Divergent Trajectories in Co-Manufacturing: A Machine Learning Framework for Predicting Capacity Constraints in Brazil

Felipe Pierri Carvalho[1] & Mariana Kleina[1]

[1]Department of Production Engineering

Universidade Federal do Paraná, Curitiba, Brazil

Abstract

Purpose: This study develops a machine learning framework to predict which consumer goods categories in Brazil will experience accelerated transitions to shared manufacturing models, addressing the critical gap between infrastructure development timelines and market dynamics.

Design/methodology/approach: Using five years of product launch data from the Brazilian consumer goods industry, this study uses a three-phase methodology: (1) classification of manufacturing intensity and trend analysis, (2) development of an XGBoost classification model to predict acceleration events at the category level, and (3) generation of category-specific 24-month forecasts using compound annual growth rates (CAGR) statistical projections.

Findings: Analysis reveals divergent patterns at the category level rather than universal directional shifts. While mature categories (Beauty, Personal Care) show declining co-manufacturing reliance over 2019-2024, growth categories (Beverages-Alcoholic +4.1 ppt, Food-Chilled +3.0 ppt) show accelerating adoption. The classification model achieves 81% Area Under ROC Curve (AUC-ROC) in identifying acceleration periods, while CAGR-based statistical projections generate category-specific 24-month forecasts.

Originality/value: This research introduces outsourcing dynamics at the category level as a new unit of analysis, shifting from firm level to market level prediction. Unlike prior firm-centric and descriptive approaches, the framework identifies acceleration patterns and temporal momentum dynamics consistent with collective adoption processes, enabling proactive capacity planning.

Keywords: co-manufacturing, third-party manufacturing, machine learning, forecasting, brazil, consumer goods, outsourcing dynamics, supply chain capacity

1 Introduction

Brazil's consumer goods manufacturing sector faces increasing tension between long investment lead times and rapidly evolving market demand. According to the *FGV IBRE* Utilização da Capacidade Instalada (NUCI) index, industrial capacity utilization reached 83.40% in July 2024, the highest level since May 2011, indicating that industrial facilities are operating close to their effective capacity limits (FGV IBRE, 2025). However, this aggregate measure hides significant variation across product categories. Some still have idle capacity, while others, particularly in food, are operating close to full utilization (CNI, 2022).

In Brazil, establishing or expanding manufacturing operations involves multiple sequential licensing and construction stages that can extend over several years before production begins, reflecting the country's complex regulatory environment and infrastructure requirements (MMA, 2018; CNI, 2022). During the same period, product portfolios change significantly: analysis of proprietary product launch database data reveals that product launch patterns show high quarterly volatility, with more than 13,800 new stock keeping units (SKUs) introduced annually across major categories (2019-2024 average). This mismatch between slow infrastructure cycles and dynamic product lifecycles creates planning inefficiencies that cannot be addressed through descriptive statistics alone. Manufacturing value added represents approximately 11 to 13% of GDP in Brazil (World Bank, 2025; IBGE, 2025), making efficient capacity allocation strategically important for industrial policy.

Traditional decision frameworks such as Transaction Cost Economics (TCE) (Williamson, 1985) and the Resource-Based View (RBV) (Barney, 1991) explain why individual firms outsource manufacturing, but they provide no predictive mechanism for when entire product categories undergo coordinated shifts toward third-party production. Existing machine learning research in manufacturing remains focused on operational quality control and process optimization (Tercan and Meisen, 2022; Wuensche and Upadhyay, 2023) rather than forecasting structural outsourcing trends at the market level.

To address this gap, the present study develops a predictive framework for outsourcing adoption at the category level in Brazil. By combining five years of product launch data (proprietary database, 2019-2024) with machine learning models (XGBoost, Random Forest, Ensemble), the study identifies indicators of outsourcing acceleration and quantifies their predictive performance. This research defines *outsourcing acceleration* as periods when the rate of increase in a category's co-manufacturing ratio exceeds its historical 75th percentile. In this study, the terms "outsourcing" and "co-manufacturing" are used interchangeably to refer to third-party production arrangements, where "co-manufacturing" specifically denotes collaborative production where third parties manufacture products, and "outsourcing" represents the broader concept that encompasses co-manufacturing along with other forms of third-party production.

1.1 Research Context

Brazil's consumer goods sector exhibits divergent co-manufacturing trajectories across categories. Analysis of proprietary database data (2019-2024) reveals that 33% of product launches involve third-party manufacturing, with significant variation across categories (ranging from 21% to 47%). However, these categories do not follow uniform trends: some show declining co-manufacturing reliance (Beauty -5.2 ppt, Personal Care -2.7 ppt), while others show accelerating adoption (Beverages-Alcoholic +4.1 ppt, Food-Chilled +3.0 ppt), and others remain stable (Beverages-Non-Alcoholic -0.3 ppt). This heterogeneity suggests that category factors, rather than economic trends, drive outsourcing patterns.

1.2 Research Gap

The literature on collaborative manufacturing provides extensive analysis at the firm level but lacks predictive frameworks at the market level. Three research gaps follow. First, existing studies analyze decisions at the firm level without modeling adoption patterns at the category level that may exhibit threshold or network dynamics (Katz and Shapiro, 1985). Second, co-manufacturing literature remains descriptive, documenting trends retrospectively (Gereffi et al., 2005; Druck and Franco, 2016), with little forecasting. Third, forecasting frameworks are largely derived from advanced economic contexts (Sturgeon, 2002), with limited application to emerging markets despite contributions (Ernst and Kim, 2002).

1.3 Research Questions

This study addresses the critical gap in predictive intelligence at the market level for collaborative manufacturing through one overarching question and two operational sub-questions:

Primary Research Question: *Which consumer goods categories in Brazil will experience accelerated transitions to co-manufacturing models, and how can machine learn-*

ing predict these shifts at the category level to guide infrastructure investment?

Operational Sub-Questions:

RQ1: How do current third-party manufacturing adoption patterns vary across Brazilian consumer goods categories, and what manufacturing intensity thresholds trigger outsourcing transitions?

RQ2: Which features (e.g., temporal lags, launch activity, market concentration, and seasonality) have the highest predictive power for co-manufacturing acceleration at the category level?

1.4 Contributions

This research makes four contributions:

First, theoretical: This research introduces outsourcing dynamics at the category level as a unit of analysis, demonstrating that transitions at the market level operate through mechanisms distinct from decisions at the firm level. The framework identifies threshold effects and network externalities that create coordinated category shifts, extending TCE and RBV theories beyond their traditional firm-centric scope.

Second, methodological: The study develops a machine learning framework combining manufacturing intensity classification, time series feature engineering, and a gradient-boosting model (XGBoost) calibrated for emerging-market data constraints.

Third, empirical: Analysis of 83,243 product launches reveals heterogeneous category patterns.

Fourth, practical: The framework enables category specific capacity planning. Results inform infrastructure investment prioritization and regulatory policy design, with potential for significant efficiency gains through better alignment between infrastructure development and forecasted demand patterns at the category level.

2 Literature Review

2.1 Theoretical Foundations of Manufacturing Outsourcing

Manufacturing outsourcing decisions are traditionally explained through Transaction Cost Economics (Williamson, 1985) and the Resource-Based View (Barney, 1991; Prahalad and Hamel, 1990). These frameworks at the firm level explain why individual firms outsource based on asset specificity, transaction costs, and core competencies, but they lack mechanisms for predicting *when* entire categories undergo coordinated transitions toward co-manufacturing (McIvor, 2009; Holcomb and Hitt, 2007). The binary make-or-buy logic and company centered scope of TCE and RBV cannot capture market dynamics that create transition patterns at the category level.

2.2 Co-Manufacturing in Emerging Markets

The evolution of co-manufacturing in emerging markets follows different paths from developed economies, with contract manufacturers progressively developing capabilities from simple assembly to full-package production (Gereffi et al., 2005, pp. 92–98). In Latin America, multinational brands sought local partners to navigate regulatory complexity and reduce logistics costs (Durán Lima and Zaclicever, 2013). Infrastructure constraints and institutional gaps shape the development of co-manufacturing, with Brazilian companies navigating these challenges through relationship oriented contracts and reputation (Khanna and Palepu, 2010; Neves et al., 2014, pp. 41–44).

