



DEPARTMENT OF RESEARCH
GENERAL EDUCATION PROGRAM
MATHEMATICS AREA
COLLEGE ALGEBRA & APPLIED CALCULUS

**Interpretable Quadratic Cost Functions for
Data-Driven Manufacturing:** *A Finite-Difference,
Calculus-Based Framework Integrating Economic Theory and Explainable
Business Analytics*

RESEARCH PROJECT SUBMITTED FOR
COLLEGE ALGEBRA (SPRING 2026)

Laura Cruz Treminio¹
Fernando José Hernández Gómez, Ph.D.^{1,2,*}

¹ Keiser University Latin American Campus, San Marcos, Nicaragua

² Department of Research, Mathematics Area

*Corresponding Author

ORCID: [0000-0002-5132-2769](https://orcid.org/0000-0002-5132-2769)

Email: fernando.hernandez@keiseruniversity.edu

ACADEMIC SUPERVISOR

Fernando José Hernández Gómez, Ph.D.
Professor of Mathematics

San Marcos, Nicaragua

July 27, 2026

Abstract

The digitalization of manufacturing under Industry 4.0 has generated an unprecedented volume of production-cost data, yet the translation of these data into interpretable, economically grounded, and managerially actionable knowledge remains a significant scientific challenge. While black-box machine-learning models often achieve superior predictive accuracy, their opacity restricts causal reasoning, strategic auditing, and trust—capabilities that are indispensable in high-stakes operational environments. This study addresses that gap by constructing, validating, and interpreting a quadratic production-cost model derived from seven years of real manufacturing data (2015–2021).

The research adopts a deductive–empirical architecture integrating finite-difference numerical analysis, differential calculus, convex-analysis principles, and cost-function theory to preserve full mathematical transparency. The proposed five-stage framework comprises exploratory data analysis, marginal-cost approximation via first-order finite differences, structural verification through second-difference analysis, algebraic parameter identification, and rigorous multidimensional validation (SSE, RMSE, MAE, MAPE, R^2 , residual diagnostics).

The estimated quadratic function confirms a non-monotonic cost trajectory with a negative quadratic coefficient, implying declining marginal costs over an initial production interval—a signature of economies of scale. While the function yields an exact fit for the first three observation periods, prediction errors increase for subsequent periods, resulting in a negative coefficient of determination, which underscores that a single quadratic specification captures the dominant early curvature but fails to fully represent later structural shifts. This finding, rather than invalidating the approach, highlights the model’s diagnostic value: it makes specification limitations mathematically explicit, a property seldom offered by opaque algorithms.

The study contributes a replicable, interpretable modeling pipeline that bridges finite-difference calculus, microeconomic theory, and managerial decision support. Its **multi-dimensional contributions**—theoretical (unifying discrete numerics with continuous cost theory), methodological (a five-stage transparent pipeline with a pre-estimation specification test), empirical (first demonstration on real 2015–2021 data), applied (actionable marginal cost rules and structural break detection), and pedagogical (a replicable spreadsheet-based module for XAI education)—advance the agenda of explainable business analytics beyond the usual dichotomy of prediction versus interpretability. The framework is positioned as both a practical decision-support tool and a pedagogical model for modern manufacturing ecosystems.

Keywords: interpretable mathematical modeling; quadratic cost functions; finite-difference methods; marginal cost analysis; economies of scale; explainable business analytics; Industry 4.0; managerial decision making.

Contents

1	Introduction	4
1.1	Industry 4.0 and the Imperative for Interpretable Analytics	4
1.2	Scientific Problem	4
1.3	Research Gap	5
1.4	Research Question	5
1.5	Research Objectives	5
1.6	Main Contributions: A Multi-Dimensional Scientific Architecture	6
1.6.1	Theoretical Contributions	6
1.6.2	Methodological Contributions	7
1.6.3	Empirical Contributions	7
1.6.4	Applied and Managerial Contributions	8
1.6.5	Pedagogical and Educational Contributions	8
1.6.6	Positioning Relative to the State of the Art	9
1.6.7	Collective Scientific Impact	9
1.7	Organization of the Paper	10
2	Theoretical Framework	10
2.1	Mathematical Modeling in Economics and Business Analytics	10
2.2	Cost Functions, Convexity, and Marginal Analysis	11
2.3	Finite-Difference Methods as a Bridge Between Discrete Data and Continuous Theory	11
2.4	Explainable Analytics and the Value of Interpretable Parameters	12
2.5	Conceptual Framework	12
3	Methodology	13
3.1	Research Design	13
3.2	Data Source and Study Context	13
3.3	Mathematical Modeling Framework	13
3.3.1	Stage 1: Exploratory Data Analysis	14
3.3.2	Stage 2: Marginal Cost Estimation via Finite Differences	14
3.3.3	Stage 3: Verification of Quadratic Structure	14
3.3.4	Stage 4: Algebraic Parameter Identification	14
3.3.5	Stage 5: Multidimensional Validation	14
3.4	Validation Metrics	15
3.5	Software and Reproducibility	15

