviation	0.0107	0.0118
Minimum	0.9105	0.9020
First quartile	0.9920	0.9913
Third quartile	0.9986	0.9982
Maximum	1.0000	1.0000

Corrected proofs

The Pearson correlation coefficient between the scores from the two models was $r = 0.994333$, and the Spearman rank correlation coefficient was $\rho_s = 0.966226$. The reported scores exhibit a high rank correlation under the two baseline specifications. Given their concentration near unity, this statistic does not demonstrate that individual frontier labels or policy conclusions are invariant.

Of the 284 units, 256 had a lower efficiency score in the fuzzy model than in the deterministic model. The score of 28 units was unchanged, and no unit obtained a higher score under the baseline fuzzy specification than under the deterministic model. The mean change in the score was -0.00073135, which corresponds to an average decline of approximately 0.073 percentage points. The largest decline in score was observed for the Southeastern Pennsylvania Transportation Authority ($0.951098 \rightarrow 0.935611$), for which the difference was -0.015487. Even in this case, the change in the efficiency score was limited to approximately 1.55 percentage points.

Normalization sensitivity. The 34 deterministic frontier agencies remain efficient under both interquartile-range (IQR)-weighted and evaluated-unit-relative slack objectives, but the ranking of inefficient agencies is not invariant. The Spearman rank correlation with the original range-weighted D measure is 0.9965 under IQR weighting and 0.1083 under evaluated-unit-relative weighting. The latter is not an $E = 1 - D$ score bounded to $[0, 1]$, so only its ranking and efficiency-zero status are compared. Translating the raw variables leaves the $\xi = 1$ additive efficiency program invariant under range normalization and VRS convexity; it does not preserve the ξ-ray scale diagnostic, for which the origin has economic meaning. As a distinct, non-equivalent geometry check, replacing the three positive discretionary variables by their natural logarithms yields 33 efficient agencies and a Spearman correlation of 0.2021 with the baseline deterministic D ordering. This log-coordinate model is not economically equivalent to proportional scaling in original units and is reported solely to reveal the sensitivity to skew compression. Table 6 summarizes the alternative slack-weighting results.

Table 6: Alternative slack-weighting schemes and resulting rank correlations.

Slack weights	Frontier units	Shared with baseline	Spearman ρ (D)
Sample range	34	34	1.0000
Interquartile range	34	34	0.9965
Evaluated-unit values	34	34	0.1083

4.4 Changes in Frontier Membership

A direct comparison of unit status in the deterministic and fuzzy models showed that 250 units remained inefficient in both models and 28 units lay on the frontier under both technologies. Six

Corrected proofs

units that were efficient in the deterministic model left the frontier after fuzzy imprecision was introduced. None of the units that were inefficient in the deterministic model became efficient in the fuzzy model. The status transitions are summarized in Table 7.

Table 7: Transition in efficiency status from the deterministic to the fuzzy model.

Status in deterministic model	Inefficient in fuzzy model	Efficient in fuzzy model
Inefficient	250	0
Efficient	6	28

Of the 34 units efficient in the deterministic model, 28 also remained efficient in the fuzzy model. The frontier-membership retention rate was

$$\frac{28}{34} \times 100 = 82.35\%.$$

The six units that left the frontier were City of Fairfax (CUE, the agency service name), Hampton Roads Transit, Miami-Dade Department of Transportation and Public Works, Greater Lafayette Public Transportation Corporation, University of Michigan Parking and Transportation Services, and UC Davis/Unitrans. The corresponding fuzzy scores and deterministic direction intervals are reported in Table B.1 (Appendix B).

The fuzzy efficiency scores of these six units showed different degrees of distance from the frontier. For example, the City of Fairfax score was 0.999476 and the University of Michigan score was 0.998457. Therefore, leaving the frontier did not necessarily indicate a substantial loss of efficiency; in some cases, the unit moved from exact efficiency ($E = 1$) to a point extremely close to the frontier. Table B.1 in Appendix B lists all six agencies.

4.5 Scale-Direction Composition on the Deterministic and Fuzzy Frontiers

Under the conditional deterministic technology, 14 of the 34 efficient units were classified as ID, representing 41.18% of the frontier units. ZI accounted for 16 units, or 47.06% of the frontier, while 4 units, or 11.76%, were classified as DD. On the fuzzy frontier, 15 of the 28 efficient units were ID,

$$\frac{15}{28} \times 100 = 53.57\%,$$

the ZI share was

$$\frac{11}{28} \times 100 = 39.29\%,$$

and the DD share was

$$\frac{2}{28} \times 100 = 7.14\%.$$

Corrected proofs

These percentages describe the distribution of returns to scale among frontier units and should not be generalized to all 284 units, because RTS in this study was determined only for DMUs that were efficient under the corresponding technology. The agency-level classifications of the 28 fuzzy-frontier units are listed in Tables A.1 and A.2 (Appendix A).

4.6 Stability of the Scale-Direction Classification

Of the 28 units that remained efficient in both the deterministic and fuzzy models, 26 retained their previous scale-direction classification. The classification agreement rate was

$$\frac{26}{28} \times 100 = 92.86\%.$$

The individual changes are identified in Table 8. Only two units changed classification: Central Florida Regional Transportation Authority (LYNX bus service) and Capital Area Transportation Authority (CATA bus service) were ZI in the deterministic model but moved to the ID region in the fuzzy model.

Table 8: Two agencies changing from zero-improvement (ZI) to increasing-direction (ID) between the conditional deterministic and baseline fuzzy frontiers; the last two columns give the fuzzy optimal-face ξ bounds.

Unit	Deterministic direction	Fuzzy direction	Fuzzy $\xi_{\min}$	Fuzzy $\xi_{\max}$
LYNX	ZI	ID	1.165709	1.165713
CATA	ZI	ID	1.566416	1.566428

For LYNX, the entire admissible ξ interval lay above 1.165, while for CATA the lower bound of the interval was approximately 1.566. Thus, the change of these two units to ID was not caused by an extremely small numerical fluctuation around one.

No direct transition from ID to DD or from DD to ID was observed. The two DD units on the baseline fuzzy frontier also lay clearly away from the threshold of one. For Blacksburg Transit, the ξ interval was approximately 0.918, while for the San Francisco Municipal Transportation Agency it was approximately 0.589.

Among the 17 baseline fuzzy-frontier agencies with strictly positive scale improvement, no optimal-face ξ interval crossed unity. The other 11 zero-improvement (ZI) agencies had $[\xi_{\min}, \xi_{\max}] = [1, 1]$ and were classified by the zero-improvement condition rather than by a one-sided interval test. The agency-level ξ intervals are listed in Tables A.1 and A.2 (Appendix A).