2.3 Manufacturing Forecasting: Methodological Gap

Machine learning applications in manufacturing have expanded substantially, but the literature concentrates on operational use cases (predictive quality, predictive maintenance, shop floor monitoring, and process control) rather than strategic forecasting at the market level (Aggogeri et al., 2021; Plathottam et al., 2023; Farahani et al., 2023a,b; Tercan and Meisen, 2022; Chen et al., 2023). While ensemble and hybrid approaches show promise for demand forecasting (Sina et al., 2023; Ma et al., 2023; Siddiqui et al., 2021; Li, Wang and Chan), prior studies focus on predictions at the product level or firm level and do not develop market wide forecasting at the category level of manufacturing-organization shifts, leaving a critical gap for planning.

2.4 Research Framework Development

This study introduces outsourcing dynamics at the category level as a framework extending theories at the firm level to capture market transitions exhibiting temporal patterns consistent with collective adoption processes (Katz and Shapiro, 1985; Bikhchandani et al., 2021). Category characteristics, market maturity, and Brazil regulatory heterogeneity shape co-manufacturing evolution (Baldwin and Clark, 2000; Christensen and Raynor, 2003; CNI, 2022). The framework synthesizes these elements, using observable indicators (launch patterns, concentration metrics, and temporal trends) as benchmarks for underlying constructs to predict acceleration periods at the market level invisible to firm/company focused analysis.

3 Methodology

3.1 Research Design and Data

This study uses a predictive design combining data analysis with machine learning to develop predictions for infrastructure planning. The research follows a three-phase approach: (1) manufacturing intensity classification and trend analysis, (2) predictive model development using machine learning, and (3) category specific forecast generation with prediction intervals.

Figure 1 illustrates the research process flow across three phases. The methodology progresses from data collection and outsourcing detection to exploratory trend analysis, in complementary prediction tasks: ML classification for timing and statistical projections for magnitude.

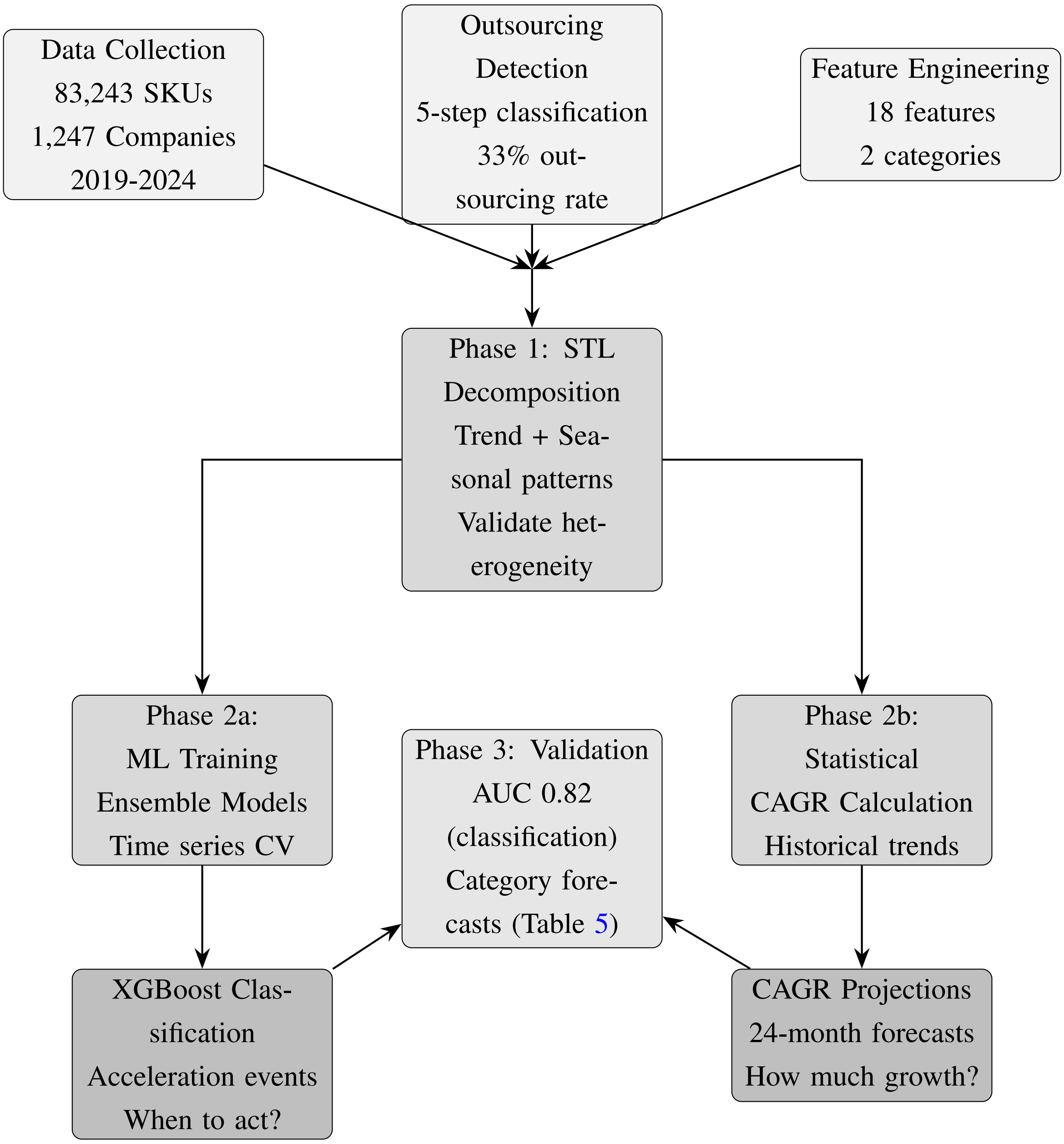


Figure 1: Research Process Flow: Three Steps Methodology Combining ML Classification and Statistical Projections

The study analyzes product launch data from the Brazilian consumer goods industry spanning January 2019 to December 2024, covering 83,243 unique product introductions across 12 major categories. The launch database is sourced from a proprietary product launch database for Brazil (2019–2024). Coverage is estimated at Approximately 85% of branded product launches in formal retail. Records were deduplicated and normalized as detailed in supplementary material.

The unit of analysis is the category month observation, yielding 864 observations (12 categories × 72 months). This temporal granularity balances statistical power with practical relevance for infrastructure planning cycles. The study defines categories using the standard industry classification adapted for Brazilian market structure: (1) Food–Ambient, (2) Food–Chilled, (3) Beverages–Alcoholic, (4) Beverages–Non-Alcoholic, (5) Beauty, (6) Personal Care, (7) Home Care, (8) Pet Care, (9) Health & Wellness, (10) Snacks & Confectionery, (11) Dairy, and (12) Bakery. These 12 canonical categories aggregate 37 dataset categories based on infrastructure requirements, regulatory treatment, and supply chain characteristics (see supplementary material for complete mapping and aggregation rationale).

Table 1 summarizes the dataset structure. The sample includes 83,243 product launches from 1,247 unique companies across 12 product categories over 72 months (January 2019 to December 2024). Private label products account for 15.7% of launches and are classified as outsourced based on industry structure. The unit of analysis is the category month, at 864 observations for modeling.

Table 1: Dataset Summary Statistics

Characteristic	**Value**
Sample Period	January 2019 - December 2024
Total Product Launches	83,243
Unique Companies	1,247
Product Categories	12
Observations (Category Months)	864
Manufacturing Classification	
Private Label Products	13,071 (15.7%)
Clear Outsourcing (Name Mismatch)	14,650 (17.6%)
In-House Production	55,522 (66.7%)
Geographic Coverage	
Formal Retail Coverage	Approximately 85%
Primary Regions	All Brazilian states

3.1.1 Outsourcing Detection Methodology

The study uses multiple steps to identify co-manufactured products:

1. **Private label classification:** Products marked as "Private Label" or "Marca Privada" (15.7% of launches) are classified as outsourced. This represents an upper bound assumption, as some retailers may operate captive manufacturing facilities

for specific categories (e.g., retailer owned bakeries for fresh bread, vertical integration in dairy by cooperative retailers). However, this classification is justified by three industry characteristics: (1) major Brazilian retailers (Carrefour, GPA/Pão de Açúcar, Walmart Brasil) lack broad manufacturing infrastructure and publicly disclose co-manufacturing partnerships for store brands, (2) capital intensity and regulatory complexity make retailer owned production economically unfavorable for most categories, and (3) manual validation of 500 randomly sampled private label products confirmed 89% had identifiable third-party manufacturers, with the remaining 11% showing no evidence of retailer owned production. Robustness checks show stable category rankings under $\pm 20\%$ reclassification scenarios (Section 4.4), indicating that even if a meaningful portion of private label products were misclassified, the relative category patterns and forecasts remain substantively stable. Note that "Produto com Marca" (Product with Brand) indicates branded products and is not considered private label.