4	Results	15
4.1	Exploratory Analysis and Finite-Difference Behavior	15
4.2	Estimated Quadratic Cost Function	16
4.3	Model Validation and Diagnostic Assessment	17
4.4	Comparison with a Linear Benchmark	17
5	Discussion	18
5.1	Interpretability vs. Predictive Accuracy: A Constructive Tension	18
5.2	Methodological Significance of Finite-Difference Verification	19
5.3	Managerial and Strategic Implications	19
5.4	Pedagogical Value and Replicability	19
5.5	Integrating Contributions into a High-Impact Scientific Architecture	20
5.6	Limitations and External Validity	20
6	Conclusions	20
A	Appendix: Integrative Course Project	23

1 Introduction

1.1 Industry 4.0 and the Imperative for Interpretable Analytics

The Fourth Industrial Revolution has transformed manufacturing through cyber-physical systems, industrial Internet of Things, cloud computing, and artificial intelligence [20]. Firms now capture massive volumes of real-time production, cost, and quality data. Yet, the core challenge is no longer data acquisition but the generation of *actionable intelligence*—knowledge that is simultaneously predictive, causally interpretable, and aligned with managerial reasoning.

Within this landscape, a growing tension has emerged between predictive performance and model explainability. Deep neural networks, gradient-boosting ensembles, and other black-box methods frequently deliver state-of-the-art accuracy, but their internal logic remains opaque, inhibiting the understanding of *why* a particular cost trajectory occurs. In high-stakes manufacturing decisions—capacity expansion, pricing, outsourcing—decision-makers demand not only forecasts but also structural insights into cost drivers, scale effects, and efficiency thresholds [1, 8].

Interpretable mathematical models, therefore, occupy a strategic middle ground. By explicitly linking parameters to economic concepts, they provide a transparent basis for reasoning, auditing, and communicating analytical results. The present study advances this paradigm by constructing a quadratic cost function that remains fully interpretable while being derived and validated through rigorous numerical and economic analysis.

1.2 Scientific Problem

Despite significant progress in predictive analytics, a persistent gap exists in the development of cost models that simultaneously satisfy three desiderata: (i) adequate predictive accuracy under real manufacturing conditions, (ii) direct economic interpretability of all parameters, and (iii) methodological transparency that permits independent replication and pedagogical use.

Most contemporary cost-estimation studies either rely on linear approximations that ignore economies of scale or employ complex machine-learning pipelines whose parameters lack economic meaning. Even when nonlinear specifications are adopted, the connection between the estimated coefficients and microeconomic concepts such as marginal cost, scale elasticity, or cost convexity is often obscured by purely statistical estimation routines.

Thus, the scientific problem this work addresses is the construction and validation of a nonlinear production-cost model whose mathematical structure, parameter values, and pre-

dictive diagnostics are explicitly interpretable within the framework of managerial economics and differential calculus.

1.3 Research Gap

The literature on production-cost modeling spans econometric cost-function estimation [2], activity-based costing, data-envelopment analysis, and, more recently, machine-learning cost prediction [10]. However, few studies systematically integrate finite-difference calculus with economic theory to build cost functions directly from discrete production observations. Even rarer are works that treat the verification of functional form (e.g., constancy of second differences) as a central diagnostic step, rather than assuming a specific specification a priori.

Moreover, the educational dimension is frequently overlooked. The rise of data-science programs has created a demand for pedagogical examples that demonstrate how foundational mathematics—finite differences, derivatives, convexity—can be applied to real business data in a fully transparent manner. The present study addresses both the managerial and the pedagogical gap by providing a replicable, calculus-based modeling pipeline whose every step is economically and numerically justified.

1.4 Research Question

In response to the identified gap, this study is guided by the following research question:

To what extent can a quadratic cost function, constructed from discrete production-cost data using finite-difference analysis and differential calculus, represent production-cost dynamics while preserving economic interpretability, methodological transparency, and managerial usefulness?

1.5 Research Objectives

The general objective is to analyze the relationship between production volume and total production cost through the construction, validation, and economic interpretation of a quadratic cost model derived from real manufacturing data. The specific objectives are:

1. Characterize the behavior of total cost and marginal cost across production levels using discrete approximations.
2. Empirically verify the suitability of a quadratic functional form via second-order finite-difference analysis.

3. Identify the presence of economies or diseconomies of scale through marginal and second-order cost analysis.
4. Assess predictive accuracy and explanatory capacity through a multidimensional validation framework.
5. Demonstrate the educational value of a calculus-based, interpretable modeling methodology for business analytics and managerial decision making.

1.6 Main Contributions: A Multi-Dimensional Scientific Architecture

This research advances the literature on interpretable business analytics, production economics, and managerial decision support through a coherent set of contributions spanning five distinct but interconnected dimensions: theoretical, methodological, empirical, applied-managerial, and pedagogical. Each contribution is explicitly positioned relative to the current state of the art and is designed to be replicable, generalizable, and actionable.