Optimal-face sensitivity. Re-solving both sign branches and the ξ-extremization programs with optimal-face objective margins of 10^{-10}, 10^{-8}, and 10^{-6} left the baseline fuzzy composition at 15 ID, 11 ZI, and 2 DD in all three cases. An objective of zero takes precedence over ξ-range tests when the model-based ZI label is assigned. This result is numerical robustness

Corrected proofs

for the selected rule; it is not a general theorem identifying derivative-based RTS under fixed service requirements.

Conditional projection extension. For the 250 agencies inefficient under the deterministic service-constrained model, one additive-optimal convex reference projection was constructed and submitted to the same crisp ξ scale-direction program, with the original 284-point technology and the evaluated agency's capacity and service bounds held fixed. The resulting projected-point labels were 67 ID, 42 ZI, and 141 DD. Because alternative optimal projections may exist, these are classifications of selected projections, not unique agency-level scale elasticities. This extension complements, rather than changes, the classifications of the 34 observed frontier agencies.

4.7 Effective-Spread Sensitivity

Table 9 reports all nine (δ, η) combinations for reproducibility but groups the rows into only five distinct optimization scenarios indexed by $\varepsilon_f = \delta(1-\eta)$: 0, 0.00625, 0.0125, 0.025, and 0.05. Repeated ε_f values yield identical LPs and are not independent sensitivity outcomes (Proposition 2). The NTD observations are actual reported transit records; the fuzzy spreads are specified hypothetical uncertainty scenarios.

Table 9: Nine nominal (δ, η) combinations representing five distinct effective-spread scenarios; identical ε_f values correspond to repeated linear programs.

δ	η	$\varepsilon_f = \delta(1-\eta)$	**Efficient units**	**% of sample**	**ID**	**ZI**	**DD**
0.025	0.50	0.01250	28	9.86	15	11	2
0.025	0.75	0.00625	31	10.92	15	12	4
0.025	1.00	0.00000	34	11.97	14	16	4
0.050	0.50	0.02500	22	7.75	10	10	2
0.050	0.75	0.01250	28	9.86	15	11	2
0.050	1.00	0.00000	34	11.97	14	16	4
0.100	0.50	0.05000	15	5.28	4	10	1
0.100	0.75	0.02500	22	7.75	10	10	2
0.100	1.00	0.00000	34	11.97	14	16	4

The repeated ε_f groups in Table 9 are 0 (three pairs), 0.0125 (two pairs), and 0.025 (two pairs). The remaining values, 0.00625 and 0.05, occur once each.

At $\eta = 1$, all three δ settings reproduce 34 efficient agencies, comprising 14 ID, 16 ZI, and 4 DD units. This equivalence follows from $\varepsilon_f = 0$, which yields the same deterministic program.

Corrected proofs

At $\eta = 0.50$, increasing δ from 0.025 to 0.05 and then to 0.10 yields 28, 22, and 15 efficient units, respectively. At $\eta = 0.75$, the corresponding counts are 31, 28, and 22. The results are indexed by ε_f, rather than by independently identifiable effects of δ and η. Since δ and η enter only through ε_f, interpreting them as separately identified effects is inappropriate. The corresponding curve in Figure 1 summarizes the five mathematically distinct scenarios.

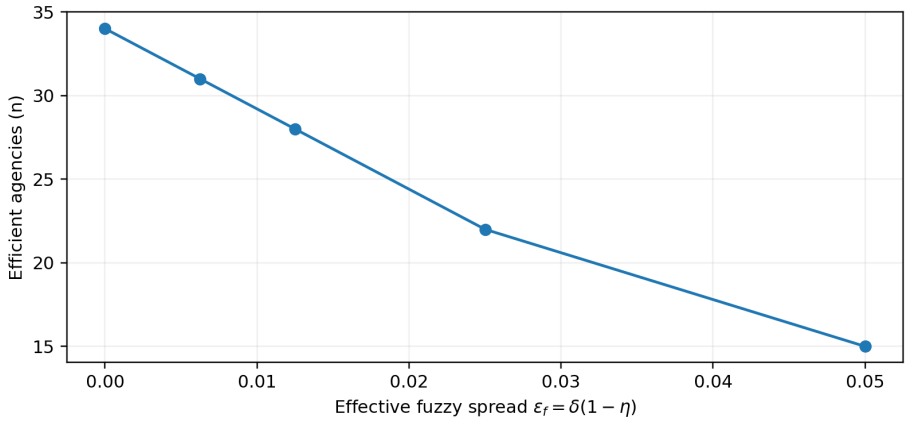


Figure 1: Number of efficient agencies by distinct effective-spread scenario.

At $\varepsilon_f = 0$, all combinations reduce to the deterministic frontier of 34 agencies. At positive ε_f, the changes in the objective and the frontier are induced by widening the same optimistic triangular η-cut, not by two independently identifiable fuzzy parameters.

Alternative uncertainty checks. A worst-case (for-all-realizations) endpoint counterpart of the shared-realization η-cut gives 42 efficient agencies at $\varepsilon_f = 0.0125$, compared with 28 under the optimistic possibility counterpart. In two heterogeneous effective-spread scenarios, assigning $\varepsilon_f = 0.025$ to smaller agencies and 0.00625 to larger ones yields 28 efficient agencies; reversing the assignment yields 27. These are controlled scenarios rather than estimates of agency-specific statistical confidence limits; the direction and size of an uncertainty effect depend on the decision rule and the distribution of imprecision.

4.8 Sensitivity to the Minimum Service-Level Constraint

At service-floor fractions $\rho = 0.90$ and 0.95, the deterministic model yields 16 efficient agencies with 3 ID, 3 ZI, and 10 DD labels. At $\rho = 1.00$, the corresponding values are 34 efficient agencies with 14 ID, 16 ZI, and 4 DD labels. The complete 21-point grid and six additional endpoint values are reported in Table 10. Figure 2 displays the 21-point grid; the six endpoint refinements appear in Table 10.

Corrected Proofs

The primary 21-point scan from $\rho = 0.80$ to 1.00 in steps of 0.01 identified 16 deterministic frontier agencies for $0.80 \leq \rho \leq 0.99$ and 34 at $\rho = 1$. The endpoint refinements in Table 10 distinguish exact optimum status from the conventional $\tau_{\text{eff}} = 10^{-7}$ numerical frontier classification. At $\rho = 0.99$ and 0.995, the smallest strictly positive D_o^* values among nonfrontier agencies were 2.20127×10^{-6} and 1.10064×10^{-6}, respectively; both exceed τ_{eff}. At $\rho = 0.9999$, 0.99999, and 0.999999, the tolerance-based counts become 17, 19, and 31 before reaching 34 at $\rho = 1$.

