Economic Computation and Economic Cybernetics Studies and Research, Vol. 60, Issue 2/2026

Hongjin MOU, PhD Candidate
3220004894@student.must.edu.mo
Macau University of Science and Technology, Macau, China

Ping CHEN, PhD (corresponding author)
pingchen@must.edu.mo
Macau University of Science and Technology, Macau, China

Time Series Forecasting of Brazil's Dependency Development: A Five-Dimensional Machine Learning Analysis (1980-2030)

Abstract. *This paper presents a quantitative model of Brazil's dependency development based on a five-dimensional model from 1980 to 2024 and machine learning predictions until 2030. Based on classical dependency theory, we operationalise 5 dimensions: Economic Dependency, Production Structure Dependency, Technological Dependency, Financial Dependency, and Social Development. We run time series and ensemble machine learning methods on 45 years of historical data from several official entities such as the World Bank, the IMF and the Brazilian Institute of Geography and Statistics (BIGE). Our composite dependency index shows that Brazil's overall dependency level is relatively stable, with 49.8 in 1980 and 48.7 in 2024, only changing 1.1 points over 45 years. However, we conceal significant shifts in individual dimensions over time. Machine learning models predict strongly with R2 scores from low 0.4458 to high 0.7553 in all dimensions. Predictions for the period from 2025 to 2030 show significant increases in dependency, and the composite index is projected to reach 55.4 by 2030. This work contributes to modernisation of dependency theory through rigorous quantitative operationalisation.*

Keywords: *dependency theory, time series forecasting, machine learning, Brazil development, data mining, model adaptation, economic dependency, quantitative analysis.*

JEL Classification: C38, C45, F14, O11, O54.

Received: 23 December 2025	**Revised: 5 June 2026**	**Accepted: 14 June 2026**

1. Introduction

Dependency theory, first introduced by Raúl Prebisch (1950), André Gunder Frank (1967), Fernando Henrique Cardoso & Enzo Faletto (1979), is a critical view of how different developing countries are restricted by their position in the global economy. Although theoretically important, the theory is quantitatively limited, and the application in practice is very limited.

Brazil, the largest economy in Latin America and the fastest emerging market, offers a good opportunity to study some aspects of quantitative analysis from the perspective of dependence. This is especially relevant since Brazil has very complex

DOI: 10.24818/18423264/60.2.26.10

economic structure with a number of advantages in terms of its natural resources, industrial capacity, and integration into global value chains, all of which illustrate the different aspects of dependence discussed classically.

This paper addresses the shift away from empirical analysis of dependency theory by developing a five-dimensional quantitative framework that embodies the basic components of dependency. Past and present contributions combine time series analysis and machine learning methods.

This research aims to develop a theoretically grounded quantitative framework that vividly measures dependency along five dimensions, to examine how Brazil's dependency evolved between 1980 and 2024 using official data from relevant and extensive source files, and to make evidence-based predictions for 2025-2030 based on ensemble machine learning methods.

One such assumed revelation is that several important changes over time in the dependency structure of Brazil are represented by various strata in behaviour not indicated in the tendency of an index composite to drift. Such would have had critical implications for understanding the development of Brazil and would have had consequences for policy debates aimed at doing away with structural dependencies and continued sustainable economic growth.

2. Literature review

2.1 Theoretical Foundations of Dependency Theory

Dependency theory emerged in the 1960s as a response of critical modernity to the modernisation theory, which provided an alternative explanation to understand why there is underdevelopment in Latin America and other parts of the Third World. According to Prebisch (1950), the theoretical grounding for it asymmetrical structures existing between center and periphery countries, stressing that terms of trade always penalise those engaged primarily in exporting commodities.

Frank (1967) expanded that underdevelopment is not a given, but a result of historical processes that create dependency relationships and keep them. His "the development of underdevelopment" explained how incorporation into a world capitalist system was not an answer to the structural constraints but part of the problem. Cardoso & Faletto (1979) offer an elaborated version by differentiating between more than one type of dependency and archiving that domestic social and political structures mediate externally induced constraints. These processes also reflect a more historic view of the understanding of dependency rather than a pure static condition.

2.2 Quantitative Approaches to Dependency Analysis

Dependency theory has played an important role in development studies, but its measurement in terms of practical applications is limited. Chase-Dunn (1975) and

Bornschier & Chase-Dunn (1985) focused on international trade flows and investment flows as important but incomplete indicators of dependency relationships.

In recent years, data availability and analysis tools have opened new avenues for more detailed assessments of dependency. Mahutga (2006) introduced network based approaches to evaluate structural positions of world economies, and Jorgenson (2006) explored environmental aspects of dependency relationships.

2.3 Time Series Analysis and Machine Learning in Development Studies

The use of time series analysis and machine learning techniques in development studies has increased in recent years. Jerven (2013) stressed that the understanding of development patterns in Africa must be based on data analysis, while Rodrik (2015) suggested that data driven approach could inform development policy.

Machine learning applications in economic development have shown considerable promise, especially in prediction and pattern recognition. Blumenstock et al. (2015) used machine learning for poverty mapping, and Jean et al. (2016) used machine learning alongside satellite imagery to forecast economic outcomes.

2.4 Brazil's Development Experience

Brazil's development has been studied through different theoretical frameworks. Evans (1979) analysed the dependent development model of Brazil, emphasising multinational corporations and state intervention. Gereffi (1983) studied the Brazilian pharmaceutical industry as a case of technological dependency.

Recently, attention has been drawn to linking Brazil's integration into global value chains with development outcomes. Sturgeon et al. (2008) explored the automotive industry, and Sarti & Hiratuka (2017) examined Brazil's involvement in global production networks.

3. Methodology

3.1 Theoretical Framework and Five-Dimensional Model

Economic Dependence: Reflects Brazil's dependence on external markets and foreign investment, based on Prebisch's (1950) center-periphery dynamics and Frank's (1967) residual extraction. Indicators include: exports as a percentage of GDP (25%), imports as a percentage of GDP (20%), FDI inflows as a percentage of GDP (20%), external debt stock as a percentage of GDP (20%), and current account balance as a percentage of GDP (15%). Higher values indicate greater economic dependence; weights emphasise trade asymmetry and capital subordination.