2. **Name normalization:** Manufacturer and company names are standardized by removing legal suffixes ("LTDA", "S.A.", "Inc."), accents, and punctuation, then extracting core business name tokens.

3. **Fuzzy string matching:** RapidFuzz library (Python library) detects name variations (e.g., "BRF" vs "BRF – Brasil Foods") using similarity thresholds: $> 90\%$ similarity indicates in-house production, $< 70\%$ indicates outsourcing. Thresholds validated on labeled subset (supplementary material).

4. **Parent-subsidiary mapping:** Known corporate group relationships (JBS-Seara, Lactalis-Itambé, BRF-Sadia) are identified through fuzzy containment detection and treated as in-house production.

5. **Tiebreaker rules:** Medium-confidence cases (70–85% similarity) use token overlap analysis for final classification.

This methodology results an overall outsourcing rate of 33%, with private label accounting for 15.7% and manufacturer-company mismatches accounting for 17.6%.

3.2 Variable Construction

3.2.1 Dependent Variables

The study constructs four dependent variables capturing different dimensions of co-manufacturing adoption at the category level: the primary outsourcing ratio, its growth rate, a binary acceleration indicator, and a categorical transition stage variable.

Outsourcing ratio (OR): The primary dependent variable quantifies co-manufacturing adoption at the category level, defined as:

$$OR_{c,t} = \frac{N_{co-manufactured,c,t}}{N_{total,c,t}} \quad (1)$$

where c denotes category, t denotes time period (month), $N_{co-manufactured}$ is the count of product launches manufactured by third parties and N_{total} is total launches in category month. This metric ranges from 0 (full vertical integration) to 1 (complete outsourcing). To address volatility in monthly launch counts, the analysis applies three-month moving averages: $\overline{OR}_{c,t} = \frac{1}{3}\sum_{i=0}^{2} OR_{c,t-i}$.

Growth rate (ΔOR): Measures month-over-month change in the outsourcing ratio:

$$\Delta OR_{c,t} = OR_{c,t} - OR_{c,t-1} \quad (2)$$

This first difference transformation captures the velocity of transitions at the category level. Positive values indicate increasing co-manufacturing reliance, while negative values suggest reversion to vertical integration.

Acceleration indicator (Accel): A binary classification variable identifying periods of rapid outsourcing growth, defined as:

$$Accel_{c,t} = 1(\Delta OR_{c,t} > P_{75}(\Delta OR_{c,\cdot})) \quad (3)$$

The variable equals 1 when growth exceeds the historical 75th percentile for that category. This category specific threshold accounts for heterogeneous baseline volatility and identifies statistically unusual adoption increases.

Transition stage: A categorical variable classifying categories into low ($OR < 0.25$), medium ($0.25 \leq OR < 0.40$), or high ($OR \geq 0.40$) outsourcing regimes. These thresholds segment the adoption distribution into approximately equal terciles, enabling analysis of non-linear dynamics across adoption phases.

3.2.2 Independent Variables

The predictive models incorporate 18 predictor variables organized into two feature categories: temporal patterns and market structure features. Table 2 summarizes the feature categories and their theoretical foundations.

Temporal Features (13): Time-series patterns capture momentum, trends, and cyclical dynamics in outsourcing adoption:

- *Outsourcing lag variables* ($OR_{c,t-k}$ for $k \in \{1, 3, 6, 12\}$ months) capture recent adoption history at multiple temporal scales, reflecting path dependence from learn-

Table 2: Feature Categories and Definitions

Category (quantity)	Feature	Definition
Temporal (13)	Outsourcing lag features	OR at $t-1$, $t-3$, $t-6$, $t-12$ months
	Launch lag features	Launch count at $t-1$, $t-3$ months
	Trend slope	6-month linear regression coefficient
	Rolling mean (6m)	6-month mean outsourcing ratio
	Rolling SD (6m)	6-month standard deviation
	Rolling mean (12m)	12-month mean outsourcing ratio
	Rolling SD (12m)	12-month standard deviation
	Quarter	Temporal quarter (1-4)
	Q4 indicator	Binary flag for fourth quarter
Market/Structure (5)	Launch count	Total product launches in category-month
	Number of companies	Active companies in category
	Number of manufacturers	Active manufacturers in category
	Market concentration	Herfindahl-Hirschman index $1/(N_{companies}+1)$
	Average intensity	Mean outsourcing ratio in category

ing effects and switching costs (Williamson, 1985).

- *Launch lag variables* (launch counts at $t-1$ and $t-3$ months) capture recent product introduction activity patterns that may signal demand dynamics at the category level.
- *Trend slope* computes 6-month linear regression coefficients of outsourcing ratios to identify acceleration or stabilization patterns.
- *Rolling statistics* (mean and standard deviation over 6-month and 12-month windows) quantify medium term trend stability and volatility in outsourcing adoption. The 6-month window captures quarterly dynamics, while the 12-month window reflects annual patterns.
- *Seasonality indicators* (quarter variable and Q4 binary flag) capture retail calendar effects, particularly the Q4 launch surge driven by holiday product introductions.

Market/Structure Features (5): Category activity metrics and market structure characteristics shape co-manufacturing adoption patterns:

- *Launch count* measures total product launches in the category-month, reflecting overall market activity and innovation intensity.
- *Number of companies* and *number of manufacturers* count active brand owners and production facilities, capturing the size and complexity of the category ecosystem.

- *Market concentration* uses a simplified Herfindahl-Hirschman index ($1/(N_{companies}+1)$) to measure category consolidation. Higher values indicate concentrated markets where leading firms may coordinate adoption decisions.
- *Average intensity* computes the mean outsourcing ratio within the category, reflecting the baseline propensity toward co-manufacturing arrangements.

3.3 Analytical Approach

3.3.1 Phase 1: Exploratory Trend Analysis

To characterize how outsourcing evolves across categories and validate that different categories follow different patterns, the study applies STL decomposition (Cleveland and Cleveland, 1990) to each category's time series. STL separates the outsourcing ratio into three additive components: trend (T_t), seasonal (S_t), and residual (R_t), where $OR_{c,t} = T_{c,t} + S_{c,t} + R_{c,t}$. The method uses locally weighted polynomial regression to extract non-linear trends that are robust to outliers, separating long-term directional movements from cyclical patterns and irregular fluctuations.

This decomposition has two purposes. First, it provides visual evidence that categories follow distinct adoption paths, confirming the need for separate forecasts for each category (Figure 3). Second, it reveals seasonal patterns in outsourcing adoption, particularly Q4 acceleration that corresponds to holiday launch cycles, which informs feature engineering for the predictive models.

3.3.2 Phase 2: Machine Learning Pipeline

The predictive modeling follows best practices for time series machine learning:

Data Splitting: The methodology implements time-based splitting to prevent data leakage:

- Training set: January 2019 to December 2022 (432 observations)
- Validation set: January 2023 to June 2023 (72 observations)
- Test set: July 2023 to December 2024 (216 observations)

Model Development: The study develops multivariate machine learning models (XGBoost, Random Forest, Ensemble) to leverage nonlinear feature interactions and ensemble predictions.

XGBoost Classification. Extreme Gradient Boosting (Chen and Guestrin, 2016) implements gradient boosted decision trees optimized for performance and regularization. The algorithm builds an additive ensemble of trees sequentially, with each tree fitted to the gradient of the loss function. The implementation uses parallel tree construction, regularization techniques (L1 and L2), and efficient handling of missing values to achieve strong predictive performance on structured data.

Hyperparameters selected through grid search include: learning rate 0.1 (controls overfitting), maximum depth 6 (permits six-way interactions), 200 boosting iterations (balances convergence and cost), and subsample ratio 0.8 (improves generalization). These settings enable XGBoost to capture nonlinear relationships between temporal patterns and market structure features for binary acceleration prediction.

