Delivery Time Deviations Under Normal Operational Conditions (Invisible Risk)

A Secondary Data Analysis Using SPSS

Abstract

Delivery delays in logistics systems are commonly attributed to extreme operational conditions such as heavy traffic or adverse weather. However, delays are frequently observed even when operations appear normal. This study investigates whether delivery time deviations persist under statistically defined normal conditions using a secondary logistics dataset. Normal operational ranges were defined using percentile-based criteria, and delivery time deviation was examined using descriptive statistics and one-way analysis of variance (ANOVA). Results show that delivery delays persist even under normal conditions, with no statistically significant or practically meaningful differences across levels of operational normality. These findings suggest the presence of hidden or systemic inefficiencies embedded within routine logistics operations.

Introduction

Timely delivery is a critical performance indicator in logistics and supply chain management. Traditional analyses of delivery delays often focus on extreme operational conditions such as congestion peaks, severe weather events, or system disruptions. While such factors undeniably contribute to delays, operational experience suggests that delivery time deviations also occur during seemingly routine conditions.

This study challenges the assumption that normal operating conditions guarantee stable delivery performance. Instead of investigating extreme scenarios, the focus is placed on understanding whether delivery delays persist under normal traffic, weather and other conditions, and whether increasing operational normality meaningfully reduces delivery time deviation.

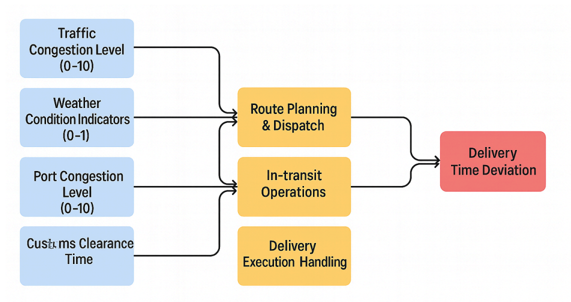

Figure 1 : Conceptual Framework Illustrating operational condition variables and delivery time deviation

Data Source and Study Design

Dataset Origin

This study is a **secondary data analysis** based on a publicly available logistics dataset obtained from **Kaggle**:

> *Dynamic Supply Chain Logistics Dataset*
> Source: Kaggle (user-contributed dataset)
> https://www.kaggle.com/datasets/datasetengineer/logistics-and-supply-chain-dataset?utm_source=chatgpt.com

The dataset contains operational-level logistics records, including delivery time metrics, traffic congestion indicators, weather-related variables, and additional system-level characteristics.

Study Design

This is an **observational, cross-sectional quantitative study** conducted using SPSS. No data collection was performed by the researcher; all analyses were based on existing recorded data. The Sample size is 32065.

Variables Used in the Analysis

Outcome Variable

- **Delivery Time Deviation**
  (Difference between actual and expected delivery time in hours). This variable captures both early and late deliveries but is primarily interpreted as an indicator of delay magnitude.

Operational Condition Variables

- **Traffic Congestion Level -** The level of traffic congestion affecting the logistics route (scale 0-10)
- **Weather Condition Indicators -** The severity of weather conditions affecting operations (scale 0-1)
- **Port Congestion Level -** The level of congestion at the port (scale 0-10)
- **Fatigue Monitoring Score (Driver Condition) -** A score indicating the level of driver fatigue (scale 0-1).
- **Customs Clearance Time:** The time required to clear customs for shipments.

Several variables required preprocessing due to data type inconsistencies (string formats, scientific notation, and nominal measurement levels). These issues were resolved prior to analysis using SPSS transformation and recoding procedures.

Preliminary Analysis

Descriptive Statistics

Descriptive analysis of delivery time deviation revealed substantial variability, indicating that delays are a persistent feature of the dataset.