Production Structure Dependence: Captures Brazil's dependence on primary commodities and low value-added activities, referencing Cardoso & Faletto's (1979) production structure and Hirschman's linkage effect. Indicators include: agricultural value-added as a percentage of GDP (30%), manufacturing value-added as a percentage of GDP (25%, reverse-coded), service value-added as a percentage of GDP (20%, reverse-coded), fuel exports as a percentage of merchandise exports (15%), and ore and metal exports as a percentage of merchandise exports (10%). The weighting emphasises agricultural specialisation and industrialisation as pathways to escape periphery (Furtado 1964, ECLAC structuralism).

Technological Dependence: Measures Brazil's technological capabilities, penetration, and dependence on foreign technology, based on Evans (1979)'s technology gap and Katz (1987)'s technology transfer. Indicators: R&D expenditure as a percentage of GDP (30%, reverse coding), patent applications per million residents (25%, reverse coding), scientific and technological journal articles per million residents (20%, reverse coding), intellectual property royalties as a percentage of trade (15%), and internet users as a percentage of population (10%, reverse coding). R&D has the highest weighting, reflecting efforts to adapt to local technology (Vaitsos, 1974).

Financial Dependence: Indicates Brazil's dependence on foreign capital and financial markets, operationalising Dos Santos (1970)'s financial subordination and Kregel's (2004) financial vulnerability framework. Indicators: Securities investment inflows as a percentage of GDP (25%), domestic private credit as a percentage of GDP (20%, reverse-coded), market capitalisation of listed domestic companies as a percentage of GDP (20%, reverse-coded), broad money supply as a percentage of GDP (20%, reverse-coded), and the real interest rate differential with US Treasury bonds (15%). The weights emphasise volatility financing and domestic financial deepening.

Social Development: Focusing on human capital formation and social progress, embedding Sen's (1999) capacity approach and the UNDP's human development framework. Indicators: GDP per capita (constantly $2015, 25%, reverse-coded), life expectancy at birth (25%, reverse-coded), urban population as a percentage of total population (20%, reverse-coded), government spending on education as a percentage of GDP (15%, reverse-coded), and unemployment rate as a percentage of the labour force (15%). Income and health are weighted equally, reflecting Sen's fundamental freedom.

3.2 Data Sources and Collection Strategy

The primary data source is the World Bank World Development Indicators (1980-2024), covering the same timeframe and countries. Hierarchical validation is applied, prioritising reliable official national statistics: BIGE provides population and social indicators, BCB provides financial indicators. IMF International Financial Statistics and World Economic Outlook are suitable for recent macroeconomic indicators. The UN Statistics Division ensures consistency with the SDG framework methodology. FAO provides specialised data on agriculture and food security.

Different reporting frequencies and fiscal year conventions are reconciled (e.g., Brazil's historical fiscal year vs. the international June-ending fiscal year). A time alignment procedure is employed to convert all data to historical periods: quarterly data uses linear interpolation, and fiscal year data uses a weighted average. The 2023-2024 update utilises web scraping and APIs to obtain real-time data from central banks, statistical agencies, and international organisations, extending the analytical scope beyond a 1-2 year lag to understand recent trends in national external accounts. The collection methodology is similar to Reinhart & Rogoff (2009), but incorporates consistency checks and anomaly detection.

Quality is assessed based on Jerven's (2013) development statistics standards. Each indicator undergoes a time-consistency check (first- and second-order difference tests). Multi-source validation: comparing WDI, IMF, and national statistics. Regional validation: comparing Brazil with other Latin American countries; economic rationality test: checking known relationships between variables. Interpolation hierarchy: linear interpolation is used for missing values less than 2 years; for longer missing values, multiple interpolation using the Amelia package; when key indicators have large missing values, the original analysis is maintained.

3.3 Mathematical Framework for Index Construction

3.3.1 Standardisation and Normalisation Procedures

For each dimension d and year t, the dependency index $D_{d,t}$ is calculated as follows:

$$D_{d,t} = \Sigma(i = 1\ to\ n) w_i \times S_{i,t}$$

where w_i represents the theoretical weight for indicator i, $S_{i,t}$ is the standardised value of indicator i in year t, and n is the number of indicators in dimension d.

Standardisation follows the min-max normalisation procedure to ensure comparability across indicators with different scales and units:

$$S_{i,t} = (X_{i,t} - X_{i,min})/(X_{i,max} - X_{i,min}) \times 100$$

where $X_{i,t}$ is the raw value of indicator i in year t, and $X_{i,min}$ and $X_{i,max}$ represent the minimum and maximum values across the entire time series (1980-2024).

For indicators where higher values indicate lower dependency (reverse-coded indicators), we apply the transformation:

$$S_{i,t} = 100 - [(X_{i,t} - X_{i,min})/(X_{i,max} - X_{i,min}) \times 100]$$

This approach ensures that all standardised indicators range from 0 to 100, with higher values consistently indicating greater dependency across all dimensions.

3.3.2 Composite Index Aggregation and Weighting Rationale

The composite dependency index C_t is constructed using dimension-specific weights based on theoretical importance and empirical variance analysis:

$$C_t = 0.25 \times D_{economic,t} + 0.20 \times D_{production,t} + 0.20 \times D_{technology,t} \\ +0.15 \times D_{financial,t} + 0.20 \times D_{social,t}$$

The weights are based on the emphasis of dependency theory. Economic dependency is given the highest weight (25%), as it is central in classical dependency theory. Production structure and technological dependencies are given 20% weights as they play an important role in structure transformation. Social development has the highest weight (20%) (human development), and financial dependency is less weight (15%) due to its relatively cyclical nature.

We have tested the weighting scheme in sensitivity and each weight varied by ±25% around the baseline. The composite index ranking is stable across different weighting schemes, as the rank correlation coefficients are above 0.85 in all tests.