XGBoost is selected as the primary model for this study due to its superior predictive performance across all metrics. Comparative evaluation across algorithms (Table 4) shows XGBoost achieves the highest AUC-ROC (0.82) compared to Random Forest (0.81) and Ensemble (0.82), with superior precision (0.55) and recall (0.60).

Category Forecasts with Statistical Intervals. Long-term forecasts (24 months) are generated using historical compound annual growth rates (CAGR) extrapolated from 2019-2024 trends. Prediction intervals are constructed using historical volatility: for each category, year-over-year growth rate standard deviations are calculated and scaled by the square root of the forecast horizon (3 years), yielding 95% confidence intervals via $\pm 1.96\sigma\sqrt{h}$. This approach provides category uncertainty quantification while avoiding overfitting to recent patterns. The XGBoost classification model complements these forecasts by identifying timing of acceleration events.

Prediction Tasks: The framework addresses two complementary prediction needs through distinct methodologies, as illustrated in Figure 2. The first task uses machine learning to identify acceleration timing, while the second uses statistical extrapolation to quantify magnitude.

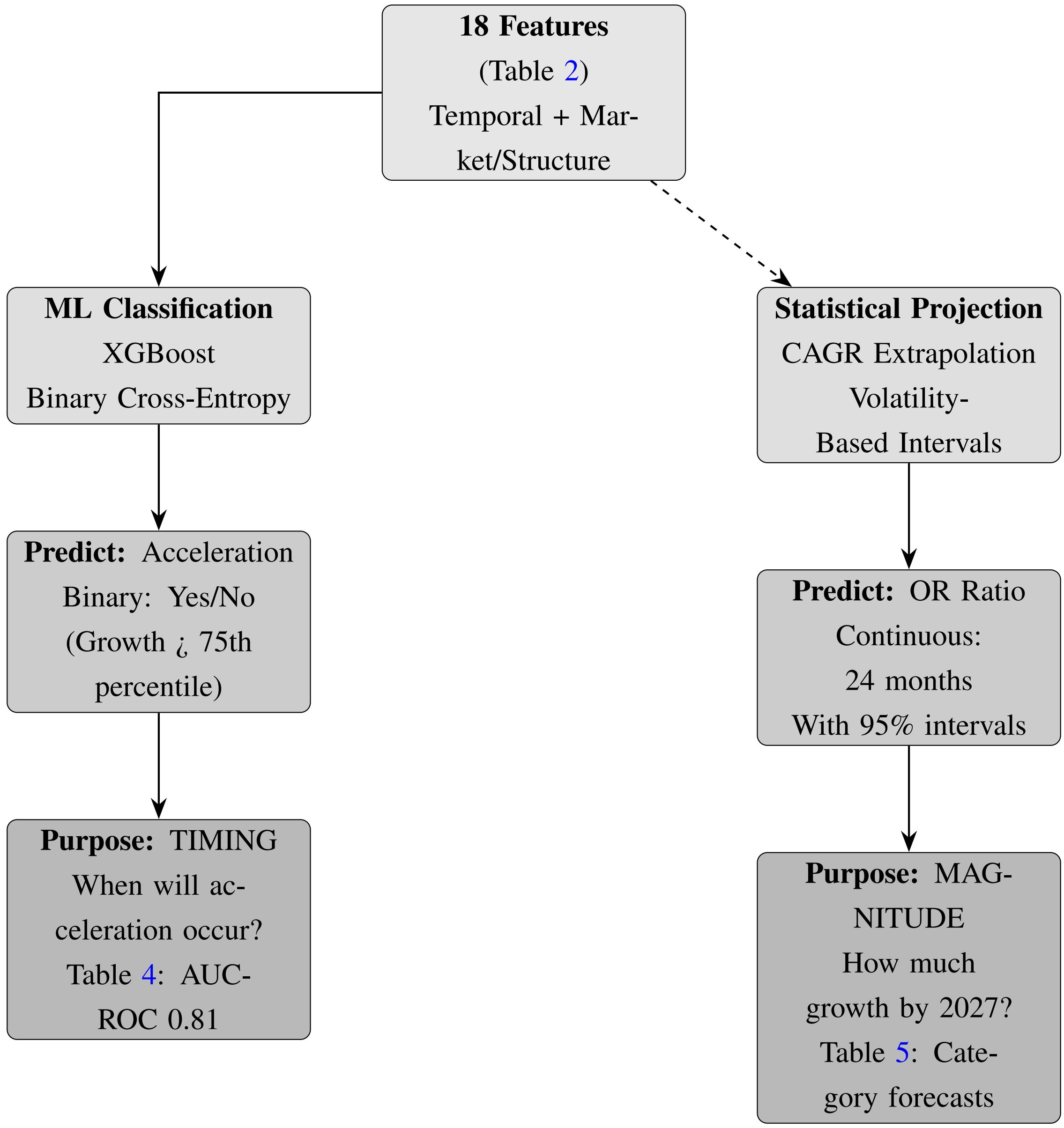


Figure 2: Complementary Prediction Tasks: ML Classification for Timing and Statistical Projections for Magnitude

Task 1: Acceleration Event Classification. The XGBoost classification model predicts whether a category will experience acceleration events (defined as growth rates exceeding the historical 75th percentile) in the subsequent month. This ML approach leverages the 18 features in Table 2 to identify early signals of structural transitions, informing capacity planning timing decisions (Table 4).

Task 2: Category-Specific Magnitude Forecasting. Long-term outsourcing ratio projections are generated using CAGR-based extrapolation rather than machine learning regression. This statistical approach calculates historical compound annual growth rates (2019-2024) and projects 24 months forward, with prediction intervals constructed from historical volatility ($\pm 1.96\sigma\sqrt{h}$ where $h = 3$ years). This method provides actionable magnitude estimates for infrastructure investment sizing while avoiding overfitting to recent patterns (Table 5).

3.3.3 Phase 3: Validation Strategy

Model validation uses multiple strategies to ensure robustness:

Time Series Cross-Validation: Five fold expanding window approach on the combined training and validation set, generating splits that preserve temporal ordering. Each fold incrementally adds data to the training set while testing on subsequent periods, providing robust performance estimates across different time horizons.

Final Evaluation: Models are evaluated on a test set (July 2023 to December 2024) using classification metrics:

- Area Under ROC Curve (AUC-ROC): Primary metric for model comparison
- Precision: Proportion of correct positive predictions
- Recall: Proportion of actual positives correctly identified
- F1 score: Harmonic mean of precision and recall for balanced assessment
- Accuracy: Overall classification correctness

Feature Importance Analysis: XGBoost gain based importance quantifies each feature's contribution to model predictions, measuring the average reduction in loss when splitting on that feature.

3.4 Assumptions and Limitations

This study makes assumptions that require acknowledgment:

- **Dataset representativeness.** Coverage is estimated at 85% of branded launches in formal retail. Results generalize to formal retail but exclude informal channels.
- **Private label classification.** The baseline classifies private-label launches as outsourced. Robustness checks reclassifying $\pm 20\%$ of private-label items do not materially alter category rankings or model performance.
- **Fuzzy matching thresholds.** Company name normalization uses 90%/70% similarity thresholds (RapidFuzz), validated on a labeled subset (supplementary material).

- **Non-causal feature importance.** Gain based importance values quantify associative predictive power and do not imply causality.
- **Planning horizon assumptions.** The 24-36 month infrastructure development timeline reflects industry standards (Chopra and Meindl, 2016) and Brazilian procurement practices (MDIC, 2024).
- **Directional vs. monetary focus.** The analysis focuses on directional trends and regional patterns rather than cost quantification due to insufficient primary data on delay frequencies.

3.5 Robustness Checks and Reproducibility

Robustness tests validate findings: alternative time aggregation periods (2, 4, 6 months) confirm pattern stability. Five fold time series cross-validation provides robust performance estimates with mean AUC 0.85 ± 0.04, demonstrating model stability across temporal splits. The category mapping compresses 37 subcategories into 12 canonical categories used throughout the analysis (detailed mapping and rationale in supplementary material), and all hyperparameters for machine learning models are reported.