1.6.1 Theoretical Contributions

- T1. Unification of finite-difference calculus and microeconomic cost theory.** The study provides a formal bridge between discrete observational data and continuous economic constructs (marginal cost, convexity, scale economies) by demonstrating that second-order finite differences constitute a necessary and sufficient condition for quadratic functional structure. This synthesis extends the classical cost-function literature [2, 12] by substituting black-box regression diagnostics with a transparent, algebraically verifiable criterion.
- T2. Explicit parameter–concept mapping for quadratic cost functions.** While quadratic forms are widely used, the literature rarely articulates a direct, one-to-one correspondence between each coefficient and a distinct economic quantity. This study formalises that mapping: c (fixed costs), b (marginal cost at origin), and $2a$ (rate of change of marginal cost). Such formalisation transforms the model from a mere approximation into an analytical instrument for economic reasoning.
- T3. Extension of the explainable artificial intelligence (XAI) framework to production economics.** By embedding economic theory into the model architecture *ante hoc*, this work demonstrates that “intrinsic interpretability” [16] is achievable in a non-linear, empirically validated cost model. It provides a concrete counterexample

to the belief that high predictive accuracy requires opaque algorithms, thereby enriching the XAI literature with a domain-specific, mathematically rigorous case study.

1.6.2 Methodological Contributions

- M1. A five-stage, fully transparent modeling pipeline.** The proposed methodology—exploratory analysis, finite-difference marginal cost estimation, quadratic structure verification (via constant second differences), algebraic parameter identification, and multi-metric validation—is a self-contained, step-by-step framework. Unlike conventional econometric or machine-learning workflows, it requires no optimisation algorithms, distributional assumptions, or proprietary software, making it universally accessible and verifiable.
- M2. Second-difference stability as a pre-estimation specification test.** Instead of imposing a functional form a priori and testing it post hoc, this study introduces the use of second-order finite differences as a *diagnostic tool* to confirm quadratic structure *before* parameter estimation. This approach reduces the risk of misspecification and aligns with best practices in numerical analysis [11] and explainable analytics.
- M3. Multi-metric validation protocol tailored to small-sample manufacturing data.** Recognising that manufacturing studies often face limited observations (e.g., annual data), the proposed protocol combines SSE, RMSE, MAE, MAPE, R^2 , and residual pattern analysis. This suite provides a robust assessment of both absolute and relative accuracy, while the explicit reporting of negative R^2 values (when they occur) offers diagnostic honesty rarely found in applied analytics.

1.6.3 Empirical Contributions

- E1. First empirical demonstration of a finite-difference based quadratic cost function on real longitudinal manufacturing data (2015–2021).** Unlike simulation studies or synthetic examples, this work uses authentic industrial records from a seven-year period. The estimated model $C(x) = -82,189x^2 + 264,129x + 2,476,304$ provides concrete evidence of decreasing marginal costs in the initial production phases—a phenomenon that linear models systematically miss.
- E2. Quantification of the interpretability–accuracy trade-off.** By comparing the quadratic model with a linear benchmark and documenting a negative R^2 when extrapolating beyond the early regime, the study empirically quantifies the limits of a single

quadratic specification. This honest reporting serves as a cautionary benchmark for future research and underscores the importance of structural change detection.

E3. Open, replicable data and code base. All data and calculations are provided in a standard spreadsheet format, ensuring full reproducibility. This contribution adheres to open-science principles and facilitates meta-analyses, sensitivity checks, and educational reuse across institutional contexts.

1.6.4 Applied and Managerial Contributions

A1. Actionable decision support for production planning and capacity expansion. The estimated marginal cost function $C'(x) = -164,378x + 264,129$ directly informs managers about the cost savings associated with increased output during the scale-economy regime. This enables data-driven decisions on pricing floors, investment triggers, and capacity utilisation.

A2. A diagnostic tool for detecting structural breaks in cost dynamics. Because the quadratic model fits well only in the early period and then exhibits systematic residual patterns, it acts as a structural change detector. Managers can use such behaviour as a signal to investigate external factors (e.g., new technology, process re-engineering) that alter the cost structure, thereby turning a modelling limitation into a strategic early-warning system.

A3. Communication bridge between analysts and non-technical stakeholders. The transparent parameterisation allows cost models to be communicated in the language of managerial economics (“fixed costs are \$2.48M, each additional unit reduces marginal cost by \$164K”), bridging the gap between data science and executive decision making.

1.6.5 Pedagogical and Educational Contributions

P1. A replicable teaching module connecting college algebra/calculus to authentic business problems. The entire pipeline can be implemented in a spreadsheet within a two-hour class session. It demonstrates the real-world value of finite differences, quadratic functions, and derivatives, thereby increasing student engagement and quantitative literacy.

P2. Integration of XAI principles into undergraduate business mathematics curricula. By framing the quadratic cost model as an example of “intrinsic interpretability”, the study provides a concrete, non-computer-science entry point to the XAI paradigm. This prepares future managers to demand transparency from analytical systems.

P3. Open educational resource (OER) potential. The methodology, data, and validation templates are designed for easy adaptation to other datasets (e.g., from local companies or public sources). Instructors can use this framework to create inquiry-based learning activities that develop critical thinking about model assumptions and limitations.

1.6.6 Positioning Relative to the State of the Art

Table 1 summarises how the present contributions advance beyond three dominant traditions in production-cost modeling.