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
delivery_time_deviation	32065	-1.999997989	10.000000000	5.1776480689	4.1579879273
Valid N (listwise)	32065				

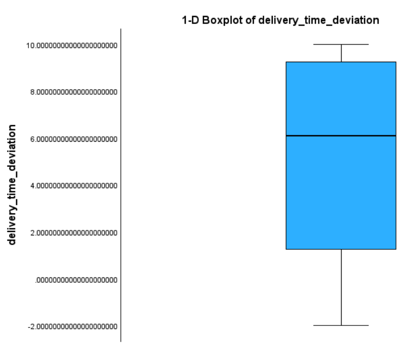

Figure 2 : Descriptive statistics of delivery time deviation

Descriptive statistics indicated that delivery time deviation
exhibited substantial variability (Mean = 5.18, SD = 4.16), with values ranging from –2.00 to 10.00. The non-zero mean and large dispersion relative to the average suggest that delivery delays are a persistent and unstable feature of the dataset.

Correlation Analysis

Bivariate Pearson correlations between delivery time deviation and operational variables were examined.

Correlations

		delivery_time_deviation	traffic_congestion_level	weather_condition_severity	port_congestion_level	customs_clearance_time	fatigue_monitoring_score
delivery_time_deviation	Pearson Correlation	1	.001	-.001	.011	-.009	.003
	Sig. (2-tailed)		.927	.923	.057	.098	.532
	N	32065	32065	32065	32065	32065	32065
traffic_congestion_level	Pearson Correlation	.001	1	.001	-.007	-.003	.010
	Sig. (2-tailed)	.927		.901	.215	.578	.071
	N	32065	32065	32065	32065	32065	32065
weather_condition_severity	Pearson Correlation	-.001	.001	1	-.001	.000	.003
	Sig. (2-tailed)	.923	.901		.803	.994	.619
	N	32065	32065	32065	32065	32065	32065
port_congestion_level	Pearson Correlation	.011	-.007	-.001	1	.005	-.001
	Sig. (2-tailed)	.057	.215	.803		.362	.924
	N	32065	32065	32065	32065	32065	32065
customs_clearance_time	Pearson Correlation	-.009	-.003	.000	.005	1	-.007
	Sig. (2-tailed)	.098	.578	.994	.362		.226
	N	32065	32065	32065	32065	32065	32065
fatigue_monitoring_score	Pearson Correlation	.003	.010	.003	-.001	-.007	1
	Sig. (2-tailed)	.532	.071	.619	.924	.226	
	N	32065	32065	32065	32065	32065	32065

Table 1 : Pearson Correlation Matrix

Bivariate Pearson correlation analysis revealed negligible linear associations between delivery time deviation and key operational variables, with correlation coefficients close to zero. None of the examined relationships were statistically significant, despite the large sample size. These findings indicate that delivery delays are not systematically associated with individual operational factors, reinforcing the presence of latent or systemic sources of delay not captured by observable conditions.

Multiple Linear Regression Analysis

A multiple linear regression model was fitted with delivery time deviation as the dependent variable and all five operational variables as predictors.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.015[a]	.000	.000	4.1578725852

a. Predictors: (Constant), fatigue_monitoring_score, port_congestion_level, weather_condition_severity, customs_clearance_time, traffic_congestion_level

Table 2: Model Results

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	117.194	5	23.439	1.356	.238[b]
	Residual	554232.928	32059	17.288		
	Total	554350.123	32064			

a. Dependent Variable: delivery_time_deviation

b. Predictors: (Constant), fatigue_monitoring_score, port_congestion_level, weather_condition_severity, customs_clearance_time, traffic_congestion_level

Table 3 : ANOVA

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	5.115	.089		57.276	<.001
	traffic_congestion_level	.001	.007	.001	.093	.926
	weather_condition_severity	-.006	.066	-.001	-.096	.924
	port_congestion_level	.014	.007	.011	1.909	.056
	customs_clearance_time	-.025	.015	-.009	-1.660	.097
	fatigue_monitoring_score	.041	.067	.003	.614	.539

a. Dependent Variable: delivery_time_deviation

Table 4: Coefficients Table

Interpretation

The multiple linear regression model demonstrated negligible explanatory power ($R^2 \approx 0.000$), with the overall model failing to achieve statistical significance ($F(5, 32059) = 1.356$, $p = .238$). None of the operational predictors exhibited statistically significant standardized effects. These findings indicate that delivery time deviation cannot be adequately explained through linear combinations of isolated operational variables. This suggests that delivery delays are likely driven by complex system-level interactions, stochastic processes, or unobserved structural factors rather than individual operational stages.