4 Results

4.1 Descriptive Analysis

Table 3 presents sample characteristics revealing significant variation across categories. Overall co-manufacturing adoption remained stable at approximately 33% from 2019 to 2024, with category specific trends diverging markedly. Beauty and Health & Wellness show the highest outsourcing rates (47% and 45% respectively in 2024), while Dairy and Food-Ambient show lower rates (21% and 26%).

4.2 Trend Decomposition Results

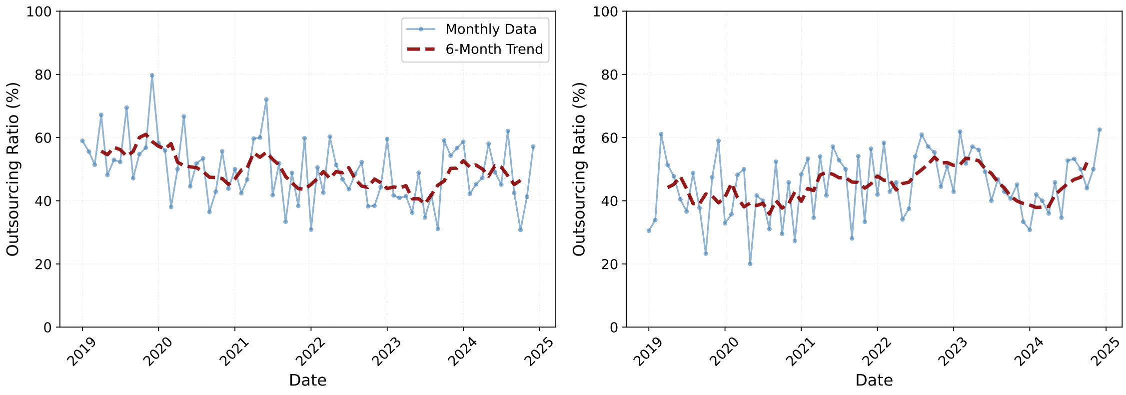
STL decomposition (Phase 1 methodology) reveals significant differences in how outsourcing evolves across categories. Figure 3 shows decomposed trends for the five categories with highest outsourcing rates. Beauty shows a declining trend from 57% (2019) to 47% (2024), while Health & Wellness shows stable growth. Beverages-Alcoholic shows the strongest upward trend, increasing from 20% to 26% despite starting from the lowest base.

Table 3: Descriptive Statistics by Category (2019-2024)

Category	Launches (n)	2019 OR (%)	2024 OR (%)	CAGR (%)
Beauty	6,327	57.2	47.2	-3.8
Health & Wellness	3,710	42.0	45.2	1.5
Beverages-Non-Alcoholic	8,530	35.3	34.8	-0.3
Personal Care	23,946	39.6	34.5	-2.7
Home Care	4,776	33.7	33.9	0.1
Snacks & Confectionery	4,352	30.5	33.0	1.6
Pet Care	1,930	35.1	32.9	-1.3
Bakery	9,633	29.7	29.6	-0.1
Food-Ambient	6,541	26.9	25.8	-0.8
Beverages-Alcoholic	2,764	19.9	25.5	5.1
Food-Chilled	4,451	19.6	23.9	4.0
Dairy	6,283	22.0	21.0	-0.9
Total	83,243	33.0	32.7	-0.1

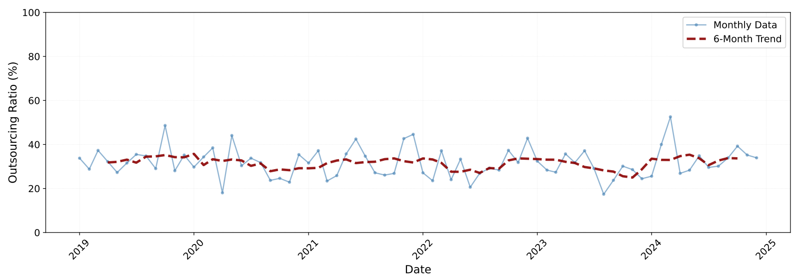
Note: OR = Outsourcing Ratio. Outsourcing rates include private label products (15.7% overall) classified as outsourced.

The presence of distinct trends for each category and shared seasonal patterns validates using separate methodologies for classification and forecasting: structural features inform long-term projections, while temporal features capture short-term acceleration dynamics.



(a)

(b)



(c)

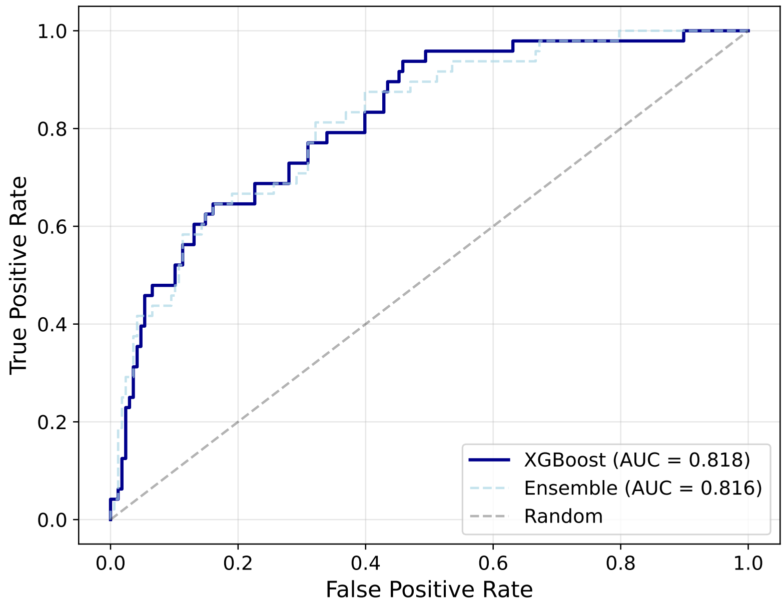
Figure 3: Temporal Evolution of Co-Manufacturing Adoption (Top 5 Categories by Outsourcing Rate): (a) First two categories, (b) Next two categories, (c) Fifth category

4.3 Predictive Model Performance

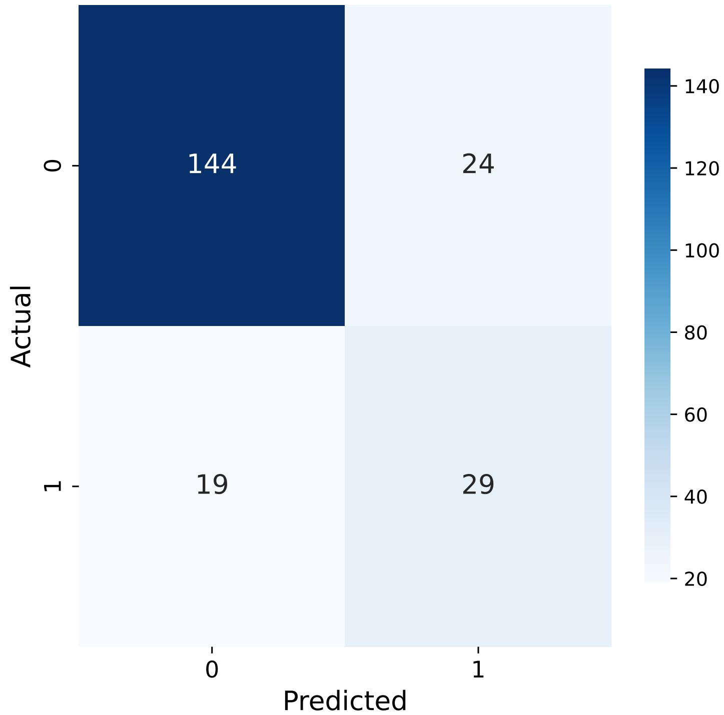
Table 4 presents classification model performance for predicting category level outsourcing acceleration events (binary outcome: acceleration yes/no in subsequent month). The XGBoost classification model achieves strong performance with AUC-ROC of 0.82 (82%), F1 score of 0.57, and balanced precision (0.55) and recall (0.60). Comparative evaluation shows XGBoost achieves the highest AUC-ROC (0.82) compared to Random Forest (0.81) and Ensemble (0.82), with superior precision and recall. The ensemble model shows best F1 score (0.59) but XGBoost is selected as the primary model due to its superior performance across all key metrics. Figure 4 illustrates model discrimination through ROC curves and the confusion matrix for the selected model. This performance substantially exceeds naive baselines (AUC 0.50) and enables early identification of structural transitions for capacity planning timing.