Endpoint geometry. At $\rho = 1$, 18 agencies are newly on the exact deterministic frontier relative to $\rho = 0.99$. For each such agency, any improving comparator feasible at $\rho = 0.99$ must have $g_{\text{ref}} < g_o$; otherwise, the same positive slack improvement would remain feasible at $\rho = 1$. For any $0.99 < \rho < 1$, take a strictly positive convex mixture of that comparator and the evaluated agency itself, with mixing weight $\theta \leq (1 - \rho)g_o/(g_o - g_{\text{ref}})$. The fleet ceiling is preserved by convexity, the service floor is satisfied, and the original positive slacks are multiplied by $\theta > 0$. Hence, those agencies remain strictly inefficient for every $\rho < 1$ in exact arithmetic, although their improvement approaches zero as ρ increases to one. Conversely, each agency efficient at $\rho = 0.80$ remains efficient for all larger ρ, because tightening the floor only shrinks the feasible reference set. Together with the observed equality of the 16-agency sets at $\rho = 0.80$ and 0.99, these facts account for the exact 16-to-34 endpoint transition; the earlier tolerance-based changes near one are classification-threshold effects, not a discontinuity in the optimal objective.

This change at $\rho = 1$ is related to the structure of the conditional production possibility set (Equation (10)). As long as the reference combination is allowed to reduce even a limited portion of the observed service level, a set of reference combinations remains feasible. Requiring complete reproduction of the observed service level excludes some of these points from the reference technology and creates a new frontier.

Across the full grid, the average score changes continuously (0.991958 at $\rho = 0.80$; 0.992848 at $\rho = 0.99$; 0.992987 at $\rho = 1$). Under the baseline deterministic optimal references, the capacity constraint is binding for 267 agencies, the service constraint for 164, and both for 150, using a 10^{-7} relative residual cutoff. These binding counts pertain to one selected optimal reference solution; alternative optima may yield different counts.

Service-floor scaling check. When the service constraint is changed from $\sum_j \lambda_j g_j \geq \rho\, g_o$ to $\sum_j \lambda_j g_j \geq \rho\, \xi\, g_o$ in the scale program, while the efficiency-frontier selection remains fixed at $\xi = 1$, the 28 baseline fuzzy-frontier agencies are classified as 3 ID, 5 ZI, and 20 DD, using the optimal scale-program objective and the sign of its ξ solution. The original fixed-floor composition is 15/11/2. These sharply different results establish that the scale-direction labels are conditional on how service obligations evolve with the scale parameter. Table 11 compares the two service-floor conventions.

Corrected proofs

Table 10: Deterministic service-floor sensitivity: 21-point primary grid and six additional endpoint levels ($N = 284$).

ρ	Efficient units	VRH binding	Mean score
0.80	16	37	0.991958
0.81	16	41	0.991975
0.82	16	52	0.991995
0.83	16	54	0.992017
0.84	16	60	0.992042
0.85	16	60	0.992068
0.86	16	69	0.992095
0.87	16	75	0.992124
0.88	16	84	0.992157
0.89	16	87	0.992193
0.90	16	90	0.992231
0.91	16	97	0.992272
0.92	16	99	0.992315
0.93	16	104	0.992361
0.94	16	109	0.992412
0.95	16	114	0.992475
0.96	16	120	0.992545
0.97	16	125	0.992625
0.98	16	135	0.992724
0.99	16	144	0.992848
0.995	16	147	0.992917
0.999	16	148	0.992973
0.9995	16	148	0.992980
0.9999	17	148	0.992986
0.99999	19	148	0.992987
0.999999	31	148	0.992987
1.00	34	164	0.992987

Table 11: Short-run scale classifications under fixed and proportionally scaled service floors.

Service floor in direction program	ID	ZI	DD
Fixed at ρg_o	15	11	2
Scaled to $\rho\xi g_o$	3	5	20

Corrected proofs

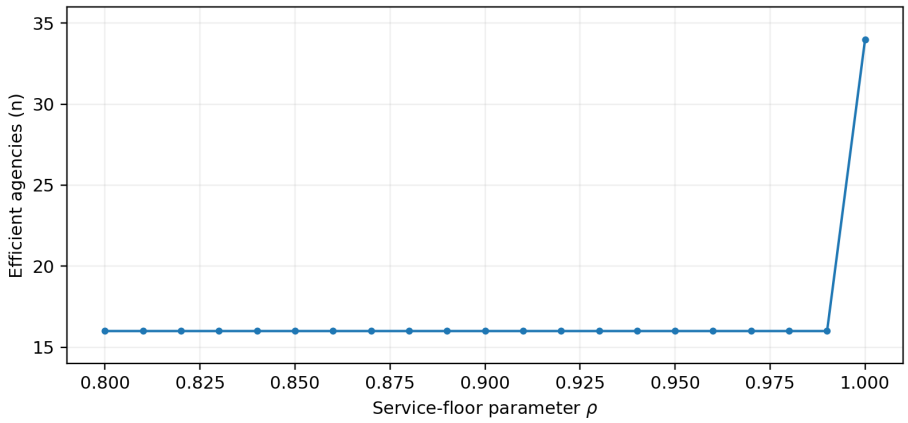


Figure 2: Conditional frontier membership across the 21-point grid ($\rho = 0.80$–1.00); six finer endpoint values are reported in Table 10.

4.9 Independent Check of the Unrestricted RTS Classification

As an independent check of the returns-to-scale classification, the eight efficient units in the standard additive model were also evaluated using a BCC supporting-hyperplane approach. The classifications from the ξ model and the BCC check were identical for all eight units. The final composition of this group consisted of one ID unit, two ZI units, and five DD units.

MTA New York City Transit (NYCT), New Jersey Transit Corporation (NJ TRANSIT), and Miami-Dade Department of Transportation and Public Works (DTPW) were classified as DD under both methods. Mass Transportation Authority (MTA; NTD ID 50032) and University of Iowa Campus Bus System (CAMBUS) were classified as ZI. Shreveport Area Transit System and Ames Transit Agency were also DD, while Paratransit, Inc. was the only ID unit in this group. Accordingly, the agreement rate between the two methods was

$$\frac{8}{8} \times 100 = 100\%.$$

4.10 Numerical Efficiency-Threshold Robustness

The efficiency-identification threshold was varied over the range

$$10^{-10} \leq \tau_{\text{eff}} \leq 10^{-6}.$$

Across all these values, the conditional deterministic model identified 34 efficient units and the baseline fuzzy model identified 28 efficient units. The number of frontier units did not change anywhere within this range.

Corrected Proofs

The smallest positive inefficiency among inefficient units in the deterministic model was approximately

$$5.95 \times 10^{-5},$$

and in the fuzzy model approximately

$$2.17 \times 10^{-4}.$$

This distance from the threshold of 10^{-7} indicates that distinguishing exactly frontier units from inefficient units did not depend on an extremely fine choice of the numerical cutoff.

4.11 Extreme-Observation Assessment

Paratransit, Inc. had a very large ξ value in the RTS model. Its ξ was approximately 504 in the deterministic model and approximately 517 under the baseline fuzzy specification (Table A.2). To evaluate the influence of this observation on the frontier, the analysis was repeated after removing it.