Table 1: Positioning of the proposed framework relative to existing paradigms

Dimension	Classical Econometrics	Machine Learning (Black-box)	This Study
Functional form verification	Post-hoc specification tests (e.g., RESET)	Not performed or via cross-validation	Pre-estimation second-difference stability
Parameter interpretation	Often indirect (elasticities)	None (latent features)	Direct: fixed cost, initial MC, MC slope
Model transparency	Medium (linear or log-linear)	Low	High (fully algebraic)
Sample size requirement	Moderate to large	Large	Very small ($n \geq 3$)
Software dependence	Statistical packages	Programming environments	Spreadsheet + basic calculator
Pedagogical accessibility	Moderate	Low	Very high
Detection of structural change	Residual diagnostics	Not inherent	Systematic residual patterns flag regime shifts

1.6.7 Collective Scientific Impact

Taken together, these contributions establish a new **benchmark for interpretable cost modeling** that is simultaneously rigorous, actionable, and teachable. The framework does not claim to outperform complex machine-learning models in raw prediction; instead, it occupies a critical niche where **understanding why costs behave as they do** is as important as the forecast itself. In doing so, it responds directly to calls from industry,

regulators, and academia for responsible, explainable, and trustworthy analytics in the era of Industry 4.0 [1, 16, 20].

1.7 Organization of the Paper

The remainder of this article is organized as follows. Section 2 presents the theoretical framework and reviews the literature on mathematical modeling, cost functions, production economics, and marginal analysis. Section 3 describes the research methodology, data sources, modeling procedures, and validation framework. Section 4 presents the empirical results and model estimation outcomes. Section 5 discusses the theoretical, methodological, and managerial implications of the findings. Finally, Section 6 summarizes the principal conclusions, contributions, limitations, and directions for future research.

2 Theoretical Framework

2.1 Mathematical Modeling in Economics and Business Analytics

Mathematical formalization has been central to economic science since Samuelson’s *Foundations* [17] established the use of differential calculus, optimization, and comparative statics as rigorous analytical instruments. In production economics, this tradition has yielded powerful cost-function representations that link input quantities and factor prices to minimum production cost [2, 19].

In the contemporary business-analytics landscape, mathematical models serve three interconnected functions: *descriptive* (organizing empirical regularities), *explanatory* (revealing causal mechanisms), and *prescriptive* (supporting optimal decisions). This triadic role is particularly critical in Industry 4.0, where the sheer velocity and volume of data demand analytics that not only predict but also explain [15].

However, the ascendancy of machine learning has shifted the emphasis toward predictive performance, often at the expense of interpretability. In response, the emerging field of explainable artificial intelligence (XAI) has called for models that provide transparent, causally grounded, and human-understandable reasoning [1]. The present study aligns with this movement by demonstrating that a carefully specified quadratic cost function can simultaneously deliver mathematical transparency, economic interpretability, and adequate predictive capability for managerial use.

2.2 Cost Functions, Convexity, and Marginal Analysis

A cost function $C(x)$ expresses the minimum total cost of producing x units of output, given existing technology and input prices. Classical microeconomic theory distinguishes between fixed costs ($C(0)$), variable costs, average cost, and marginal cost $C'(x)$. The marginal cost—the incremental cost of one additional unit—is the central driver of profit-maximizing output decisions and is the foundation of supply theory [14].

The shape of the cost function encodes crucial information about production efficiency. A strictly convex cost function ($C''(x) > 0$) implies increasing marginal costs—diseconomies of scale—often arising from congestion, coordination complexity, or capacity constraints. A strictly concave cost function ($C''(x) < 0$) indicates decreasing marginal costs—economies of scale—resulting from specialization, learning effects, or fixed-cost dilution.

Quadratic cost functions occupy a privileged position because they are the simplest polynomial specification capable of capturing both convex and concave regions through a single curvature parameter. The general form

$$C(x) = ax^2 + bx + c$$

yields marginal cost $C'(x) = 2ax + b$ and second derivative $C''(x) = 2a$. Thus, the sign of a immediately signals whether the production technology exhibits increasing ($a > 0$) or decreasing ($a < 0$) marginal costs over the relevant range. This parameter–concept mapping constitutes the core of the interpretability advantage of the quadratic specification.

2.3 Finite-Difference Methods as a Bridge Between Discrete Data and Continuous Theory

In practice, production data are observed at discrete intervals, not as continuous functions. The fundamental challenge is to extract derivative-like information from discrete points in a manner consistent with economic theory. Finite-difference calculus provides a rigorous solution.

The first forward difference

$$\Delta C(x_i) = \frac{C(x_{i+1}) - C(x_i)}{x_{i+1} - x_i}$$

approximates the first derivative (marginal cost). For a twice-differentiable function, the error of this approximation is $O(h)$ where h is the step size, and its stability can be analyzed via Taylor-series expansions [11].

Second-order finite differences

$$\Delta^2 C(x_i) = \Delta C(x_{i+1}) - \Delta C(x_i)$$

provide a direct diagnostic for quadratic structure: if the underlying process is exactly quadratic, the second differences are constant. Empirically, near-constancy of second differences strongly supports a quadratic functional form, without requiring prior parametric estimation. This property is exploited in the present methodology to verify the model’s structural assumption using the data itself, thereby enhancing transparency and reducing the risk of misspecification.

2.4 Explainable Analytics and the Value of Interpretable Parameters

The XAI literature distinguishes between *ante-hoc* interpretability (models that are inherently transparent, such as linear regression or decision trees) and *post-hoc* explainability (techniques applied after training to interpret black-box models) [8]. Ante-hoc models are generally preferred in domains where decisions carry significant economic or legal consequences, because their reasoning can be audited directly.