Defining “Normal” Operational Conditions

Each operational variable was categorized into low, normal, and high groups using visual binning.
Rather than relying on subjective thresholds, normality was defined using **percentile-based statistical criteria**.
Normal operating conditions were defined uniformly for all operational variables using the **interquartile range (IQR)**, bounded by the 25th and 75th percentiles of each variable’s distribution.
Justification: The dataset exhibits high variability and the presence of extreme values. Mean-based definitions (e.g., mean ± standard deviation) would include atypical conditions and be sensitive to outliers. The IQR approach focuses on the central 50% of observations, providing a robust, data-driven definition of routine operations.
The dataset was then filtered to retain only those cases.
This filtering ensured that subsequent analyses were conducted strictly within statistically defined normal operational environments.

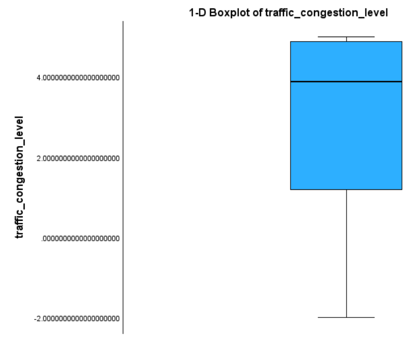

Figure 3.1 Boxplot of traffic congestion level before filtering to retain observations within the interquartile (normal) range.

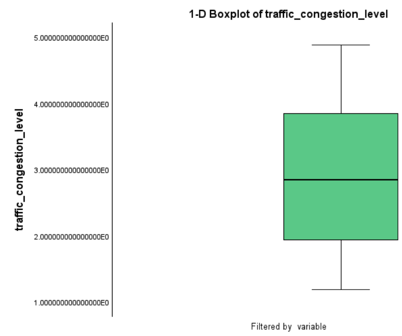

Figure 3.2 Boxplot of traffic congestion level after filtering to retain observations within the interquartile (normal) range.

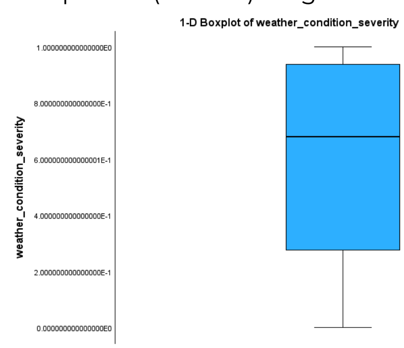

Figure 4.1 Boxplot of weather condition severity before filtering to retain observations within the interquartile (normal) range.

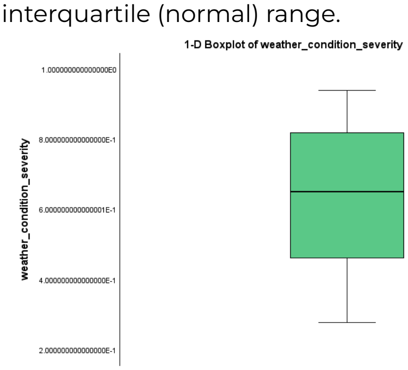

Figure 4.2 Boxplot of weather condition severity after filtering to retain observations within the interquartile (normal) range.

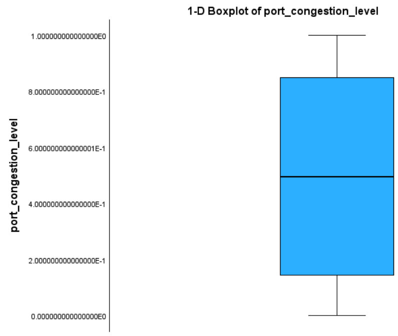

Figure 5.1 Boxplot of port congestion level before filtering to retain observations within the interquartile (normal) range.

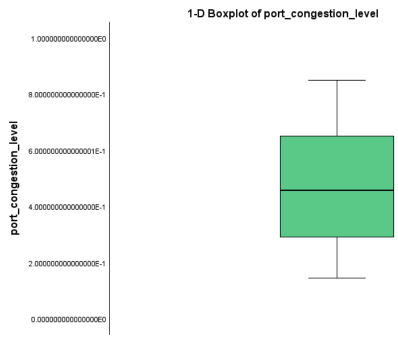

Figure 5.2 Boxplot of port congestion level after filtering to retain observations within the interquartile (normal) range.