After Paratransit, Inc. was removed, the number of efficient units decreased from 34 to 33 in the deterministic model and from 28 to 27 in the fuzzy model. In both models, the reduction was attributable exactly to the removal of that same ID unit, and the scale-direction classification of all other frontier units common to the two analyses remained unchanged.

A broader extreme-observation screen used the natural logarithms of the following positive operating ratios: VRH per VOMS and PMT per VOMS. For each of these two log ratios, observations beyond $Q_1 - 3 \times \text{IQR}$ or $Q_3 + 3 \times \text{IQR}$ were flagged; the union comprised Western Maine Transportation Services, Clermont Transportation Connection, and Paratransit, Inc. As a sensitivity check rather than a revised sample definition, the model was re-estimated without the flagged observations.

After the simultaneous removal of these three units, the scale-direction classification of every unit common to both the original and adjusted samples remained unchanged. Accordingly, the extreme observations were not removed from the main analysis, and the 284-unit sample remained the basis for reporting the results.

Sample-stratified analysis. Splitting the agencies at the overall median VOMS (35.5 vehicles) yields 142 lower-capacity and 142 higher-capacity agencies. Re-estimating a separate reference frontier within each stratum identifies 19 versus 39 deterministic efficient agencies and 18 versus 34 baseline fuzzy efficient agencies, respectively. Because each stratum uses its own range denominators and reference technology, the within-stratum D values are not directly comparable with pooled scores. Paratransit, Inc. (NTD ID 90223) is confirmed as a Motor Bus, Directly Operated observation with cost $108,845, VOMS 12, VRH 1,001, UPT 2,242, and

Corrected proofs

PMT 17,263; its large ξ is an extrapolation diagnostic and not a recommended multiplicative expansion of operations. The fleet-stratified frontier counts are reported in Table 12.

Table 12: Separate conditional-frontier estimates by fleet-capacity stratum.

Sample	N	Deterministic frontier	Fuzzy frontier
Pooled	284	34	28
VOMS ≤ 35.5	142	19	18
VOMS > 35.5	142	39	34

Market-size sensitivity. The 284 agencies were additionally split at the median reported urbanized-area population (388,721.5), forming two groups of 142. When each group is evaluated against its own conditional reference technology, the lower-population stratum has 25 deterministic and 22 baseline fuzzy frontier agencies; the higher-population stratum has 31 and 23, respectively. These classifications are conditional on separate reference sets and range denominators. Urbanized-area population characterizes the service-market context, not the legal or organizational type of the operator. Table 13 shows the market-size stratification.

Table 13: Additional service-market stratification by reported urbanized-area population.

Urbanized-area population	Agencies	Deterministic frontier	Fuzzy frontier
At or below sample median	142	25	22
Above sample median	142	31	23

5 Discussion

5.1 Technology Structure and Efficiency Discrimination

Changing the reference technology, while holding the 284 observed agencies fixed, increased exact frontier membership from 8 in the unrestricted model to 16 under the fleet ceiling and 34 under the full-service floor (Table 4). These numbers measure changes in admissible comparator sets, not improvements in actual operator performance. The additional restrictions can reduce discrimination and alter which agencies admit strictly improving reference combinations.

The quasi-fixed VOMS restriction removes peer combinations that would require an infeasible fleet allocation over the short-run horizon. A unit that is inefficient under the unrestricted

Corrected proofs

comparison can therefore lie on the frontier of a smaller, conditionally admissible reference set. The service-floor restriction imposes a separate condition requiring an adequate level of VRH.

This result is consistent with the theoretical basis established by Banker and Morey [3], who distinguished between controllable factors and fixed or managerially uncontrollable factors. If a variable cannot be adjusted freely over the analytical horizon, including it in the same structure as the discretionary inputs can create a reference that is mathematically feasible but inconsistent with the unit's actual constraints. VOMS plays this role in the present study. Fleet capacity may change in the long run, but in the short run it is governed by capital decisions, vehicle acquisition and delivery, replacement planning, and maintenance constraints.

Noori et al. [18] address the identification of a single most efficient decision-making unit, not a complete ranking of agencies. Their procedure is therefore cited here as a distinct discrimination task rather than as a ranking method for every efficient agency.

The largest shift in the frontier occurred after the full service-level constraint was imposed. The number of efficient units in the deterministic model increased from 16 to 34. This change is consistent with the logic of the conditional technology. The constraint

$$\sum_j \lambda_j g_j \geq g_o,$$

prevents a reference combination that supplies less service than the observed level of the evaluated unit from appearing to deliver superior performance. Excluding such combinations reduces the number of available benchmarks, and units that were inefficient under the unconstrained technology may lie on the frontier under the conditional technology.

This finding is conceptually consistent with the argument of Ouellette, Quesnel, and Vigeant [20]. They showed that restrictions imposed on a technology can alter the structure of the RTS evaluation and that conventional RTS rules cannot be assumed to retain the same interpretation regardless of such constraints. An important distinction in the present study is that VRH is not used as a direct indicator of a specific legal regulation. It only defines the minimum service level within the production possibility set. Therefore, the increase in efficient units at $\rho = 1$ should be attributed to the restriction of the reference technology rather than to the causal effect of an observed regulation.

The increase from 16 to 34 frontier agencies at the full service requirement should not be attributed to a demonstrated operational benefit of the constraint. It is an expected implication of excluding reference combinations. If VOMS and VRH are strongly associated (Table 3), the constraint can also act as a utilization filter, particularly near $\rho = 1$. The observed jump is therefore descriptive rather than causal.

If a transit agency is compared with a unit that has more favorable input–output ratios but achieves those ratios while providing fewer service hours, the resulting benchmark is not necessarily appropriate for the managerial question of maintaining service. The conditional model restricts comparisons to units or combinations that satisfy the specified service condition.

Corrected proofs

The 21-point sensitivity scan and the scaled-floor counterfactual (Section 4.8) show that both frontier discrimination and the scale-direction diagnosis change with the definition of service obligations. In particular, the discontinuity in frontier membership at $\rho = 1$ occurs alongside smoothly varying mean scores (Table 10 and Figure 2). The model should therefore be used as a conditional benchmarking device, not as evidence that requiring full VRH improves actual productivity.

5.2 Role of Possibility-Based Fuzziness

Under the optimistic possibility formulation, increasing the effective spread relaxes the deterministic inequalities and cannot increase $E = 1 - D$ (Proposition 4). Thus, the observed nonpositive score changes and the shrinking exact frontier are consequences of the model's nesting property. They must not be presented as evidence that the chosen imprecision model represents actual sampling uncertainty.