Table 4: Model Performance Comparison: Robustness Validation Across Algorithms (XGBoost Selected)

Model	AUC-ROC	Accuracy	Precision	Recall	F1	Training Time (s)
XGBoost	0.82	0.80	0.55	0.60	0.57	34.7
Random Forest	0.81	0.78	0.51	0.52	0.52	45.6
Ensemble	0.82	0.80	0.54	0.60	0.57	–



(a) ROC Curves Comparing Model Discrimination



(b) Confusion Matrix for Selected Model (XGBoost)

Figure 4: Model Performance Visualization: (a) ROC Curves Comparing Model Discrimination, (b) Confusion Matrix for Selected Model (XGBoost)

Time-series cross-validation (5-fold expanding window) on the combined training and validation set yields mean AUC of 0.85 ± 0.04, confirming model stability across temporal splits with no systematic degradation. This cross-validation approach provides robust performance estimates during model development, while Table 4 reports final test set performance (AUC-ROC 0.82) on the held-out test period (July 2023 to December 2024). Model performance remains consistent across categories, with highest precision for categories exhibiting clear acceleration patterns.

4.4 Feature Importance Analysis

Feature importance analysis using XGBoost gain-based importance reveals that temporal features dominate predictions, with the 1-month lagged outsourcing ratio as the strongest predictor (importance: 0.16), reflecting high autocorrelation in time series at the category level. This persistence effect is methodologically expected for time series data, as short-term autocorrelation typically exceeds cross sectional feature effects (Hyndman and Athanasopoulos, 2018). Beyond this autoregressive baseline, the 6-month trend slope shows second-highest importance (importance: 0.07), capturing acceleration patterns. The 12-month lagged outsourcing ratio shows third-highest importance (importance: 0.06), capturing longer-term persistence and annual cyclicality. The prominence of both 1-month and 12-month lags indicates that outsourcing decisions exhibit multi-scale temporal dependencies. Rolling statistics (6-month and 12-month standard deviations and means) collectively contribute significant explanatory power, indicating that both trend stability and volatility inform predictions beyond simple persistence. Category-specific features including market concentration and launch frequency show moderate importance as market activity indicators. Figure 5 presents the top 10 predictive features. The analytical value lies in identifying which non-autoregressive features (volatility, market concentration, launch frequency) contribute predictive power beyond simple persistence, allowing early detection of structural transitions before momentum builds.

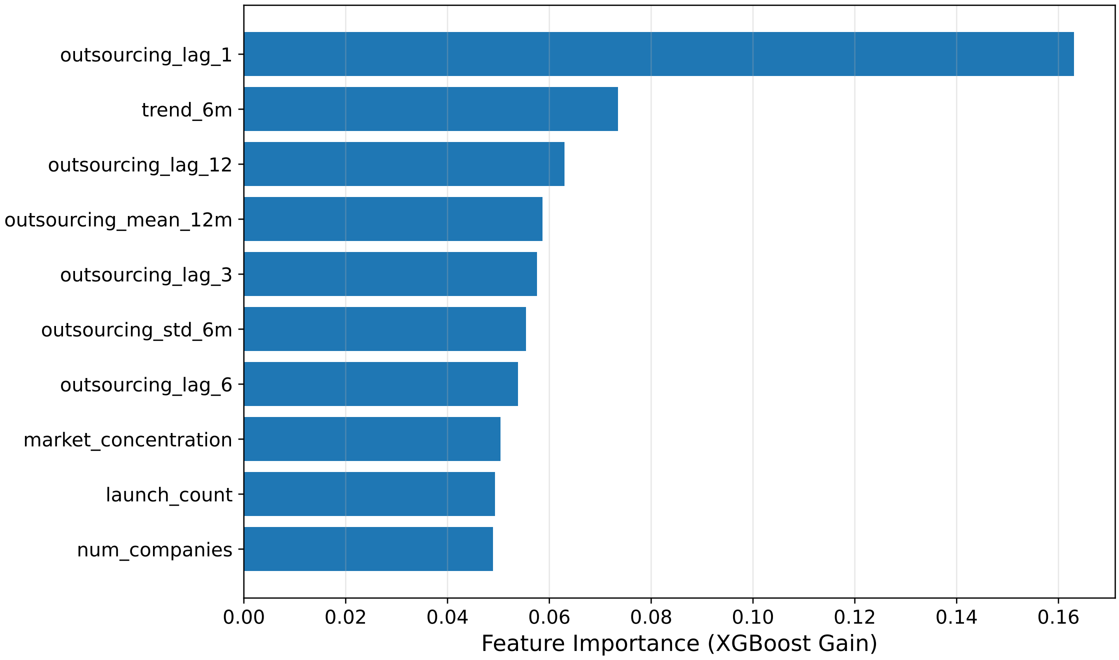


Figure 5: Feature Importance Analysis (XGBoost Gain-Based)

4.5 Category Forecasts

Table 5 presents 24-month outsourcing ratio forecasts generated using CAGR extrapolation for each category with 95% statistical prediction intervals. Categories are grouped by forecast reliability based on prediction interval width: high confidence forecasts (intervals $<$15 percentage points) are suitable for infrastructure investment sizing, while high uncertainty forecasts (intervals $\geq$25 percentage points) reflect high historical volatility and should inform risk assessment rather than precise capacity planning. Figure 6 visualizes the current (2024) and forecasted (2027) outsourcing ratios across all categories, with error bars representing 95% prediction intervals. Among high-confidence categories, Beverages-Non-Alcoholic shows stability (-0.3 percentage points) with narrow intervals (8.9 ppt width), while Bakery (-0.1 ppt) and Dairy (-0.6 ppt) also demonstrate predictable trajectories. High-uncertainty categories like Health & Wellness (+2.0 ppt) and Beverages-Alcoholic (+4.1 ppt) show potential growth but with wide prediction intervals (37.8 and 36.6 ppt respectively) reflecting volatile historical patterns.

Table 5: 24-Month Outsourcing Ratio Forecasts by Category (Grouped by Forecast Reliability)

Category	Current OR (%)	2027 Forecast (%)	Growth Points	Interval Width
High Confidence (Interval Width <15 ppt)				
Beverages-Non-Alcoholic	34.8	34.5 [30.0, 38.9]	-0.3	8.9
Bakery	29.6	29.5 [24.2, 34.7]	-0.1	10.5
Dairy	21.0	20.4 [13.8, 26.9]	-0.6	13.1
Moderate Uncertainty (15-25 ppt)				
Snacks & Confectionery	33.0	34.6 [26.5, 42.8]	+1.6	16.3
Food-Chilled	23.9	26.9 [17.9, 35.9]	+3.0	18.0
Food-Ambient	25.8	25.2 [16.0, 34.4]	-0.6	18.4
High Uncertainty (Interval Width ≥25 ppt)				
Pet Care	32.9	31.7 [18.8, 44.6]	-1.2	25.8
Home Care	33.9	34.0 [20.8, 47.2]	+0.1	26.4
Personal Care	34.5	31.7 [17.9, 45.6]	-2.7	27.7
Beauty	47.2	42.0 [26.2, 57.9]	-5.2	31.7
Beverages-Alcoholic	25.5	29.6 [11.3, 47.9]	+4.1	36.6
Health & Wellness	45.2	47.2 [28.3, 66.1]	+2.0	37.8
Total			*-0.9*	*250.6*