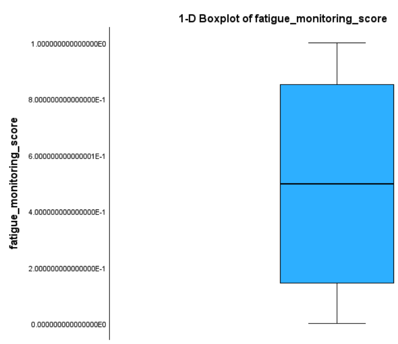

Figure 6.1 Boxplot of fatigue monitoring score before filtering to retain observations within the interquartile (normal) range.

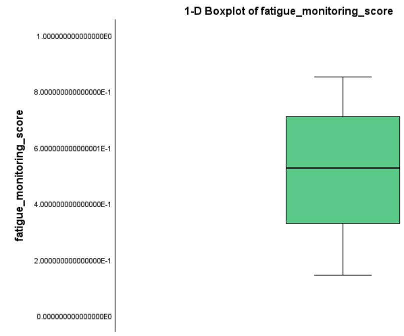

Figure 6.2 Boxplot of fatigue monitoring score after filtering to retain observations within the interquartile (normal) range

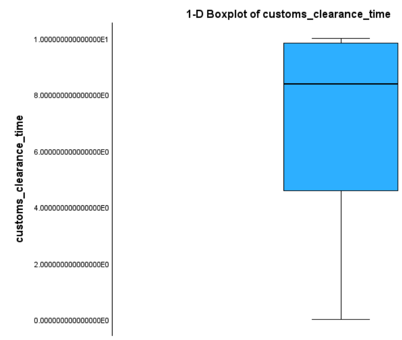

Figure 7.1 Boxplot of customs clearance time before filtering to retain observations within the interquartile (normal) range

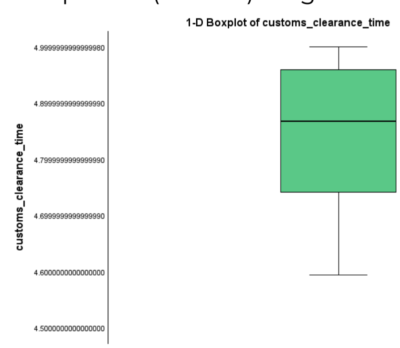

Figure 7.2 : Boxplot of customs clearance time after filtering to retain observations within the interquartile (normal) range

Ex: Normal Traffic Conditions

Normal traffic was defined as values falling within the **interquartile range (IQR)** of traffic congestion levels:

- Lower bound: 25th percentile
- Upper bound: 75th percentile

The normal traffic range was defined using the 25th and 75th percentiles (interquartile range) of traffic congestion levels. This method was chosen because the data show high variability (SD = 2.27) and extreme values (min = –2, max = 5), making mean-based ranges (e.g., mean ±

SD) overly broad and influenced by outliers. The 25th–75th percentile approach focuses on the central 50% of observations, representing typical day-to-day traffic conditions while excluding unusual low or high congestion, providing a more robust and realistic definition of “normal traffic.”

Statistics

traffic_congestion_level

N	Valid	32065
	Missing	1
Percentiles	25	1.1854283752
	75	4.8843667196

Table 5: Percentile table showing 25th and 75th percentiles for traffic congestion level.

Statistics

weather_condition_severity

N	Valid	32065
	Missing	1
Percentiles	25	.27708646856
	75	.93816825089

Table 6 : Percentile table showing 25th and 75th percentiles for Weather Condition Severity

Statistics

port_congestion_level

N	Valid	32065
	Missing	1
Percentiles	25	.14397475890
	75	.84983720494

Table 7 : Percentile table showing 25th and 75th percentiles for Port Congestion Level

Statistics

fatigue_monitoring_score

N	Valid	32065
	Missing	1
Percentiles	25	.14435602707
	75	.85109081618

Table 8 : Percentile table showing 25th and 75th percentiles for fatigue monitoring score