The deterministic and baseline fuzzy scores were strongly correlated (Pearson $r = 0.9943$; Spearman $\rho_s = 0.9662$), with a mean score difference of approximately 0.000731 (Table 5). Because the scores were concentrated near unity (mean about 0.993), these coefficients do not establish uniform stability of individual rankings or classifications. Six of the 34 deterministic frontier agencies lost exact frontier status (Table 7), and two of the 28 shared frontier agencies changed their model-based scale-direction label (Table 8). The conclusions are confined to these observed comparisons and must be read together with the marked differences between the separately estimated stratum frontiers (Tables 12 and 13).

Agreement between the deterministic and fuzzy scores describes the selected comparison only; the narrow concentration of the range-adjusted efficiency scores around one limits the substantive content of a high Pearson correlation. Rank agreement is likewise sensitive to ties and to the chosen score definition. Heterogeneous agency-specific precision, interval or necessity treatments, and relative numerical tolerances may produce different classifications.

Under the common-realization optimistic construction, the fuzzy scenarios depend on δ and η only through $\varepsilon_f = \delta(1-\eta)$, as established in Proposition 2. The three $\eta = 1$ combinations are the identical $\varepsilon_f = 0$ deterministic program, rather than independent validation cases. For the five distinct positive and zero effective spreads, the reported frontier changes follow the nesting property in Proposition 4. Hence, neither δ nor η has a separately identified effect within this specific deterministic equivalent.

Increasing ε_f from 0 through 0.00625, 0.0125, 0.025, and 0.05 produces the reported frontier counts 34, 31, 28, 22, and 15 (Table 9 and Figure 1). This is a scenario comparison, not an empirical estimate of agency-specific measurement precision.

Corrected proofs

The shared-realization, independent-occurrence, and worst-case interval specifications answer different uncertainty questions. Their efficiency frontiers are not interchangeable: the independent-occurrence optimistic comparator yields zero exactly efficient agencies, whereas the shared-realization optimistic and worst-case versions yield 28 and 42, respectively, at $\varepsilon_f = 0.0125$. This demonstrates that the dependence convention and the orientation of the fuzzy inequalities must be reported as part of the estimand.

5.3 Stability and Alternative Optima

Among the 28 agencies that remained on both the deterministic and fuzzy frontiers, 26 retained their model-specific scale-direction label, whereas LYNX and CATA changed from ZI to ID (Section 4.6). This conditional agreement does not account for the six agencies leaving the frontier or establish label stability under different service floors or separate reference populations.

The ξ intervals of the two agencies were separated from one, rather than being indistinguishable from it at numerical precision: the lower bounds were 1.165709 for LYNX and 1.566416 for CATA. No direct transition between the ID and DD categories occurred among the 28 agencies common to the two baseline frontiers.

5.4 Interpretation of Conditional Scale Results

The 15 ID, 11 ZI, and 2 DD labels refer to the 28 observed agencies on the baseline fuzzy frontier, rather than to local economic returns to scale. The separate classification of the 250 inefficient deterministic agencies in Section 4.6 was obtained only after selecting one additive-optimal reference projection for each agency. Those projection-based labels describe selected points in the conditional technology and are not unique agency-level scale elasticities.

For a baseline fuzzy frontier agency assigned ID, the constrained scale LP has a strictly positive attainable slack improvement and an entire near-optimal ξ interval above one. This is a program-specific response to the proportional scaling of adjustable expenditure and passenger outputs while fleet capacity and the fixed service floor are unchanged. It does not imply that real production moves to a locally more productive region or justify fleet expansion.

The two DD units were also clearly separated from $\xi = 1$. The ξ value for Blacksburg Transit was approximately 0.918 and that for the San Francisco Municipal Transportation Agency approximately 0.589 (Tables A.1 and A.2). This distance reduces the likelihood that the DD classification is merely the result of numerical fluctuation around the ZI boundary.

Corrected proofs

Previous public-transport studies have also shown that scale status differs among operators. Odeck and Alkadi [19] reported increasing returns for a subset of units in the Norwegian bus industry, and Yu [25] showed that RTS status differs across bus-service structures. The present results likewise support the view that organizational size alone is insufficient to determine scale status and that the definition of technology and operational constraints must be incorporated into the assessment.

5.5 Quasi-Fixed Capacity

VOMS is not scaled by ξ in the present model. This choice directly affects the interpretation of returns to scale. If fleet size were also increased proportionally along with cost and outputs, the model would answer a different question concerning changes in the scale of all factors. Holding fleet capacity fixed limits the analysis to the question of where the unit lies, given its existing capacity, along the scale dimension of the adjustable component of the technology.

With VOMS held fixed, the diagnostic varies only the discretionary input–output ray inside the specified conditional reference set. Thus, an ID label identifies a positive scale-program optimum with ξ values above unity under this convention; it is not evidence that the underlying multioutput production function is locally in an increasing-returns region. Scaling fleet capacity or minimum service alongside ξ would define a different optimization experiment.

5.6 Comparing Conditional and Unconditional Reference Sets

The difference between 8 and 34 efficient units shows that the choice of technology constraints directly changes the scope of the reference set. This gap indicates that efficiency evaluation without accounting for capacity and service-level restrictions can generate a very different frontier. Neither frontier can be considered inherently “more correct” independently of the evaluation question; each represents a different technology.

The standard model answers how a unit performs relative to the observed set without additional capacity or service restrictions. The conditional model asks where the same unit stands when compared with combinations that respect its quasi-fixed-capacity and minimum-service constraints. The greater number of efficient units under the conditional technology is a natural consequence of a more restricted comparison set.

A low score under the unconditional technology may ignore some of the unit’s actual constraints. Conversely, an excessively restrictive technology can narrow the comparison set so much that many units lie on the frontier. Constraints should therefore be selected in accordance with the decision problem and the actual restrictions of production.

Corrected proofs

5.7 Interpretation of Validation Checks

Equality at $\eta = 1$ and agreement with the unrestricted BCC benchmark are preliminary implementation checks (Section 3.13 and Section 4.9). The independently assembled conditional and fuzzy matrices provide additional computational validation, while the synthetic experiment (Section 3.14) reveals the limits of the ξ convention under known curvature. The sign and size of a finite optimal ξ cannot in general be read as a one-sided production elasticity (Proposition 5). The original primal optimal-face interval should therefore be distinguished from classifications derived from strong complementary slackness and from alternative admissible projections.

The BCC supporting-hyperplane comparison concerns only the eight frontier agencies of the unrestricted additive model, for which the reported labels agreed. The independent assembly of the deterministic conditional and fuzzy coefficient matrices and the agreement of the HiGHS algorithms verify the numerical implementation of those finite LPs; they are not an independent economic validation of the conditional or fuzzy scale-direction categories.

For the 17 positive-improvement baseline fuzzy-frontier agencies, the ξ intervals lay strictly on one side of unity; the other 11 units had $[\xi_{\min}, \xi_{\max}] = [1, 1]$. Changing the objective-face margin from 10^{-10} to 10^{-6} did not change their model-based labels. This is a solver-robustness result conditional on the ξ criterion, not a proof of local RTS identification. A chosen additive projection of each of the 250 deterministically inefficient agencies was classified separately; alternative projections may lead to different local or global scale directions.