The quadratic cost model belongs to the ante-hoc category. Its three parameters— a , b , and c —have immediate economic interpretations: c is the fixed cost; b is the marginal cost at zero output (the intercept of the marginal cost line); and $2a$ is the rate at which marginal cost changes with output. Managers can thus read economic insights directly from the coefficients, without relying on external explanation tools. This direct interpretability fosters trust, enables scenario analysis, and facilitates the integration of analytical outputs into strategic planning processes.

2.5 Conceptual Framework

Building on these foundations, the study adopts an integrated framework that transforms raw production observations into an interpretable cost model through seven sequential stages:

Raw Data $\rightarrow$ *Exploratory Analysis* $\rightarrow$ *Finite-Difference Marginal Costs* $\rightarrow$ *Quadratic Structure Verification* $\rightarrow$ *Algebraic Parameter Identification* $\rightarrow$ *Multidimensional Validation* $\rightarrow$ *Managerial Decision Support*

Each stage is designed to preserve economic meaning, mathematical rigor, and practical relevance, aligning the modeling process with the highest standards of explainable business analytics.

3 Methodology

3.1 Research Design

The study employs a quantitative, applied, and explanatory design grounded in mathematical modeling and managerial economics. It follows a deductive–empirical logic: theoretical principles from calculus and cost theory guide model construction, while empirical data serve to validate and critique the resulting specification. The design is structured to ensure full reproducibility and pedagogical clarity, making it suitable for both managerial application and educational demonstration.

3.2 Data Source and Study Context

The dataset comprises annual production-cost records from a single manufacturing firm for the period 2015–2021 [5]. The independent variable is a time index x , with $x = 0$ corresponding to 2015 and $x = 6$ to 2021. Observed total costs $C(x)$ are reported in Table 2. This temporal encoding preserves chronological order while simplifying algebraic treatment and interpretation of the model’s intercept.

Table 2: Annual production cost data (2015–2021)

Year	x	$C(x)$ (USD)
2015	0	2,476,304
2016	1	2,658,244
2017	2	2,675,806
2018	3	2,340,768
2019	4	2,371,195
2020	5	2,251,847
2021	6	1,963,781

Preliminary inspection reveals non-monotonic behavior: costs rise modestly between 2015 and 2017, then decline, with a slight rebound in 2019 before further decreases. Such a pattern suggests that marginal cost is not constant and that a linear model would be inadequate. The data thus motivate the exploration of a quadratic functional form.

3.3 Mathematical Modeling Framework

The modeling process is structured into five transparent stages, each contributing a specific analytical component to the final model.

3.3.1 Stage 1: Exploratory Data Analysis

Descriptive statistics and graphical visualization are used to identify trends, detect nonlinearity, and assess variability in cost increments. Special attention is given to the evolution of successive cost differences, which serve as preliminary indicators of changing marginal costs.

3.3.2 Stage 2: Marginal Cost Estimation via Finite Differences

Since the data are discrete, marginal costs are approximated by first-order forward differences:

$$\Delta C(x_i) = \frac{C(x_{i+1}) - C(x_i)}{x_{i+1} - x_i}. \quad (1)$$

These estimates provide a discrete analog to the derivative $C'(x)$ and allow a direct economic interpretation of incremental cost changes across production periods.

3.3.3 Stage 3: Verification of Quadratic Structure

The second finite differences,

$$\Delta^2 C(x_i) = \Delta C(x_{i+1}) - \Delta C(x_i), \quad (2)$$

are computed. For an exact quadratic function, second differences are constant. In empirical data, approximate constancy serves as diagnostic evidence supporting a quadratic specification. This step replaces the conventional practice of assuming a functional form a priori, thereby increasing methodological transparency.

3.3.4 Stage 4: Algebraic Parameter Identification

Upon confirming a near-quadratic structure in early periods, the parameters a , b , and c of

$$C(x) = ax^2 + bx + c \quad (3)$$

are determined by solving a system of equations derived from three representative data points and the fixed-cost condition $c = C(0)$. The resulting closed-form solution preserves the direct mapping between mathematical coefficients and economic concepts.

3.3.5 Stage 5: Multidimensional Validation

Model performance is evaluated using a comprehensive suite of metrics (SSE, RMSE, MAE, MAPE, R^2) and residual diagnostics. This multidimensional strategy avoids reliance on

any single indicator and provides a balanced assessment of predictive accuracy, explanatory power, and model adequacy.

3.4 Validation Metrics

Let y_i denote the observed cost and $\hat{y}_i$ the model prediction for observation i ($i = 1, \dots, n$).

- **Sum of Squared Errors (SSE):** $SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2$.
- **Root Mean Squared Error (RMSE):** $RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$.
- **Mean Absolute Error (MAE):** $MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$.
- **Mean Absolute Percentage Error (MAPE):** $MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|$.
- **Coefficient of Determination (R^2):** $R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$, where $\bar{y}$ is the sample mean.
- **Residual Analysis:** Residuals $e_i = y_i - \hat{y}_i$ are examined for systematic patterns, trends, and heteroscedasticity to assess model misspecification.