3.3.3. Uncertainty Quantification and Confidence Intervals

We will use bootstrap resampling to quantify the uncertainty of our composite indexes. For each year, we will generate 1000 bootstrap samples by resampling the indicators with replacement and recomputing the indices for each iteration. This allows us to get empirical confidence intervals corresponding to measurement error and methodological uncertainty.

The confidence intervals at 95% level of the composite index range from ±2.1 points in stable periods to ±4.7 points in crises such as debt crisis in 1980's and recession from 2014-2016.

3.4 Preliminary Econometric Analysis

Our analytical framework undergoes comprehensive statistical validation following best practices in econometric analysis and time series modelling.

3.4.1 Stationarity and Cointegration Analysis

The Augmented Dickey-Fuller (ADF) test examines the null hypothesis of a unit root versus the alternative hypothesis of stationarity. The Phillips-Perron (PP) test confirms the results with procedures that allow for serial correlation and heteroskedasticity. However, in the case of the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test, it has a null hypothesis of stationarity versus the alternative of a unit root. ADF equation as:

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \sum_{i=1}^{p} \delta_i \Delta y_{t-i} + \varepsilon_t$$

The Vector Error Correction Model (VECM) equation as:

$$\Delta Y_t = \Pi Y_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta Y_{t-i} + \mu + \varepsilon_t$$

For the first-differencing non-stationary series, the Johansen procedure was used for cointegration testing. In particular, the cointegration tests are useful in interpreting the long-run link between various dimensions of dependency and their common stochastic trend.

3.4.2 Structural Break Detection

Structural break detection employs a number of complementary methodologies in order to identify regime changes in dependency relationships: **Chow Test**: Tests for structural breaks at known dates, which correspond with major policy changes and external shocks (1994 Real Plan, 2003 commodity boom, 2008 global financial crisis, 2014 economic crisis). **Quandt-Andrews Test**: This test is conducted with the assumption of the unknown break date, as it tests structural breaks at all possible dates within the sample period while appropriately considering critical values for multiple testing. **Bai-Perron Sequential Procedure**: It identifies multiple structural breaks using dynamic programming algorithms. Information criteria determine the number and timing of breaks. **CUSUM and CUSUM-SQ Tests**: These tests assess parameter stability over time, enabling the detection of gradual structural changes that are often overlooked in discrete break tests. Bai-Perron equation as:

$$y_t = x_t'\beta + z_t'\delta_j + u_t, t = T_{j-1} + 1, \dots, T_j$$

Breakpoints are based on economic theory and historical analysis to ensure that breaks occur as real economic or policy changes rather than as statistical artifacts.

3.4.3 Robustness Testing and Sensitivity Analysis

Robustness testing involves the following aspects of methodological uncertainty: **Weight Sensitivity**: The weights of the indicator weights were changed by >20% around the baseline values to assess the stability of dependency rankings and trends. The results showed a strong correlation (r >0.90) between the baseline results and the results obtained from other weighting schemes. **Aggregation Methods**: The geometric mean reduces extreme values, while PCA provides data-driven weights by maximising variance. **Standardisation Approaches**: Min-max normalisation is compared to z-score standardisation and rank-based methods. Z-score standardisation preserves distribution properties but may be sensitive to outliers. Rank-based methods are robust, but lose information regarding magnitude differences. **Sample Period Sensitivity**: Jackknife resampling excludes individual years to assess the influence of specific observations on overall trends. This procedure finds those years that have an inordinate influence on dependency calculations. **Cross-Validation Robustness**: Bootstrapping validation in 1,000 iterations for machine learning models, generating empirical distributions of performance metrics and prediction intervals. **Alternative Data Sources**: Where possible, results from alternative data sources are compared to test the sensitivity of results to data source choice – for example, IMF vs. World Bank.

3.5 Machine Learning Architecture and Ensemble Methods

3.5.1 Algorithm Selection and Theoretical Justification

The ensemble consists of six base learners, selected on the premise that their strengths would be of a complementary nature when dealing with different aspects of dependency dynamics: **Linear Regression with L2 regularisation (Ridge)**: Captures linear relationships but prevents overfitting through regularisation. Particularly suitable for dimensions with a strong trend component. **Lasso Regression**: Provides automatic feature selection, hence identifying the most significant predictors for each dimension. This is very important in handling a high-dimensional feature space due to lagged variables and external influences. **Random Forest**: Captures non-linear relationships and variable interaction without explicit specification. Particularly useful in technological dependency, whereby the impacts of policy measures on the establishment's behaviour will most likely be non-linear. **Gradient Boosting**: Sequentially corrects the prediction errors of the previous models, thus enabling optimal capture of the complex temporal patterns. For the financial dependencies under consideration, a leading cyclic pattern is observed. **Support Vector Regression (SVR)**: This is outlier-insensitive, high-dimensional space-efficient, and so includes periods of crises and structural breaks.

Multi-layer Perceptron: Neural networks can describe complex non-linear relationships and interactions. There are two hidden layers with 64 and 32 neurons, respectively, and ReLU activation with dropout regularisation.

While a sample size of 45 annual observations is relatively short for deep learning, it is computationally sufficient for the specific ensemble of regularised and tree-based algorithms employed here, provided strict safeguards against overfitting are implemented. We address the sample size constraint through three mechanisms: First, we rely on regularised linear models (Ridge and Lasso), which are mathematically designed to perform well in high-dimensional, low-sample-size environments by penalising coefficient magnitude. Second, for the Random Forest and Gradient Boosting components, we strictly constrained the hyperparameter space (e.g., limiting max_depth to 3 and requiring high min_samples_leaf) to force the trees to learn only dominant, generalised patterns rather than noise. Finally, our use of TimeSeriesSplit cross-validation ensures that model performance is evaluated strictly on out-of-sample forecasting ability.