5.8 Numerical Tolerances and Outliers

Changing the efficiency-identification threshold from 10^{-10} to 10^{-6} did not change the number of efficient units in either the deterministic or the fuzzy model (Section 4.10). The distance between the smallest positive inefficiency value and the baseline threshold also showed that the separation of frontier units did not depend on a highly specific choice of τ_{eff}.

The very large ξ value for Paratransit, Inc. could raise the possibility that an extreme observation affected the structure of the frontier. Removing this unit only removed that same unit from the frontier set, while the scale-direction classifications of all other common units remained unchanged. The simultaneous removal of the three extreme observations identified from the logarithmic ratios produced the same conclusion (Section 4.11).

These checks show that the main scale-program values for the other units do not depend on the presence of a single very small or otherwise unusual observation. At the same time, Paratransit, Inc. was not removed from the main analysis because a large ξ value alone is not sufficient grounds for excluding a valid observation from the sample.

Corrected proofs

5.9 Transit-System Implications

The efficiency of a public-transport agency can be defined under different technologies, each answering a different question. If the objective is simply to compare input–output relationships without operational constraints, the standard technology is sufficient. If the decision maker wants to know where the unit stands given its existing fleet capacity and a requirement to preserve its service level, the conditional technology is more appropriate.

Within this framework, a managerial benchmark is more usable when achieving its performance does not require capacity greater than the quasi-fixed capacity of the evaluated unit and does not reduce service below the specified level. This restriction prevents a simple reduction in service from appearing as a route to improved efficiency.

For transport planners, the sensitivity findings imply that a reported efficient status should be accompanied by the chosen minimum-service floor, by whether the fleet is fixed over the planning horizon, and by the effective spread assigned to the passenger and expenditure measures. Alternative service obligations may lead to different program directions even for the same observed agencies; a change of label is a signal to revisit the benchmarking assumptions rather than an automatic recommendation to expand or contract service.

The fuzzy scenario analysis helps transit managers identify where exact frontier membership is sensitive to an explicit imprecision convention. Since six deterministic frontier agencies became inefficient and the within-stratum frontiers differ substantially, managers should not infer robust efficiency from an aggregate score correlation alone; benchmarking recommendations must specify the retained capacity, the service threshold, and the assumed measurement spread.

5.10 Incremental Methodological Contribution

The methodological contribution of the model lies in combining three features within a single conditional RTS framework. The quasi-fixed input prevents short-run capacity from being scaled like a discretionary input. The minimum service constraint excludes reference combinations that do not preserve the specified service level. The fuzzy representation relaxes the assumption that input and output values are known exactly.

The precise incremental contribution is a fully specified conditional LP diagnostic with common-realization sign-dependent fuzzy counterparts, a global optimal-face ξ interval, and an auditable sensitivity analysis with respect to the service-floor policy and the uncertainty convention. The capacity and service inequalities themselves are established nondiscretionary-variable devices, not new constraints. The assessment distinguishes finite-program scale directions from local economic RTS and makes the identification boundary explicit in Proposition 5.

Corrected proofs

The empirical contribution is an auditable comparison of the consequences of changing the specified technology and the uncertainty convention. The high score correlations and the 26/28 agreement among common frontier units are descriptive and do not establish the invariance of the full frontier, of the scale-direction labels, or of their economic interpretation.

6 Conclusion

This study develops a conditional range-adjusted additive DEA scale-direction diagnostic with quasi-fixed fleet capacity, minimum service provision, and a common-realization triangular fuzzy model. The feasible sign-branch formulation (Section 3.10) and the optimal-face interval (Section 3.7) define a reproducible global LP decision rule; they are not an estimator of local multioutput economic elasticity. In the 284-agency NTD sample, the successive reference technologies yield 8, 16, 34, and 28 frontier agencies (Table 4). The full-service condition reduces the admissible comparator set and thereby weakens discrimination, while ε_f governs the changes induced by the selected optimistic fuzzy convention. The 28 baseline fuzzy frontier agencies comprise 15 ID, 11 ZI, and 2 DD classifications. Changing the service floor from fixed to ξ-proportional changes the corresponding labels to 3, 5, and 20 (Table 11), confirming the dependence on the policy definition. The refined service-grid analysis (Section 4.8) shows that near $\rho = 1$, exact frontier membership and tolerance-based decisions must be distinguished.

Limitations and Future Work

The analysis uses a cross-sectional, heterogeneous transit sample and scenario-based rather than empirically calibrated fuzzy widths. Rankings depend on the normalization and the specified reference technology; projected-point labels depend on the chosen optimal projection. The ξ diagnostic is global and conditional, and its sign is not a local elasticity without further identifying assumptions. Future work should examine agency-specific uncertainty, alternative projection bounds, and local directional elasticity under explicitly binding service and fleet constraints, for example using bootstrap-based sensitivity analysis [23].

Corrected proofs

Appendix

A Scale-Direction Classification of the Fuzzy Frontier Agencies

Tables A.1 and A.2 report the scale-direction classification and the ξ interval for each of the 28 agencies that are efficient in the baseline fuzzy model.

Table A.1: Scale-direction classification and ξ interval for the 28 efficient units in the fuzzy model, part 1 of 2 (Det. = deterministic).

NTD ID	Decision-making unit	Det. direction	Fuzzy direction	Fuzzy ξ_{min}	Fuzzy ξ_{max}
20002	Capital District Transportation Authority (CDTA)	ID	ID	1.481201	1.481207
20008	MTA New York City Transit (NYCT)	ZI	ZI	1.000000	1.000000
20080	New Jersey Transit Corporation (NJ TRANSIT)	ZI	ZI	1.000000	1.000000
30035	Ohio Valley Regional Transportation Authority (OVRTA / EORTA)	ID	ID	1.284868	1.284914
30091	Town of Blacksburg (BT) – Transit	DD	DD	0.918091	0.918100
30096	Shore Transit	ID	ID	1.646360	1.646391
40002	City of Knoxville (KAT) – Mass Transit	ID	ID	1.376881	1.376889
40022	Metropolitan Atlanta Rapid Transit Authority (MARTA)	ZI	ZI	1.000000	1.000000
40032	County of Volusia	ID	ID	1.310446	1.310489
40035	Central Florida Regional Transportation Authority (LYNX)	ZI	ID	1.165709	1.165713
50032	Mass Transportation Authority (Flint, Michigan; NTD ID 50032)	ZI	ZI	1.000000	1.000000
50033	Interurban Transit Partnership (ITP)	ID	ID	1.578385	1.578405
50036	Capital Area Transportation Authority (CATA)	ZI	ID	1.566416	1.566428
50066	Chicago Transit Authority (CTA)	ID	ID	1.308350	1.308356

The final scale-direction composition of this set consists of 15 ID, 11 ZI, and 2 DD units. The two classification changes relative to the deterministic model involve LYNX and CATA, both of which moved from ZI to ID (Table 8).