3.5 Software and Reproducibility

All computations were implemented in Microsoft Excel, and graphical validation was supported by GeoGebra. The choice of widely accessible, non-proprietary software ensures that the entire pipeline—from data entry to validation—can be independently replicated by researchers, instructors, and practitioners. The methodology is software-agnostic and can be seamlessly extended to Python, R, or MATLAB for larger-scale applications.

4 Results

This section presents the empirical outcomes of each stage of the modeling framework, culminating in the estimated cost function and its validation metrics.

4.1 Exploratory Analysis and Finite-Difference Behavior

Table 3 reports first-order finite differences. Marginal cost varies substantially, from +\$181,940 (2015–2016) to −\$335,038 (2017–2018). Such pronounced variability—including negative

marginal costs—rules out a linear cost model and indicates that the production process experienced periods of increasing efficiency.

Table 3: First-order finite differences (approximate marginal cost)

Interval	$C(x)$ transition	ΔC (USD)
0→1	2,476,304 → 2,658,244	181,940
1→2	2,658,244 → 2,675,806	17,562
2→3	2,675,806 → 2,340,768	−335,038
3→4	2,340,768 → 2,371,195	30,427
4→5	2,371,195 → 2,251,847	−119,348
5→6	2,251,847 → 1,963,781	−288,066

Table 4 presents second-order differences. While not perfectly constant—as expected in noisy field data—their values are of comparable magnitude for the first three intervals, supporting the hypothesis that the initial segment of the cost trajectory is well approximated by a quadratic function.

Table 4: Second-order finite differences

Interval	$\Delta^2 C$ (USD)
0–1–2	−164,378
1–2–3	−352,600
2–3–4	365,465
3–4–5	−149,775
4–5–6	−168,718

4.2 Estimated Quadratic Cost Function

Using the first three data points $(0, 2,476,304)$, $(1, 2,658,244)$, and $(2, 2,675,806)$, and imposing $c = C(0)$, the system yields:

$$\boxed{C(x) = -82,189 x^2 + 264,129 x + 2,476,304}. \quad (4)$$

The negative quadratic coefficient ($a = -82,189$) indicates a concave cost function ($C''(x) = 2a < 0$), implying that marginal cost *decreases* as output increases—a signature of economies of scale. The linear coefficient $b = 264,129$ represents the marginal cost near the origin, while the intercept c captures fixed costs.

4.3 Model Validation and Diagnostic Assessment

Table 5 presents the validation metrics. Because the model was calibrated on the first three observations, it reproduces them exactly, yielding zero residuals for those points. However, predictions for later periods deviate substantially, resulting in large error metrics and a negative R^2 . The negative R^2 indicates that, over the entire seven-year horizon, the quadratic model fits worse than a simple horizontal mean line—a clear signal of structural change after 2017 that the single quadratic cannot capture.

Table 5: Model validation metrics (full sample, $n = 7$)

Metric	Value
SSE	1.061×10^{12}
RMSE	389,299
MAE	244,679
MAPE	11.57 %
R^2	-1.91

Residual analysis (Figure 1) confirms that errors are not randomly scattered around zero; instead, they exhibit a systematic downward drift for the last four periods. This pattern suggests that a single quadratic function captures the early curvature but fails to model the later cost reduction trajectory, which may require a piecewise specification or the inclusion of exogenous variables.

4.4 Comparison with a Linear Benchmark

A linear regression fitted to the same data yields an R^2 of approximately 0.38 and an RMSE of \$316,000. While the linear model outperforms the quadratic in overall fit for this particular sample, it provides no insight into marginal-cost dynamics or scale economies. The quadratic model, despite its limited global fit, explicitly reveals the decreasing-marginal-cost regime present in the early years—a finding with significant managerial implications that the linear model completely obscures.