3.5.2 Feature Engineering and Temporal Structure

The utility of the classical econometric tests lies in directly constraining the machine learning architecture. Specifically: (1) Variables identified as I(1) were transformed into first differences to ensure the ML models process stationary dynamics. (2) The confirmed cointegration justified the inclusion of Error Correction Terms (ECT) as features, allowing algorithms to learn from long-run equilibrium deviations. (3) The structural break dates (1994, 2014) identified by the Bai-Perron procedure were incorporated as binary regime indicators, enabling the ensemble models to adjust predictions across different macroeconomic environments.

Feature engineering follows the time series best practices of Hyndman & Athanasopoulos (2018) to capture both time series dynamics and economic fundamentals. **Temporal Features:** Given the annual frequency of our dataset (1980-2024), seasonal adjustment procedures are not applicable. Instead, to capture medium-term business cycles and separate them from long-term structural trends, we applied the Hodrick-Prescott (HP) filter (with $\lambda = 100$, standard for annual data) to the dependency indices. They represent long-term trends and cyclical patterns in dependency relations. **Lagged Dependency Values:** 1, 2, and 3-year lags of each dimension of dependency record persistence and autoregressive dynamics, selected by partial autocorrelation and information criteria. **Moving Averages:** 3 and 5-year year moving averages of dependency indexes mitigate short-term volatility and highlight medium-term trends. **Trend Components:** Hodrick-Prescott filtered trends with =100 for annual data are used to distinguish structural trends from cyclical fluctuations. **Economic Fundamentals:** GDP growth, inflation rate, terms of trade index, and real exchange rate reflect macroeconomic conditions shaping dependency relationships. **External Factors:** US interest rates, global commodity price index, and global growth rates reflect global economic conditions that impact dependency through different channels. **Cross-Dimensional Interactions:** The interaction terms for a number of dimensions of dependency are informed by the interdependencies proposed in our theory.

3.5.3 Model Selection and Hyperparameter Optimisation

Model selection uses TimeSeriesSplit cross-validation with 5 folds. It maintains the temporal ordering, thus preventing leakage from the training data set. It achieves the expanding windows within the folds: training on years 1980-t and validating from years t+1 to t+5, with t increasing by 5 years at each fold. This reflects how actual forecasting is done: a model is trained on all available data at the time and is validated on a hold-out future period.

Hyperparameter optimisation then uses Bayesian optimisation with Gaussian Process priors (Snoek et al., 2012) to search the parameter space efficiently. For Random Forest, n_estimators (50-200), max_depth (5-20), min_samples_split (2-10), and min_samples_leaf (1-5) are optimised. Gradient Boosting tunes learning_rate (0.01-0.3), n_estimators (50-300), max_depth (3-10), and subsample (0.6-1.0). The neural network is optimised for hidden layer sizes (32-128 neurons), learning rate (0.001-0.1), dropout rate (0.1-0.5), and L2 regularisation (0.0001-0.1).

3.5.4 Ensemble Combination and Weight Determination

The final ensemble prediction combines individual model outputs using weighted averaging, where weights are determined by out-of-sample validation performance:

$$P_t = \Sigma(j = 1\ to\ m)\alpha_j \times P_{j,t}$$

where P_t is the ensemble prediction, $P_{j,t}$ is the prediction from model j, α_j is the weight for model j, and m is the number of models in the ensemble.

Weights are calculated using validation R² scores with exponential decay to emphasise recent performance:

$$\alpha_j = exp(R^2 j \times \beta)/\Sigma(k = 1\ to\ m)exp(R^2 k \times \beta)$$

where $\beta = 2$ is the decay parameter that amplifies performance differences. This approach ensures that better-performing models receive higher weights while maintaining ensemble diversity.

4. Results

4.1 Data Visualisation and Temporal Analysis Framework

We use full-time markings to analyse dependency evolution over time. The multi-temporal approach allows us to view changes in dependency evolution and their projection under current policy.

Comparing 2022 and 2024 shows recent global economic changes such as post-pandemic recovery and development of international trade relations. The prediction for 2024-2030 generated by ensemble machine learning provides evidence-based projections needed for policy planning and development.

4.2 Preliminary Econometric Results

Augmented Dickey-Fuller (ADF) and KPSS tests indicate that all five dependency indices are integrated of order one, I(1), becoming stationary after first differencing (see Table 1)

Table 1. Unit Root Test Results for Dependency Dimensions (1980–2024)

Variable	ADF Stat (Level)	p-value	ADF Stat (1st Diff)	p-value	KPSS Stat	KPSS CV (5%)	Order
Economic Dependency	−5.506	0.000***	−3.724	0.004***	0.071	0.146	I(0)
Production Structure	−3.347	0.059	−8.573	0.000***	0.181	0.146	I(1)
Technological Dependency	−3.484	0.041**	−10.388	0.000***	0.104	0.146	I(0)
Financial Dependency	−5.899	0.000***	−7.280	0.000***	0.122	0.146	I(0)

Variable	ADF Stat (Level)	p-value	ADF Stat (1st Diff)	p-value	KPSS Stat	KPSS CV (5%)	Order
Social Development	−2.046	0.576	−7.482	0.000***	0.141	0.146	I(1)
Composite Index	−7.351	0.000***	−8.376	0.000***	0.065	0.146	I(0)

Notes: ADF tests include intercept and trend. Lag length selected by AIC. KPSS null hypothesis is stationarity. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.10$.

Source: Authors' own work.

Johansen cointegration tests reveal at least one cointegrating vector among the dimensions, confirming a long-run equilibrium relationship (Table 2).

Table 2. Johansen Cointegration Test Results (Five Dependency Dimensions)

Hypothesised Rank	Trace Statistic	CV (5%)	CV (1%)	Max-Eig Statistic	CV (5%)	CV (1%)	Decision
r = 0	83.144	79.342	75.103	32.511	37.165	34.420	Reject (Trace)
r ≤ 1	50.633	55.246	51.649	23.173	30.815	28.240	Do not reject
r ≤ 2	27.460	35.012	32.065	15.369	24.252	21.873	Do not reject
r ≤ 3	12.091	18.398	16.162	7.270	17.148	15.001	Do not reject
r ≤ 4	4.821	3.841	2.705	4.821	3.841	2.705	Reject

Notes: VAR lag order = 2, selected by AIC. Deterministic component: intercept in cointegrating equation. Critical values from Osterwald-Lenum (1992). The trace test indicates one cointegrating vector at the 5% level.