Corrected proofs

Table A.2: Scale-direction classification and ξ interval for the 28 efficient units in the fuzzy model, part 2 of 2 (Det. = deterministic).

NTD ID	Decision-making unit	Det. direction	Fuzzy direction	Fuzzy ξ_{min}	Fuzzy ξ_{max}
50113	Pace, Suburban Bus Division	ID	ID	1.395197	1.395200
50145	City of Kokomo (COK) – Transit	ID	ID	1.617752	1.617890
50163	Licking County Public Transportation (LCT)	ID	ID	1.982106	1.982647
60011	VIA Metropolitan Transit	ZI	ZI	1.000000	1.000000
60024	Shreveport Area Transit System	ZI	ZI	1.000000	1.000000
60032	New Orleans Regional Transit Authority (NORTA)	ZI	ZI	1.000000	1.000000
60114	STAR Transit	ZI	ZI	1.000000	1.000000
60273	University of Texas Rio Grande Valley – Parking & Transportation Services	ID	ID	3.745172	3.745369
70019	University of Iowa CAMBUS	ZI	ZI	1.000000	1.000000
70041	Ames Transit Agency	ZI	ZI	1.000000	1.000000
70049	River Bend Transit (RBT)	ID	ID	15.034547	15.036789
90015	San Francisco Municipal Transportation Agency – Transit Division	DD	DD	0.589309	0.589310
90154	Los Angeles County Metropolitan Transportation Authority (LACMTA)	ZI	ZI	1.000000	1.000000
90223	Paratransit, Inc.	ID	ID	517.004925	517.007055

Corrected proofs

B Agencies Leaving the Baseline Fuzzy Frontier

Table B.1 lists the six agencies that were efficient in the deterministic model but left the frontier in the baseline fuzzy model (Section 4.4).

Table B.1: Six units efficient in the deterministic model that left the frontier in the baseline fuzzy model (Det. = deterministic).

NTD ID	Decision-making unit	Det. direction	Det. ξ_{min}	Det. ξ_{max}	Fuzzy efficiency
30058	City of Fairfax (CUE)	DD	0.786459	0.786475	0.999476
30083	Hampton Roads Transit (HRT)	ID	1.601584	1.601605	0.987418
40034	Miami-Dade DTPW	ZI	1.000000	1.000000	0.993230
50051	Greater Lafayette Public Transportation Corporation (GLPTC)	ZI	1.000000	1.000000	0.996685
50158	University of Michigan Parking and Transportation Services	ZI	1.000000	1.000000	0.998457
90142	UC Davis / Unitrans	DD	0.933348	0.933357	0.998069

Five of these six units had fuzzy efficiency scores above 0.99. Thus, leaving the frontier was not necessarily associated with a large distance from full efficiency; in several cases, the status change occurred very close to the frontier.

Declarations

Availability of Supporting Data

Data are from the Federal Transit Administration National Transit Database, TS2.1 Service Data and Operating Expenses Time Series by Mode (reporting year 2024, 2015–2024 time-series workbook). Original observations are publicly available at https://www.transit.dot.gov/ntd/ntd-data (website consulted October 26, 2026). The analytical selection was restricted to Active, Full Reporter, Urban, Operating, Motor Bus (MB), Directly Operated (DO) records; the supplied workbook yielded 288 eligible records and 284 after excluding four zero-PMT observations (Section 3.1). The complete analysis-ready data set—the 284 NTD agency identifiers, the analysis-ready input–output matrix, and the documented extraction and exclusion filters—is publicly archived in a persistent repository at https://doi.org/10.5281/zenodo.23987779 (Zenodo record 23987779).

Corrected proofs

The identification key is the NTD ID joined across the 2024 cost, VOMS, VRH, UPT, and PMT records. The complete 284-ID extract, the exact zero-PMT exclusions, and the scenario outputs are included in the computational supplement. The source-file version and download metadata are retained with the original supplied workbook.

Code Availability

The computational supplement contains the 284 NTD identifiers, the analysis-ready input–output matrix, the documented NTD extraction and sample filters, the Python model scripts, the sensitivity-grid outputs, the synthetic experiment settings, and the agency-level results. Replication uses Python/NumPy 2.3.5 and SciPy 1.17.0 with the bundled HiGHS 1.8.0 solver and 10^{-9} primal/dual feasibility tolerances (Section 3.12). The self-contained computational supplement includes the analysis-ready data, scripts, model settings, synthetic experiment, and numerical outputs, allowing the calculations to be reproduced from the archived computational supplement at https://doi.org/10.5281/zenodo.23987779 (Zenodo record 23987779). The archive includes the Python implementations of the deterministic and fuzzy additive DEA linear programs, the model specifications, the sensitivity-analysis outputs, the synthetic-experiment configurations, and all numerical results reported in this manuscript.

Funding

This research received no specific grant from any public, commercial, or nonprofit funding agency.

Conflict of Interest

The authors declare that they have no known financial or personal conflicts of interest that could have influenced this research.

Authors' Contributions

Marzieh Hossein Nakhai: Conceptualization; Methodology; Data curation; Software; Formal analysis; Investigation; Visualization; Writing – original draft preparation. **Mohsen Rostamy Malkhalifeh**: Conceptualization; Methodology; Validation; Supervision; Writing – review & editing. **Farhad Hosseinzadeh Lotfi**: Methodology; Validation; Supervision; Writing – review & editing. All authors reviewed and approved the final manuscript.

Artificial Intelligence Statement

OpenAI ChatGPT (GPT-6) assisted with English-language editing, consistency checks, and mathematical typesetting cross-checks. The authors reviewed the mathematical statements, numerical calculations, and resulting interpretations and retain responsibility for the complete scientific content and conclusions.

Corrected proofs

Publisher's Note

The publisher remains neutral regarding jurisdictional claims in published maps and institutional affiliations.