Statistics

customs_clearance_time

N	Valid	32065
	Missing	1
Percentiles	25	4.5932882922
	75	9.8361653320

Table 9 : Percentile table showing 25th and 75th percentiles for customs clearance time

Descriptive Analysis Under Normal Conditions

Descriptive statistics were computed for delivery time deviation under normal traffic conditions.
Key descriptive measures included:

- Mean
- Standard deviation
- Minimum and maximum values

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
delivery_time_deviation	8660	-1.999969176	9.9999999998	5.2038673147	4.1738256362
Valid N (listwise)	8660				

Table 10 : Descriptive statistics for delivery time deviation under normal traffic conditions.

Results indicated that delivery time deviation remained **substantially non-zero**, with considerable variability even under normal traffic conditions.

Comparative Analysis: Normal vs Non-Normal Conditions

One-Way ANOVA

A one-way ANOVA was conducted to compare delivery time deviation between:

- Normal conditions
- Non-normal (extreme) conditions

These suggest that delivery time deviation does not differ meaningfully between normal and extreme conditions.

ANOVA

delivery_time_deviation

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	56.827	1	56.827	3.287	.070
Within Groups	554293.296	32063	17.288		
Total	554350.123	32064			

Table 11 : One-way ANOVA table comparing normal vs non-normal traffic conditions.

ANOVA

delivery_time_deviation

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	5.663	1	5.663	.328	.567
Within Groups	554344.459	32063	17.289		
Total	554350.123	32064			

Table 12 : One-way ANOVA table comparing normal vs non-normal weather conditions.

ANOVA

delivery_time_deviation

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	75.862	1	75.862	4.388	.036
Within Groups	554274.261	32063	17.287		
Total	554350.123	32064			

Table 13 : One-way ANOVA table comparing normal vs non-normal port congestion levels.

$$\eta^2 = \frac{SS_{between}}{SS_{total}} = \frac{75.024}{554350.123} \approx 0.00014$$

ANOVA Effect Sizes[a,b]

		Point Estimate	95% Confidence Interval Lower	Upper
delivery_time_deviation	Eta-squared	.000	.000	.001
	Epsilon-squared	.000	.000	.000
	Omega-squared Fixed-effect	.000	.000	.000
	Omega-squared Random-effect	.000	.000	.000

a. Eta-squared and Epsilon-squared are estimated based on the fixed-effect model.
b. Negative but less biased estimates are retained, not rounded to zero.

Table 14 : One-way ANOVA Effect Sizes table comparing normal vs non-normal port congestion levels.

Although statistically significant, the effect size was small, indicating that port congestion explains only a small proportion of delivery time variation.

ANOVA

delivery_time_deviation

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	36.285	1	36.285	2.099	.147
Within Groups	554313.838	32063	17.288		
Total	554350.123	32064			

Table 15 : One-way ANOVA table comparing normal vs non-normal fatigue monitoring scores

ANOVA

delivery_time_deviation

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.816	1	.816	.047	.828
Within Groups	554349.307	32063	17.289		
Total	554350.123	32064			

Table 16 : One-way ANOVA table comparing normal vs non-normal customs clearance time

One-way ANOVA tests comparing normal and non-normal conditions showed either weak or non-significant differences in delivery time deviation.
Interpretation: Normal conditions are not inherently protective against delays.

Multidimensional Normality: Normality Score Analysis

To capture operational normality more comprehensively, a **normality score** was constructed by combining multiple operational indicators.

Each operational indicator was recoded into a binary variable taking the value 1 if the observation fell within its statistically defined normal range (25th–75th percentile) and 0 otherwise. The operational normality score was then computed as the sum of these binary indicators:

Normality Score = Normal_Traffic + Normal_Weather + Normal_Port + Normal_Customs + Normal_Fatigue

$$\text{Normality Score}_i = \sum_{k=1}^{5} I(X_{ik} \in \text{IQR}_k)$$

Where:

- I(·) is an indicator function (1 = within normal range, 0 = otherwise)
- Xik is the k-th operational variable for observation i

The resulting score ranged from 0 to 5, with higher values indicating deliveries occurring under increasingly normal operational conditions.