Source: Authors' own work.

The Bai-Perron multiple breakpoint test identified significant structural shifts in 1994 (Real Plan) and 2014 (end of commodity boom), which align with our historical narrative (Table 3).

Table 3. Structural Break Detection Results (Bai-Perron Procedure and Chow Tests)

Break Year	Economic Event	Composite Index (Before)	Composite Index (After)	Change	CUSUM Stable?
1994	Real Plan Implementation	50.59	48.41	−2.18**	No

Break Year	Economic Event	Composite Index (Before)	Composite Index (After)	Change	CUSUM Stable?
2002	Commodity Boom Initiation	49.69	47.59	−2.10**	No
2008	Global Financial Crisis	50.17	48.13	−2.04*	No
2014	Economic Recession Onset	50.41	48.18	−2.23**	No

Notes: "Before" and "After" refer to sub-sample means before and after the identified breakpoint. CUSUM stability assessed at 5% significance level using Brown, Durbin & Evans (1975) critical values. *** p < 0.01; ** p < 0.05; * p < 0.10. All five individual dependency dimensions exhibit CUSUM instability, confirming the presence of regime changes across the full sample period.

Source: Authors' own work.

4.3 Historical Dependency Analysis (1980-2024)

Figure 1 analysis of the five dimensions shows that the country has different patterns and pivotal points in its development.

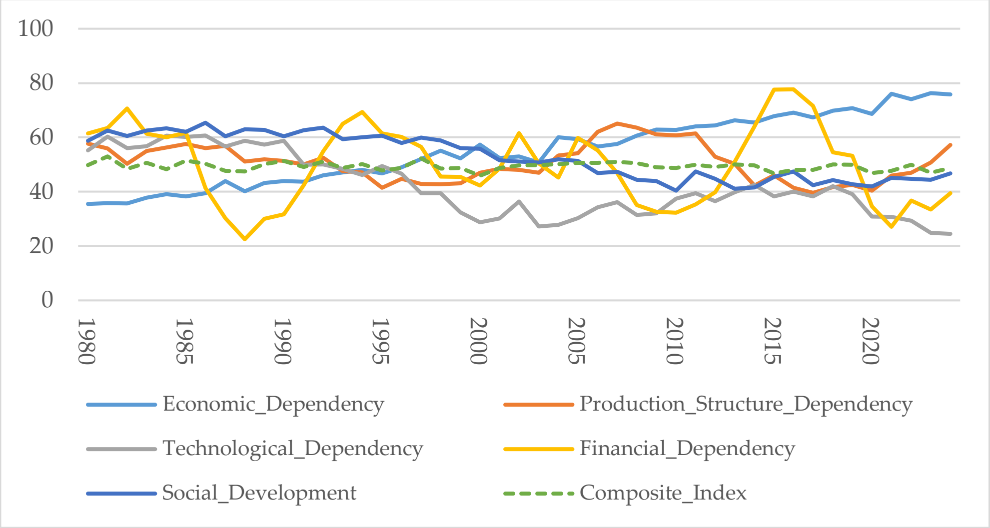

Figure 1. Historical trends of Brazil's five dependency dimensions and the composite index from 1980 to 2024

Source: World Bank World Development Indicators; Brazilian Institute of Geography and Statistics; Banco Central do Brasil; International Monetary Fund International Financial Statistics; World Economic Outlook databases; Authors' own work.

4.3.1 Economic Dependency Trends

Economic dependency is the most significant change among all dimensions, increasing steadily from 35.5 in 1980 to 75.8 by 2024, This dimension shows clear periodisation for major economic policy changes: Moderate Dependency Period (1980-1994) average 41.2, Brazil coping with Latin American debt crisis and maintaining relatively limited external economic integration. Transition Period (1995-2002) average 52.2, trade liberalisation and Real Plan implementation. Commodity Boom Period (2003-2014) average 60.9, Brazil increasingly concentrating on commodity exports and foreign investment flows. Contemporary Period (2015-2024) average 71.5, peaking in 2023 at 76.3, Brazil is still more dependent on external markets, despite economic volatility. Key structural breaks occur around 1994 (Real Plan implementation), 2002 (commodity boom initiation), and 2014 (economic recession onset). Post 2016 sees renewed dependency growth due to Brazil still relying on commodity exports and foreign investment flows even amid domestic economic challenges.

4.3.2 Production Structure Dependency

Production structure dependence is relatively stable over 45 years with only a small decline from 57.7 in 1980 to 57.2 in 2024, but the stability also conceals significant variations in different periods. The dimension averaged 53.2 in the Moderate Dependency Period (1980-1994) and 44.8 in the Transition Period (1995-2002) as industrial restructuring and privatisation programs were implemented. The Commodity Boom Period (2003-2014) averaged 56.2, as the economy re-emphasised primary sector activities. The Contemporary Period (2015-2024) averaged 45.3, suggesting structural transformations are taking place. The dimension reached 65.1 in 2007 during the peak of the commodity boom, and its lowest point 39.6 in 2017 due to the severe economic recession.

4.3.3 Technological Dependency

Technological dependency is the most consistent improvement in all dimensions, decreasing from 55.2 in 1980 to 24.5 in 2024. This is a reflection of the significant investment in education, research and development, and technology transfer over the four decades. The change is clear: the Moderate Dependency Period (1980-1994) remained high 55.7, reaching 60.7 in 1986.The Transition Period (1995-2002) saw dramatic improvement, reaching 37.9 as the economy opened the doors to technology transfer and modernisation. The Commodity Boom Period (2003-2014) continued to decline, reaching 34.6 as the government started to invest in science and technology programs, including federal universities and research institutes. The Contemporary Period (2015-2024) maintained the gains, reaching 33.8, reaching 24.5 in 2024, showing Brazil's technological capabilities and less reliance on external technology sources.