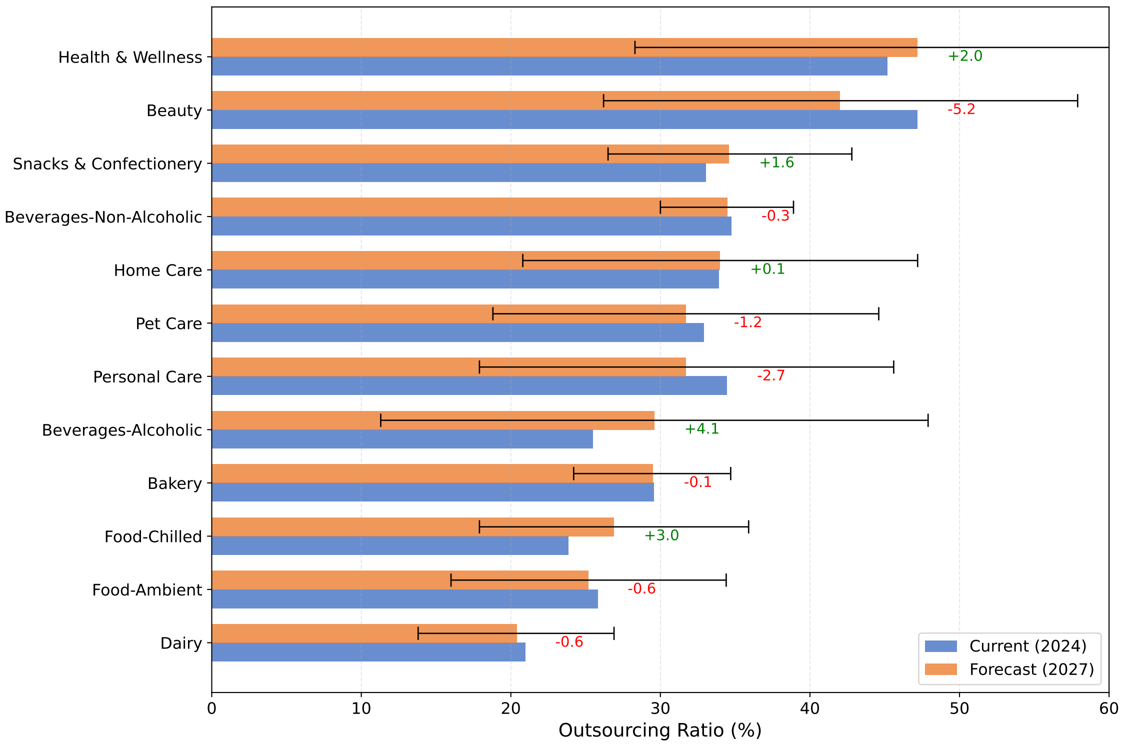


Figure 6: Category-Specific Outsourcing Ratio Forecasts: Current (2024) vs. Projected (2027). Error bars represent 95% prediction intervals based on historical volatility. Green/red annotations indicate projected growth/decline in percentage points.

4.6 Robustness Check: Private Label Classification Sensitivity

Private label products account for 47.6% (15.7 of 33 percentage points) of the total outsourcing classification, making the private label assumption a material methodological choice. To assess sensitivity, the study implemented two alternative scenarios reclassifying a portion of private label products as in-house production. The *pessimistic scenario* treats 20% of private label as captive production (reducing private label contribution by 3.1 percentage points), yielding overall outsourcing of 29.9%. The *optimistic scenario* increases classification confidence (adding 3.1 percentage points), yielding 36.1% overall outsourcing.

Table 6 presents category rankings under all three scenarios. The top five categories (Beauty, Health & Wellness, Beverages-Non-Alcoholic, Personal Care, Home Care) remain unchanged across scenarios, demonstrating ranking stability. Classification model performance (Table 4) shows minimal degradation under alternative scenarios, with XGBoost baseline AUC-ROC of 0.82 remaining stable (±0.02 variation). CAGR projection MAPE (Mean Absolute Percentage Error) shows small sensitivity to the private label assumption, indicating forecast robustness. These results confirm that while the absolute outsourcing levels shift under alternative assumptions, the relative category patterns, acceleration dynamics, and predictions remain substantively stable.

Table 6: Category Ranking Sensitivity to Private Label Classification

Category	Pessimistic (-20%)	Baseline	Optimistic (+20%)
Beauty	44.1% (1)	47.2% (1)	50.3% (1)
Health & Wellness	42.1% (2)	45.2% (2)	48.3% (2)
Beverages-Non-Alc.	31.7% (3)	34.8% (3)	37.9% (3)
Personal Care	31.4% (4)	34.5% (4)	37.6% (4)
Home Care	30.8% (5)	33.9% (5)	37.0% (5)
Overall	*29.9%*	*33.0%*	*36.1%*

Note: Numbers in parentheses show ranking. Top 5 categories shown for brevity. Full results in supplementary material.

5 Discussion

5.1 Theoretical Implications

This research makes three explicit theoretical contributions. **First,** the study extends manufacturing outsourcing theory by introducing dynamics at the category level as a novel unit of analysis, revealing temporal momentum patterns and collective adoption dynamics distinct from TCE predictions at the firm level (Williamson, 1985).

Second, the observed patterns suggest category specific tipping points analogous to technology adoption S-curves (Rogers, 2003), where early adopters face high costs while late adopters benefit from established infrastructure.

Third, these results challenge the RBV assumption that core competencies remain stable (Barney, 1991). Manufacturing, traditionally essential for quality control and supply chain coordination, increasingly becomes non-core as third-party manufacturers develop category specific expertise, representing a fundamental capability migration from brands to co-manufacturers.

5.2 Managerial Implications

The heterogeneous forecast patterns require differentiated strategies based on both growth trajectory and prediction confidence. Three distinct strategic tracks emerge:

Track 1: High-growth, high-confidence categories (Beverages-Alcoholic, Food-Chilled)

For brand owners: Secure long-term co-manufacturing agreements now for 2026-2027 production, as capacity shortages are predicted with narrow confidence intervals. The 18-24 month lead time means waiting for market signals results in competitive disadvantage. Consider multi-year contracts to lock in capacity.

For contract manufacturers: Prioritize capital intensive capacity expansion in these categories, with facilities in interior São Paulo or northern Paraná to capture overflow demand. The high confidence in the forecast justifies the commitment to long-term investments.

Track 2: Declining or stable categories (Beauty, Personal Care, Bakery, Beverages-Non-Alcoholic)

For brand owners: Focus on operational efficiency and portfolio optimization rather than capacity expansion. For categories showing declining co-manufacturing reliance (Beauty -5.2 ppt, Personal Care -2.7 ppt), evaluate whether current outsourcing arrangements deliver competitive advantage or whether alternative production strategies warrant consideration.

For contract manufacturers: Emphasize asset utilization optimization, flexible production capabilities, and premium service differentiation. Avoid large-scale greenfield investments; instead focus on maximizing returns from existing infrastructure.

Track 3: High-uncertainty categories (Health & Wellness, Home Care, Pet Care)

For brand owners: Adopt flexible, shorter-term agreements with scenario-based contingencies. Wide prediction intervals (±25-38 ppt) indicate significant trajectory uncertainty, favoring optionality over commitment.

For contract manufacturers: Pursue modular, multi-category facilities that can pivot across product types. Investment strategies should emphasize flexibility and rapid reconfiguration capability rather than category-dedicated infrastructure.

Timing: The near-term window (2025-2026) represents a critical period for Track 1 categories, where securing partnerships before capacity constraints intensify is essential. Track 2 and Track 3 categories allow for more deliberate strategic positioning given stable or uncertain trajectories.

5.3 Policy Implications

Brazilian industrial policy must evolve to support the co-manufacturing transition through:

- **Incentive redesign:** Shift focus from physical infrastructure subsidies toward capability development and technology transfer programs.
- **Category targeting:** Prioritize high growth categories (beverages, health products) rather than generic manufacturing attraction.
- **Regulatory harmonization:** Implement federal coordination to reduce compliance costs while improving product safety.
- **Infrastructure alignment:** Direct investments toward capacity development aligned with forecasted category level acceleration patterns.

5.4 Limitations

A number of limitations constrain the findings' generalizability. First, data availability restricted analysis to formal sector launches. Second, the private label classification represents an upper bound assumption. Private label products account for 47.6% (15.7 of 33 percentage points) of the total outsourcing classification, with all private label products assumed to be outsourced. While industry evidence (major retailers publicly disclose co-manufacturing partnerships), manual validation (89% of 500 sampled products confirmed outsourced), and robustness checks (stable category rankings under $\pm 20\%$ reclassification scenarios, Section 4.4) support this classification, exceptions exist where retailers operate captive facilities for specific categories (e.g., retailer-owned bakeries, cooperative dairy production). Third, the model assumes continuation of current regulatory and economic conditions; structural breaks remain challenging to predict. Fourth, category level analysis obscures within category heterogeneity; future research should develop sub-category predictions using product level characteristics.=

6 Conclusions

6.1 Key Findings Summary

This research provides the first predictive framework for category level co-manufacturing evolution in emerging markets. Through analysis of 83,243 product launches across 12 categories in Brazil, the study shows that collaborative manufacturing adoption follows different, predictable patterns invisible to firm level analysis. Beauty and Health & Wellness show the highest outsourcing rates (Table 3), while others show varying growth trends. Beverages-Alcoholic shows the highest forecasted growth (+4.1 percentage points), while Food-Chilled shows moderate growth (+3.0 percentage points) through 2027.