References

[1] Amirteimoori, A., Allahviranloo, T., Fakhr Mousavi, S. M. (2025). "Returns-to-scale and scale economies of two-stage production processes using a fully fuzzy range-adjusted measure model with strong complementary slackness conditions". *Expert Systems with Applications*, 262, 125606. https://doi.org/10.1016/j.eswa.2024.125606

[2] Banker, R. D., Charnes, A., Cooper, W. W. (1984). "Some models for estimating technical and scale inefficiencies in data envelopment analysis". *Management Science*, 30(9), 1078–1092. https://doi.org/10.1287/mnsc.30.9.1078

[3] Banker, R. D., Morey, R. C. (1986). "Efficiency analysis for exogenously fixed inputs and outputs". *Operations Research*, 34(4), 513–521. https://doi.org/10.1287/opre.34.4.513

[4] Banker, R. D., Thrall, R. M. (1992). "Estimation of returns to scale using data envelopment analysis". *European Journal of Operational Research*, 62(1), 74–84. https://doi.org/10.1016/0377-2217(92)90178-C

[5] Blavet, T., Ouellette, P., Vigeant, S. (2023). "α-returns to scale with quasi-fixed inputs: An application to Québec hospitals". *Applied Economics*, 55(58), 6922–6938. https://doi.org/10.1080/00036846.2023.2166664

[6] Charnes, A., Cooper, W. W., Rhodes, E. (1978). "Measuring the efficiency of decision making units". *European Journal of Operational Research*, 2(6), 429–444. https://doi.org/10.1016/0377-2217(78)90138-8

[7] Chu, X., Fielding, G. J., Lamar, B. W. (1992). "Measuring transit performance using data envelopment analysis". *Transportation Research Part A: Policy and Practice*, 26(3), 223–230. https://doi.org/10.1016/0965-8564(92)90033-4

[8] Cooper, W. W., Park, K. S., Pastor, J. T. (1999). "RAM: A range adjusted measure of inefficiency for use with additive models, and relations to other models and measures in DEA". *Journal of Productivity Analysis*, 11(1), 5–42. https://doi.org/10.1023/A:1007701304281

Corrected proofs

[9] Daraio, C., Diana, M., Di Costa, F., Leporelli, C., Matteucci, G., Nastasi, A. (2016). "Efficiency and effectiveness in the urban public transport sector: a critical review with directions for future research". *European Journal of Operational Research*, 248(1), 1–20. https://doi.org/10.1016/j.ejor.2015.05.059

[10] Dyson, R. G., Allen, R., Camanho, A. S., Podinovski, V. V., Sarrico, C. S., Shale, E. A. (2001). "Pitfalls and protocols in DEA". *European Journal of Operational Research*, 132(2), 245–259. https://doi.org/10.1016/S0377-2217(00)00149-1

[11] Emami, F., Nasirzade Tabrizi, T. (2016). "An additive model for estimation return to scale in regulated environment with quasi-fixed inputs in data envelopment analysis (DEA)". *International Journal of Data Envelopment Analysis*, 4(1), 893–901. https://sanad.iau.ir/ar/Article/826433

[12] Emami, F., Rostamy Malkhalifeh, M., Safdari, H. (2014). "Estimation of returns to scale in the presence of undesirable (bad) outputs in DEA when the firm is regulated". *International Journal of Data Envelopment Analysis*, 2(4), 487–495. http://ijdea.srbiau.ac.ir

[13] Federal Transit Administration (2024). "TS2.1: Service data and operating expenses time series by mode" [Data set]. National Transit Database, U.S. Department of Transportation. Retrieved November 5, 2026, from https://www.transit.dot.gov/ntd/data-product/ts21-service-data-and-operating-expenses-time-series-mode-2

[14] Hatami-Marbini, A., Emrouznejad, A., Tavana, M. (2011). "A taxonomy and review of the fuzzy data envelopment analysis literature: two decades in the making". *European Journal of Operational Research*, 214(3), 457–472. https://doi.org/10.1016/j.ejor.2011.02.001

[15] Huangfu, Q., Hall, J. A. J. (2018). "Parallelizing the dual revised simplex method". *Mathematical Programming Computation*, 10(1), 119–142. https://doi.org/10.1007/s12532-017-0130-5

[16] Khodabakhshi, M., Gholami, Y., Kheirollahi, H. (2010). "An additive model approach for estimating returns to scale in imprecise data envelopment analysis". *Applied Mathematical Modelling*, 34(5), 1247–1257. https://doi.org/10.1016/j.apm.2009.08.011

[17] Lertworasirikul, S., Fang, S. C., Joines, J. A., Nuttle, H. L. W. (2003). "Fuzzy data envelopment analysis (DEA): A possibility approach". *Fuzzy Sets and Systems*, 139(2), 379–394. https://doi.org/10.1016/S0165-0114(02)00484-0

Corrected Proofs

[18] Noori, Z., Zhiani Rezai, H., Davoodi, A., Kordrostami, S. (2022). "Finding the most efficient DMU in DEA: A model-free procedure". *Control and Optimization in Applied Mathematics*, 7(1), 15–29. https://doi.org/10.30473/coam.2022.62871.1193

[19] Odeck, J., Alkadi, A. (2001). "Evaluating efficiency in the Norwegian bus industry using data envelopment analysis". *Transportation*, 28(3), 211–232. https://doi.org/10.1023/A:1010333518966

[20] Ouellette, P., Quesnel, J. P., Vigeant, S. (2012). "Measuring returns to scale in DEA models when the firm is regulated". *European Journal of Operational Research*, 220(2), 571–576. https://doi.org/10.1016/j.ejor.2012.01.048

[21] Payan, A., Hajihosseini, A., Nosrat, A. (2022). "Detection of the type of left and right returns-to-scales using envelopment DEA models in crisp and fuzzy environments: an application for predicting changes in the stock market". *International Journal of Computational Intelligence Systems*, 15(1), Article 26. https://doi.org/10.1007/s44196-022-00079-4

[22] Seiford, L. M., Zhu, J. (1999). "An investigation of returns to scale in data envelopment analysis". *Omega*, 27(1), 1–11. https://doi.org/10.1016/S0305-0483(98)00025-5

[23] Simar, L., Wilson, P. W. (1998). "Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models". *Management Science*, 44(1), 49–61. https://doi.org/10.1287/mnsc.44.1.49

[24] Sueyoshi, T., Sekitani, K. (2005). "Returns to scale in dynamic DEA". *European Journal of Operational Research*, 161(2), 536–544. https://doi.org/10.1016/j.ejor.2003.08.055

[25] Yu, M. M. (2008). "Measuring the efficiency and return to scale status of multi-mode bus transit: evidence from Taiwan's bus system". *Applied Economics Letters*, 15(8), 647–653. https://doi.org/10.1080/13504850600721858

Authors' Bio-sketches

Marzieh Hossein Nakhai is a Ph.D. candidate in Mathematics at the Science and Research Branch, Islamic Azad University, Tehran, Iran. Her current research focuses on data envelopment analysis, fuzzy optimization, and mathematical modeling of efficiency and returns to scale.

Corrected proofs

Mohsen Rostamy Malkhalifeh is an Associate Professor of Applied Mathematics at the Science and Research Branch, Islamic Azad University, Tehran, Iran. His research interests include operations research, data envelopment analysis, mathematical optimization, and performance evaluation. Corresponding author. Email: Rostamy@srbiau.ac.ir

Farhad Hosseinzadeh Lotfi is a Full Professor of Mathematics at the Science and Research Branch, Islamic Azad University, Tehran, Iran. His principal research interests include operations research, data envelopment analysis, efficiency measurement, and decision-making under uncertainty.

Corrected proofs