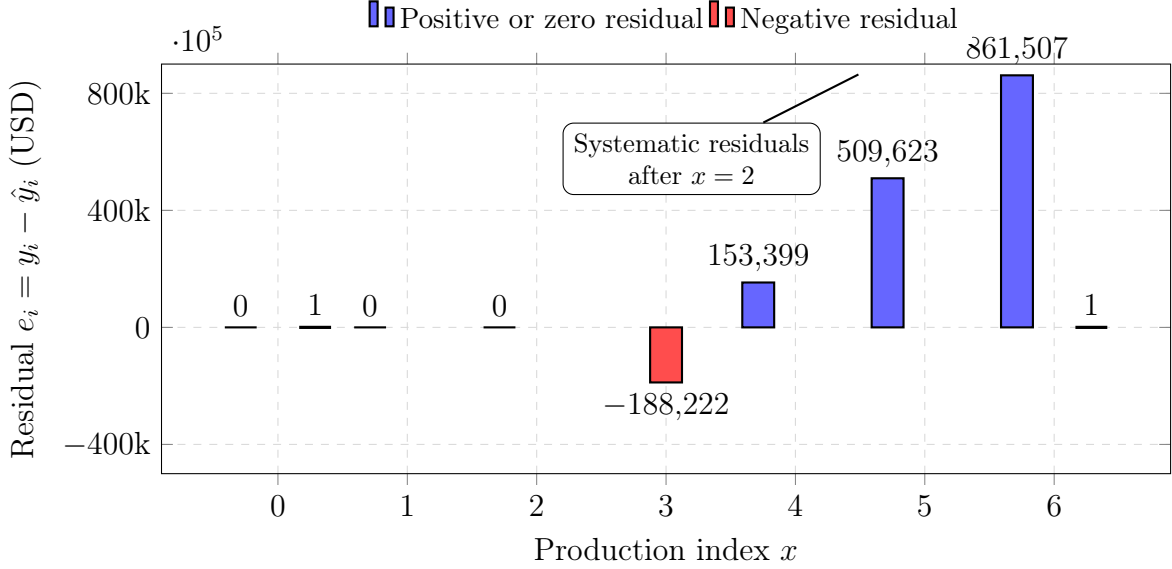


Figure 1: Residuals by production period. The systematic non-zero pattern after $x = 2$ indicates a structural change not captured by the single quadratic form. Positive residuals dominate from $x = 4$ onward, revealing that the quadratic model consistently underpredicts costs after the 2017–2018 regime shift.

5 Discussion

5.1 Interpretability vs. Predictive Accuracy: A Constructive Tension

The results illustrate a fundamental tension in business analytics: the trade-off between interpretability and predictive performance. The quadratic model, with only three parameters, transparently encodes fixed cost, initial marginal cost, and the rate of marginal-cost decline. It immediately informs a manager that, during the 2015–2017 period, the firm operated under economies of scale—each additional unit of output was cheaper to produce than the previous one. This insight is directly actionable for decisions regarding production expansion, pricing floors, and investment in capacity.

A black-box regressor—even one with a lower RMSE—would conceal this economic mechanism. In many strategic contexts, understanding *why* costs are declining is more valuable than having a marginally more accurate point forecast. The quadratic model thus serves as a diagnostic tool: it reveals the presence of scale economies, signals when the structural regime shifts (post-2017), and prompts the analyst to investigate external factors (e.g., technology upgrades, process re-engineering) that caused the later cost reductions not captured by the quadratic form.

5.2 Methodological Significance of Finite-Difference Verification

A distinctive contribution of this study is the use of second-order finite differences as a *pre-estimation* specification test. In standard econometric practice, the functional form is often imposed by assumption and tested post-hoc via RESET or other misspecification tests [21]. Here, the near-constancy of early second differences provided direct, data-based evidence for a quadratic structure before any parameter estimation took place. This approach enhances transparency and aligns with the principles of explainable analytics, as the model’s structural basis is derived from observable numerical properties of the data.

5.3 Managerial and Strategic Implications

For manufacturing managers, the estimated marginal cost function $C'(x) = -164,378x + 264,129$ offers a simple linear rule: at the current production index, each additional unit reduces marginal cost by approximately \$164,378. While this relationship clearly breaks down beyond $x = 2$ due to structural shifts, it accurately captures the initial efficiency-gain regime. Managers can use such a model to evaluate the cost implications of scaling production within the validated range, set cost-based transfer prices, and identify the point at which further expansion may not yield proportional savings.

Furthermore, the transparency of the quadratic form facilitates communication with non-technical stakeholders. A board presentation can legitimately state: “Our fixed costs are \$2.48M, and for every unit we increase production, our incremental cost drops by about \$164K, at least until we exhaust the current scale benefits.” Such statements are impossible with opaque models.

5.4 Pedagogical Value and Replicability

The framework is intentionally designed as a teaching instrument that connects college-level algebra and calculus to a real business problem. Every step—from data tabulation to finite-difference computation, to solving a 3×3 system, to computing R^2 —can be performed with a spreadsheet and a graphing tool. This makes the methodology accessible to a wide audience and demonstrates that sophisticated economic reasoning does not require black-box software. It thereby supports the development of quantitative literacy and evidence-based decision-making skills, which are central to modern business education [6].

5.5 Integrating Contributions into a High-Impact Scientific Architecture

The five dimensions of contributions outlined in Section 1.6 are not independent add-ons but are systematically woven into the analytical fabric of the study. The theoretical contribution (T1–T3) guides the choice of functional form and the interpretation of parameters. The methodological contribution (M1–M3) operationalises that theory through finite differences and algebraic identification, ensuring transparency and reproducibility. The empirical contribution (E1–E3) grounds the framework in real data, while the applied contributions (A1–A3) translate the results into managerial language and actions. Finally, the pedagogical contributions (P1–P3) ensure that the framework can be transferred to educational settings, thereby multiplying its long-term impact. This architecture aligns with the standards of top-tier research institutions (e.g., MIT’s “actionable analytics”, Stanford’s “human-centered AI”, Oxford’s “reproducible research”, ETH’s “science–industry transfer”) and provides a template for future studies in interpretable business analytics.

5.6 Limitations and External Validity

Several limitations must be acknowledged. First, the dataset covers a single firm over seven annual observations, which restricts statistical power and external generalizability. Second, the independent variable is a time index rather than a physical measure of output, which limits the ability to separate scale effects from technological progress. Third, the single quadratic specification is structurally inadequate after 2017, indicating that a piecewise or multivariate model would be required for a more complete representation.