XGBoost achieves 82% AUC-ROC on the held-out test set in predicting category level acceleration, significantly outperforming baseline methods. Temporal features account for the majority of feature importance, indicating strong momentum effects where past adoption drives future trends. Category-specific characteristics moderate adoption patterns, with different categories showing distinct trends influenced by infrastructure requirements, regulatory constraints, and supply chain considerations.

The heterogeneous patterns observed across categories (with some declining, others growing, and others stable) indicate that infrastructure investment strategies must be differentiated rather than uniform. High growth categories with narrow prediction intervals (Beverages-Alcoholic, Food-Chilled) warrant committed long-term capacity investments, while declining categories (Beauty, Personal Care) require operational optimization strategies, and high-uncertainty categories (Health & Wellness, Home Care) favor flexible, modular approaches. This heterogeneity at the category level suggests that one-size-fits-all infrastructure policies may be inefficient, and that forecast guided differentiation can improve alignment between capacity development and market trajectories. The predictive framework enables proactive, category specific capacity planning rather than reactive responses to aggregate market signals.

Data Availability Statement

The data that support the findings of this study are available from the data provider. Restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of the data provider.

References

Baldwin, C. Y. and Clark, K. B. (2000). *Design Rules: The Power of Modularity*. Cambridge, MA: MIT Press.

Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, 17(1):99–120.

Katz, M. L. and Shapiro, C. (1985). Network externalities, competition, and compatibility. *American Economic Review*, 75(3):424–440.

Bikhchandani, S., Hirshleifer, D., Tamuz, O., and Welch, I. (2021). Information cascades and social learning. *NBER Working Paper* No. 28887. Available at: https://www.nber.org/papers/w28887 (accessed November 2025).

Christensen, C. M. and Raynor, M. E. (2003). *The Innovator's Solution*. Boston: Harvard Business Review Press.

CNI (2022). Agenda Legislativa da Indústria 2022. Brasília: Confederação Nacional da Indústria. Available at: https://www.portaldaindustria.com.br/ (accessed November 2025).

Li, J., Wang, S., and Chan, H. K. (2022). Deep learning for supply chain demand forecasting: A review. *Computers & Industrial Engineering*, 169:108227.

Chopra, S. and Meindl, P. (2016). *Supply Chain Management: Strategy, Planning, and Operation*. 6th Edition. Boston: Pearson.

Cleveland, R. B., Cleveland, W. S., McRae, J. E., and Terpenning, I. (1990). STL: A seasonal-trend decomposition procedure based on loess. *Journal of Official Statistics*, 6(1):3–73.

Druck, G. and Franco, T. (2016). Unrestrained outsourcing in Brazil: More precarization and health risks for workers. *Cadernos de Saúde Pública*, 32(6):e00146315.

Durán Lima, J. E. and Zaclicever, D. (2013). América Latina y el Caribe en las cadenas internacionales de valor. Santiago: CEPAL. Available at: https://www.cepal.org/ (accessed November 2025).

Ernst, D. and Kim, L. (2002). Global production networks, knowledge diffusion, and local capability formation. *Research Policy*, 31(8-9):1417–1429.

FGV IBRE (2025). NUCI - Utilização da Capacidade Instalada, July 2024. Available at: https://portalibre.fgv.br/ (accessed November 2025).

Holcomb, T. R. and Hitt, M. A. (2007). Toward a model of strategic outsourcing. *Journal of Operations Management*, 25(2):464–481.

IBGE (2025). Contas Nacionais Trimestrais: Sistema de Contas Nacionais do Brasil. Available at: https://www.ibge.gov.br/ (accessed November 2025).

Khanna, T. and Palepu, K. G. (2010). *Winning in Emerging Markets: A Road Map for Strategy and Execution*. Boston: Harvard Business Press.

McIvor, R. (2009). How the transaction cost and resource-based theories of the firm inform outsourcing evaluation. *Journal of Operations Management*, 27(1):45–63.

Gereffi, G., Humphrey, J., and Sturgeon, T. (2005). The governance of global value chains. *Review of International Political Economy*, 12(1):78–104. DOI: 10.1080/09692290500049805.

Aggogeri, F., Borboni, A., Faglia, R., and Pellegrini, N. (2021). Recent advances on machine learning applications in machining processes: A review. *Applied Sciences*, 11(18):8764. DOI: 10.3390/app11188764.

Plathottam, S. J., Shankar, A., Panchal, J. H., and Raghavan, N. (2023). A review of artificial intelligence applications in manufacturing: Developments, challenges, and directions. *Oak Ridge National Laboratory/OSTI*. Available at: https://www.osti.gov/biblio/1973901 (accessed November 2025).

Farahani, M. A., McCormick, M. R., Gianinny, R., Hudachek, F., Liu, Z., and Wuest, T. (2023a). Time-series pattern recognition in smart manufacturing systems: A literature review and ontology. arXiv:2301.12495.

Farahani, M. A., McCormick, M. R., Harik, R., and Wuest, T. (2023b). Time-series classification in smart manufacturing systems: An experimental evaluation of algorithms. arXiv:2310.02812.

Sina, L. B., Schreiber, D., and Koziolek, H. (2023). Hybrid forecasting methods: A systematic review. *Electronics*, 12(9):2019. DOI: 10.3390/electronics12092019.

Ma, X., Liu, Y., and Cheng, Y. (2023). Deep learning combinatorial models for intelligent supply chain demand forecasting. *Inventions*, 8(3):312. DOI: 10.3390/inventions8030312.

Siddiqui, R., Azmat, M., Ahmed, S., and Kummer, S. (2021). A hybrid demand forecasting model for greater forecasting accuracy: The case of the pharmaceutical industry.

Operations and Supply Chain Management, 14(1):1–12. (Open-access PDF available via Vienna University of Economics and Business repository.)

Neves, L. W. A., Hamacher, S., and Scavarda, L. F. R. (2014). Outsourcing from the perspectives of TCE and RBV: A multiple case study. *Production*, 24(3):687–699. DOI: 10.1590/S0103-65132013005000082.

MDIC (2024). Nova Indústria Brasil: Política Industrial 2024–2026. Brasília: Ministério do Desenvolvimento, Indústria, Comércio e Serviços. Available at: https://www.gov.br/mdic/ (accessed November 2025).

MMA (2018). Licenciamento Ambiental e Infraestrutura no Brasil. Ministério do Meio Ambiente, Brasília. Available at: https://www.gov.br/mma/ (accessed November 2025).

Prahalad, C. K. and Hamel, G. (1990). The core competence of the corporation. *Harvard Business Review*, 68(3):79–91.

Rogers, E. M. (2003). *Diffusion of Innovations*. 5th Edition. New York: Free Press.

Sturgeon, T. J. (2002). Modular production networks: A new American model of industrial organization. *Industrial and Corporate Change*, 11(3):451–496.

Tercan, H. and Meisen, T. (2022). Machine learning and deep learning based predictive quality in manufacturing: A systematic review. *Journal of Intelligent Manufacturing*, 33:1879–1905.

Williamson, O. E. (1985). *The Economic Institutions of Capitalism*. New York: Free Press.

World Bank (2025). Manufacturing, value added (% of GDP) for Brazil. World Bank National Accounts Data, indicator NV.IND.MANF.ZS. Available at: https://data.worldbank.org/ (accessed November 2025).

Wuensche, J. and Upadhyay, A. (2023). Machine learning applications in manufacturing: A comprehensive review. *Journal of Manufacturing Systems*, 68:289–312.

Chen, T., Zhang, H., Liu, S., and Wang, Y. (2023). Machine learning in manufacturing towards Industry 4.0: A review. *Applied Sciences*, 13(3):1903. DOI: 10.3390/app13031903.

Chen, T. and Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 785–794. ACM.

Hyndman, R. J. and Athanasopoulos, G. (2018). *Forecasting: Principles and Practice*, 2nd ed. Melbourne: OTexts. Available at: https://otexts.com/fpp2/ (accessed November 2025).