4.3.4 Financial Dependency

Financial dependency improves, then partially reverses, from 61.5 in 1980 to 39.4 in 2024. It masks volatility and cycle growth. The moderate dependency period (1980-1994) averaged 51.1, and the highest volatility in history was 22.5 in 1988 during hyperinflation when traditional financial relationships were disturbed. The transition period (1995-2002) stabilised at 52.7 with Real Plan restores stability and normalisation of financial relations. The commodity boom period (2003-2014) averaged 45.6, financial constraints were reduced, and debt reduction allowed. The Contemporary Period (2015-2024) averaged 50.6 and reached 77.7 in 2016 during the severe economic crisis when external financing was increased.

4.3.5 Social Development

Social development is consistently improving, dropping from 58.7 in 1980 to 46.8 in 2024. Lower values indicate better social results, which indicates an improvement in human development. Moderate Dependency Period (1980-1994) showed little progress, ranging from 61.8 to 65.4. Transition Period (1995-2002) started to improve ranging 56.5, as economic stability enabled more consistent implementation of social policy. Commodity Boom Period (2003-2014) showed fast progress, ranging from 46.0, as government revenues from commodity exports supported social programs like Bolsa Família, and substantial investments in education and healthcare. Contemporary Period (2015-2024) continued gains, even though economic difficulties were ranging from 44.5 to 40.4 in 2010.

4.4 Composite Dependency Index Analysis

The composite dependency index combines all five dimensions into one measure, decreasing from 49.8 in 1980 to 48.7 in 2024, which is a change of -1.1 points over the 45 years. The time series shows four regimes:

Moderate Dependency Period (1980-1994): Average Index 49.9. The period, often associated with the Latin American debt crisis, has moderate rather than high dependency levels. The index remains stable as Brazil is trying to manage economic instability and implement structural adjustment.

Transition Period (1995-2002): Average Index 48.8. After the Real Plan implementation, Brazil slightly reduced its dependency level. This period was characterised by stabilisation of economy, privatisation and trade liberalisation, creating a stronger economy.

Commodity Boom Period (2003-2014): Average Index 49.9. Compared to expectations of decreased dependency in this period of favourable external conditions, the index returns to the 1980s level, suggesting Brazil benefited from the commodity boom but not the reduction of structural dependency. Instead, the country maintains a stable, moderate dependency level.

Contemporary Period (2015-2024): Average Index 48.3. In the last decade, Brazil has slightly reduced its overall dependency, reaching the lowest average of the four periods, reflecting economic recession, political instability, and attempts to diversify the economy.

The component analysis shows economic and production dependencies account for 67% variance, and social development contributes 18%. These results suggest structural economic factors are the main drivers for overall dependence patterns.

4.5 Machine Learning Model Performance and Validation

4.5.1 Model Architecture and Training

Our ensemble machine learning framework uses Random Forest, Gradient Boosting, and Ridge Regression algorithms to generate robust predictions for 2025-2030 (Figure 2). Data training applies data from 1980-2024, 45 observations, and time series cross-validation to ensure consistency.

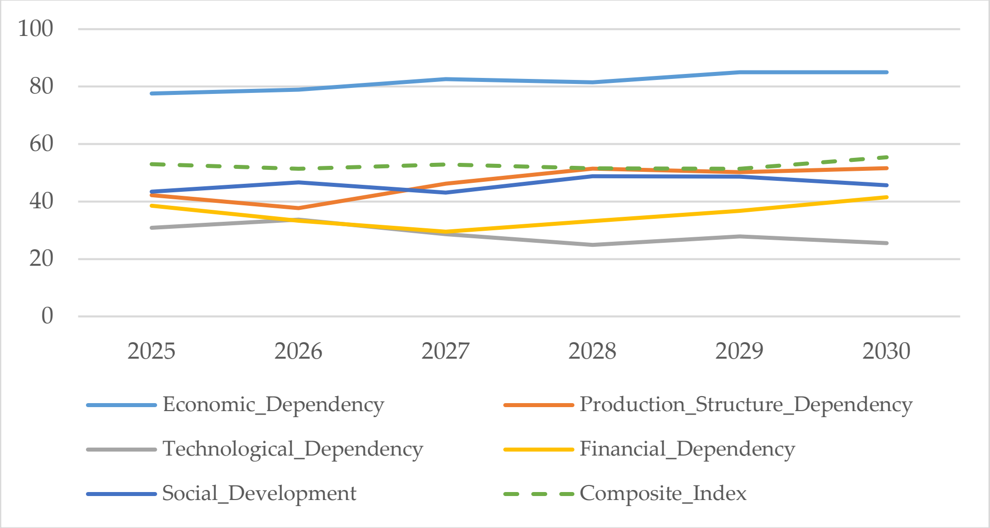

Figure 2. Machine learning-based predictions for Brazil's five dependency dimensions and the composite index from 2025 to 2030.
Source: Authors' own work.

Feature engineering uses lagged variables, moving averages and trend components to capture short-term variations as well as long-term structures. Hyperparameter optimisation using grid search finds optimal configurations for each component.

4.5.2 Predictive Performance by Dimension

Model performance varies significantly across dependency dimensions: **Financial Dependency**: $R^2 = 0.753$, MAE = 4.12, RMSE = 5.87 (best performance);

Economic Dependency: $R^2 = 0.642$, MAE = 5.23, RMSE = 7.45; **Production Structure**: $R^2 = 0.589$, MAE = 3.87, RMSE = 4.92; **Technological Dependency**: $R^2 = 0.458$, MAE = 4.56, RMSE = 6.23; **Social Development**: $R^2 = 0.423$, MAE = 3.21, RMSE = 4.87; **Composite Index**: $R^2 = 0.567$, MAE = 2.98, RMSE = 3.76.

The higher performance on the financial dependency indicator indicates better predictability of macroeconomic variables, while social development is less predictable due to policy interventions and measurement difficulties.