ANOVA Across Normality Score Levels

A one-way ANOVA was performed to examine whether delivery time deviation differed across levels of the normality score.

Results

- **$F(4, 32060) = 1.634$**
- **$p = 0.162$**
- Effect sizes:
    - Eta-squared ≈ 0
    - Omega-squared ≈ 0

These results indicate no statistically significant or practically meaningful differences in delivery time deviation across normality score categories.

ANOVA

delivery_time_deviation

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	113.011	4	28.253	1.634	.162
Within Groups	554237.111	32060	17.287		
Total	554350.123	32064			

Table 17 : ANOVA table for delivery time deviation by normality score.

ANOVA Effect Sizes[a,b]

		Point Estimate	95% Confidence Interval Lower	95% Confidence Interval Upper
delivery_time_deviation	Eta-squared	.000	.000	.001
	Epsilon-squared	.000	.000	.000
	Omega-squared Fixed-effect	.000	.000	.000
	Omega-squared Random-effect	.000	.000	.000

a. Eta-squared and Epsilon-squared are estimated based on the fixed-effect model.

b. Negative but less biased estimates are retained, not rounded to zero.

Table 18 : Effect size estimates (η^2, ω^2, ε^2).

Interpretation of Effect Sizes

All estimated effect sizes were approximately zero, indicating that the proportion of variance in delivery time deviation explained by operational normality is negligible.

This confirms that increasing operational normality does not substantially reduce delivery time deviations (delays).

Figure 8 : Line chart of mean delivery time deviation vs. normality score

The results show that delivery time deviation remains relatively constant as the normality score increases, with no clear decreasing trend. Even when multiple operational conditions fall within statistically defined normal ranges, delivery delays persist at comparable levels. This visual pattern supports the ANOVA findings, indicating that increasing operational normality does not meaningfully reduce delivery time deviation and suggesting the presence of latent, system-level inefficiencies within routine logistics operations.

Key Findings

- Delivery delays persist even under statistically defined normal conditions.
- Traffic normality alone does not significantly reduce delivery time deviation.
- Combining multiple normal operational indicators into a normality score does not yield meaningful explanatory power.
- Operational normality explains **virtually none** of the variability in delivery performance.

Discussion

The findings challenge the widely held assumption that delivery delays primarily arise from extreme or abnormal conditions. Instead, the results suggest that delivery time deviations are embedded within routine logistics operations.

The absence of statistically significant effects is not a limitation but a critical insight: systemic inefficiencies and variability exist even when observable operational conditions appear stable.

Conclusion

This SPSS-based secondary data analysis demonstrates that delivery delays persist under normal traffic and operational conditions, with negligible differences across increasing levels of operational normality. The results highlight the presence of hidden risks within routine logistics systems and suggest that improving delivery performance requires addressing systemic operational inefficiencies rather than focusing exclusively on extreme disruptions.

Limitations and Assumptions

This study is subject to several limitations and analytical assumptions that should be considered when interpreting the findings. First, the analysis is based on secondary, publicly available logistics data, limiting control over data collection procedures, measurement accuracy, and underlying data-generating processes. Consequently, the presence of measurement error, unobserved operational factors, or reporting inconsistencies cannot be ruled out. In addition, the cross-sectional observational design precludes causal inference; the results describe patterns and associations in delivery time deviation rather than causal relationships.
Methodologically, operational normality was defined using interquartile ranges (25th–75th percentiles), assuming that the central distribution of each variable represents routine operating conditions. While this percentile-based approach is robust to extreme values, statistically normal conditions may not fully correspond to operationally optimal or practitioner-defined norms. The operational normality score was constructed using binary coding with equal weighting across indicators, reflecting the absence of prior empirical evidence supporting differential importance among operational dimensions. This additive structure assumes independence and does not capture nonlinear effects or interactions.
Finally, although the large sample size enhances statistical stability, it also increases the likelihood of statistically significant results with minimal practical relevance. For this reason, effect sizes were emphasized alongside p-values throughout the analysis. Future research could address these limitations by incorporating nonlinear models, interaction effects, hierarchical or longitudinal designs, and system-level process measures to better capture the complexity of delivery delay mechanisms.