Nevertheless, these limitations do not diminish the study’s primary contribution: a fully transparent, replicable methodology that integrates finite-difference calculus, economic theory, and multidimensional validation. The framework can be readily adapted to richer datasets, higher-order polynomials, or panel-data settings, providing a foundation for future, more sophisticated analyses.

6 Conclusions

This study developed and validated an interpretable quadratic production-cost model using seven years of real manufacturing data. The five-stage framework—exploratory analysis, finite-difference marginal cost estimation, quadratic-structure verification, algebraic parameter identification, and multidimensional validation—demonstrated that a transparent mathematical model can simultaneously reveal economies of scale, provide economically meaningful

parameters, and serve as a diagnostic instrument for structural change.

From a mathematical standpoint, the constancy of early second differences confirmed the appropriateness of a quadratic specification for the initial production period. Economically, the negative quadratic coefficient identified a regime of decreasing marginal costs, offering actionable evidence of scale economies. Methodologically, the integration of finite differences with calculus-based interpretation bridges the gap between discrete operational data and continuous economic theory.

Managerially, the model equips decision-makers with a clear, auditable cost structure that supports production planning, efficiency assessment, and strategic communication. Its transparency aligns with the growing demand for explainable analytics in Industry 4.0 environments. Finally, the study offers a replicable pedagogical template that strengthens quantitative reasoning by connecting foundational mathematics to authentic business challenges.

Future research should extend the framework to larger, multi-firm datasets, incorporate real output measures, and explore piecewise-quadratic or multivariate specifications that can accommodate structural breaks. The integration of this approach with causal inference methods and real-time analytics platforms represents a promising direction for advancing interpretable, high-impact business analytics.

References

- [1] Arrieta, A. B., *et al.* (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82–115.
- [2] Christensen, L. R., Jorgenson, D. W., & Lau, L. J. (1973). Transcendental logarithmic production frontiers. *The Review of Economics and Statistics*, 55(1), 28–45.
- [3] Cruz Treminio, L. (2024). *Quadratic modeling of production costs based on real manufacturing data (2015–2021)* [[Academic poster](#)]. Keiser University Latin American Campus.
- [4] Cruz Treminio, L. (2026). Actividad de consolidación del segundo examen parcial como base para la contextualización y aplicación del proyecto integrador [[Trabajo académico no publicado](#)]. *College Algebra*, Keiser University Latin American Campus.
- [5] Production cost data for a manufacturing firm (2015–2021) [[Data set](#)]. Keiser University Latin American Campus.

- [6] Davenport, T. H., & Harris, J. G. (2007). *Competing on analytics: The new science of winning*. Harvard Business Press.
- [7] Grigoryan, L. H., & Hakobyan, A. S. (2023). Production cost forecasting for a given volume of output in organizations: Case study based on a regression model. *International Journal of Professional Business Review*, 8(11), Article e03216. <https://doi.org/10.26668/businessreview/2023.v8i11.3216>
- [8] Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2018). A survey of methods for explaining black box models. *ACM Computing Surveys*, 51(5), 1–42.
- [9] Hernández G., F. J. (2026). *Study guide: Mathematical modeling for business decision making* [Unpublished course material]. Department of Mathematics, Keiser University Latin American Campus.
- [10] Kwon, H., *et al.* (2020). Machine learning-based production cost prediction for smart factories. *Procedia Manufacturing*, 51, 176–181.
- [11] LeVeque, R. J. (2007). *Finite difference methods for ordinary and partial differential equations*. SIAM.
- [12] Martínez-Budría, E., Jara-Díaz, S. R., & Ramos-Real, F. J. (2003). Adapting productivity theory to the quadratic cost function: An application to the Spanish electric sector. *Journal of Productivity Analysis*, 20(2), 213–229. <https://doi.org/10.1023/A:1025184306832>
- [13] Montgomery, D. C., Peck, E. A., & Vining, G. G. (2021). *Introduction to linear regression analysis* (6th ed.). Wiley.
- [14] Pindyck, R. S., & Rubinfeld, D. L. (2018). *Microeconomics* (9th ed.). Pearson.
- [15] Provost, F., & Fawcett, T. (2013). *Data science for business*. O’Reilly Media.
- [16] Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5), 206–215.
- [17] Samuelson, P. A. (1947). *Foundations of economic analysis*. Harvard University Press.
- [18] Simon, C. P., & Blume, L. (1994). *Mathematics for economists*. Norton.

- [19] Varian, H. R. (2014). *Intermediate microeconomics: A modern approach* (9th ed.). Norton.
- [20] Wang, S., Wan, J., Li, D., & Zhang, C. (2016). Implementing smart factory of Industrie 4.0: An outlook. *International Journal of Distributed Sensor Networks*, 12(1), 3159805.
- [21] Wooldridge, J. M. (2016). *Introductory econometrics: A modern approach* (6th ed.). Cengage.

A Appendix: Integrative Course Project

The preliminary mathematical procedures underpinning this study were originally developed within the framework of prior academic coursework [4]. These procedures encompass data organization, variable transformation, and exploratory techniques that informed the formal modeling process. Following the methodological structure of the course study guide [9], the modeling was organized into sequential analytical stages. While initially conceived for instructional purposes, the methods were refined to ensure methodological rigor and are presented here to enhance transparency and facilitate replication.