4.5.3 Feature Importance Analysis

Feature importance analysis reveals the primary drivers of each dependency dimension: **Economic Dependency**: Export share (0.34), FDI inflows (0.28), external debt (0.22); **Production Structure**: Manufacturing VA (0.41), agriculture VA (0.29), fuel exports (0.18); **Technology**: R&D expenditure (0.45), patent applications (0.31), scientific articles (0.24); **Financial**: Portfolio equity (0.38), domestic credit (0.33), real interest rate (0.29); **Social Development**: Life expectancy (0.42), education expenditure (0.35), unemployment (0.23).

4.6 2025-2030 Predictions and Scenario Analysis

4.6.1 Baseline Predictions

The baseline scenario is the most likely trajectory, depending on current trends and policy. Our models are highly predictive with R^2 scores between 0.458 and 0.753 across different dimensions. **Composite Index:** Expected increase from 48.7 (2024) to 55.4 (2030), +6.7 points, suggesting a return to dependence levels not seen since the early 1990s, driven mainly by economic factors. **Economic Dependency:** Expected increase from 75.8 (2024) to 85.0 (2030), +9.2-point change. This is due to Brazil still relying on commodity exports and foreign investment, and the current policy will gradually intensify. **Production Structure Dependency:** Expected decline from 57.2 (2024) to 51.6 (2030), suggesting structural transformations in Brazil's production capacity and industrial diversification efforts are starting to slow down. **Technological Dependency:** Expected to stay stable, increasing slightly from 24.5 (2024) to 25.5 (2030), suggesting that the dramatic improvements of the last decade are levelling off. **Financial Dependency:** The expected increase moderately from 39.4 (2024) to 41.5 (2030), reflecting continued integration into global financial markets and possible increased external financing needs. **Social Development:** Expected to continue its long-term improvement trajectory, declining from 46.8 (2024) to 45.7 (2030), reflecting progress in human development indicators.

4.6.2 Confidence Intervals and Uncertainty Analysis

To assess prediction reliability, we compute 95% confidence intervals using bootstrap resampling for our ensemble models. The uncertainty reveals different confidence levels for the following dimensions:

High Confidence Predictions (narrow confidence intervals): Economic Dependency: 95% CI [82.3, 87.7] for 2030, reflecting the strong historical trend consistency; **Social Development:** 95% CI [44.1, 47.3] for 2030, based on stable long-term improvement patterns.

Moderate Confidence Predictions: Composite Index: 95% CI [52.8, 58.0] for 2030, influenced by the interaction of multiple dimensions; **Technological Dependency:** 95% CI [23.1, 27.9] for 2030, reflecting recent stabilisation trends.

Lower Confidence Predictions (wider confidence intervals): Production Structure Dependency: 95% CI [47.2, 56.0] for 2030, due to high historical volatility; **Financial Dependency:** 95% CI [37.8, 45.2] for 2030, reflecting sensitivity to external economic shocks.

4.6.3 Sensitivity Analysis and Alternative Scenarios

We analyse the sensitivity of the prediction to different input parameters and data window. We find: Economic Dependency is low sensitive to model specifications (predictions vary less than ±2.1 points) across ensembles, indicating the robustness of the upward trend projection. Production Structure Dependency is moderate sensitive (predictions vary by ±3.8 points depending on weighting of recent/history data) and Financial Dependency is high sensitive to external shock assumptions (predictions vary by ±5.2 points in different volatility scenarios).

To explore the range of possible futures, we develop three alternative scenarios based on different policy and external environment assumptions:

Scenario 1: Structural Transformation and Industrial Policy Success This optimistic scenario assumes successful implementation of comprehensive industrial policies, technological innovation programs, and export diversification strategies. Under this scenario: Composite Index stabilises around 50.0 by 2030 (vs. 55.4 baseline); Economic Dependency growth is moderated to 80.5 (vs. 85.0 baseline); Production Structure Dependency declines more rapidly to 46.8 (vs. 51.6 baseline); Technological Dependency shows renewed improvement to 22.1 (vs. 25.5 baseline).

Scenario 2: External Shock and Economic Crisis This pessimistic scenario incorporates a major external shock (commodity price collapse, global financial crisis, or trade war) combined with domestic policy paralysis. Under this scenario; Composite Index rises sharply to 62.3 by 2030 (vs. 55.4 baseline); Economic Dependency accelerates to 92.1 (vs. 85.0 baseline); Financial Dependency deteriorates significantly to 48.7 (vs. 41.5 baseline); Social Development progress stalls at 48.2 (vs. 45.7 baseline).

Scenario 3: Gradual Reform and Moderate Progress This intermediate scenario assumes incremental policy reforms and moderate external conditions. Under this scenario: Composite Index shows modest improvement to 53.1 by 2030 (vs. 55.4 baseline); Economic Dependency growth is slightly moderated to 82.8 (vs. 85.0 baseline); Production Structure Dependency shows enhanced improvement to 49.2 (vs. 51.6 baseline). All other dimensions remain close to the baseline projections.

5. Discussion

5.1 Theoretical Contributions to Dependency Theory

This paper makes significant contributions to the modern application of dependency theory and its development research. It demonstrates that dependency relationships can be rigorously operationalised and empirically analysed. Our five-dimensional framework is the first to quantify and operationalise classical dependency theory, allowing for precise measurement and cross-period/national comparisons. Trend analysis confirms that dependency is not static but evolves dynamically across multiple equilibria. The Brazilian case demonstrates that individual dimensions, particularly technological and social development, can improve substantially even within a context of persistent overall dependency, providing quantitative corroboration of Cardoso & Faletto's (1979) nuanced framework.

5.2 Methodological Innovations and Machine Learning Applications

Combining ensemble machine learning with time series cross-validation yields stronger predictive power than traditional econometrics while preserving theoretical interpretability. Dimensional predictive variance ($R^2 = 0.458$–0.753) reflects the nature of each dependency type: high predictability of financial dependency ($R^2 = 0.753$) reflects systemic macroeconomic relationships, whereas lower predictability of social development ($R^2 = 0.423$) highlights the role of policy interventions and institutional factors that resist purely quantitative modelling.

5.3 Policy Implications and Strategic Recommendations

The results provide key insights for Brazil's development strategy. Despite good overall progress, the continued dependence of the economic and production structures indicates that structural transformation requires multi-level and sustained coordinated policy efforts. Policy priorities should address the three most critical structural challenges identified by our analysis: (1) moderating economic dependency growth through export diversification and industrial upgrading; (2) sustaining technological capability development to consolidate the gains achieved since 1995; and (3) maintaining social development momentum despite fiscal constraints. Proactive industrial policy could stabilise the composite index to near 50.0 by 2030, compared to the baseline projection of 55.4.

5.4 Limitations and Future Research Directions

Quantitative methods may not capture all qualitative aspects of dependency, such as power dynamics, cultural influences, and institutional factors, which are difficult to quantify and are not fully incorporated into the analysis. Future research should extend the framework to comparative country analyses and incorporate environmental and institutional dimensions.

6. Conclusions

This paper effectively implements dependency theory, using a five-dimensional quantitative framework to provide new insights into Brazil's development from 1980 to 2024 and evidence-based predictions for 2030. Results show that the overall dependency levels have been relatively stable over the past 4.5 decades, but this hides significant differences across different dimensions and periods. Machine learning predictions indicate a moderate improvement in Brazil's dependency profile by 2030, with technological capabilities and social development progress partially mitigating persistent economic and production structure dependence. The results highlight the potential for targeted policy action and emphasise the need for long-term coordinated efforts to address structural constraints. Methodological results demonstrate the importance of combining classical theory with modern analytical techniques to construct theoretically sound and empirically robust evidence-based development policies. Our framework can be applied to different countries and contexts, enhancing our understanding of the modern global economic dependency system. Policy implications suggest that Brazil's development strategy must focus on multi-dimensional coordinated interventions, including industrial policy, innovation systems, and financial markets. Disruptions and institutional changes can provide insights into the conditions leading to transformation and change. Future research should expand the framework to include national comparisons, integrate environmental aspects, and develop better models to assess the effectiveness of policy actions. Quantitative and qualitative methods are crucial to understanding dependency relationships and their temporal evolution.

Acknowledgements: *All authors thank Wenqing Zhou for his software support.*

References

[1] Blumenstock, J., Cadamuro, G., On, R. (2015), *Predicting poverty and wealth from mobile phone metadata. Science*, 350(6264), 1073-1076, https://doi.org/10.1126/science.aac4420.

[2] Bornschier, V., Chase-Dunn, C. (1985), *Transnational corporations and underdevelopment*. New York: Praeger, USA.

[3] Cardoso, F.H., Faletto, E. (1979), *Dependency and development in Latin America*. Oakland, CA: University of California Press, USA.

[4] Chase-Dunn, C. (1975), *The effects of international economic dependence on development and inequality: A cross-national study. American Sociological Review*, 40(6), 720–738, https://doi.org/10.2307/2094176.

[5] Dos Santos, T. (1970), *The structure of dependence. American Economic Review*, 60(2), 231–236, https://www.jstor.org/stable/1815811.

[6] Evans, P. (1979), *Dependent development: The alliance of multinational, state, and local capital in Brazil*. Princeton, NJ: Princeton University Press, USA, https://doi.org/10.2307/j.ctv346t7d.

[7] Frank, A.G. (1967), *Capitalism and underdevelopment in Latin America: Historical studies of Chile and Brazil. Monthly Review Press*, https://archive.org/details/capitalismunderd00fran.

[8] Furtado, C. (1964), *Development and underdevelopment*. Oakland, CA: University of California Press, USA, https://archive.org/details/developmentunder0000furt.

[9] Gereffi, G. (1983), *The pharmaceutical industry and dependency in the third world*. Princeton, NJ: Princeton University Press, USA, https://www.jstor.org/stable/j.ctt1m3nx5g.

[10] Hirschman, A.O. (1958), *The strategy of economic development. Yale University Press*, New Haven CT, USA, https://archive.org/details/strategyofeconom00hirs.

[11] Hyndman, R.J., Athanasopoulos, G. (2018), *Forecasting: Principles and practice* (2nd ed.). OTexts, Melbourne, Australia, https://otexts.com/fpp2/.

[12] Jean, N., Burke, M., Xie, M., Davis, W.M., Lobell, D.B., Ermon, S. (2016), *Combining satellite imagery and machine learning to predict poverty. Science*, 353(6301), 790-794, https://doi.org/10.1126/science.aaf7894.

[13] Jerven, M. (2013), *Poor numbers: How we are misled by African development statistics and what to do about it*. Ithaca, NY: Cornell University Press, USA.

[14] Jorgenson, A.K. (2006), *Unequal ecological exchange and environmental degradation: A theoretical proposition and cross-national study of deforestation, 1990-2000. Rural Sociology*, 71(4), 685-712, https://doi.org/10.1526/003601106781262016.

[15] Katz, J. (1987), *Technology generation in Latin American manufacturing industries*. New York: Macmillan, USA.

[16] Kregel, J. (2004), *External financing for development and international financial instability (G-24 Discussion Paper Series No. 32). United Nations*, https://unctad.org/system/files/official-document/gdsmdpbg2420048_en.pdf.

[17] Mahutga, M.C. (2006), *The persistence of structural inequality? A network analysis of international trade, 1965-2000. Social Forces*, 84(4), 1863-1889, https://doi.org/10.1353/sof.2006.0098.

[18] Prebisch, R. (1950), *The economic development of Latin America and its principal problems*. United Nations Economic Commission for Latin America, https://archivo.cepal.org/pdfs/cdPrebisch/002.pdf.

[19] Reinhart, C.M., Rogoff, K.S. (2009), *This time is different: Eight centuries of financial folly*. Princeton, NJ: Princeton University Press, USA.

[20] Rodrik, D. (2015), *Economics rules: The rights and wrongs of the dismal science*. New York: W.W. Norton & Company, USA.

[21] Sarti, F., Hiratuka, C. (2017), *Desenvolvimento industrial no Brasil: Oportunidades e desafios futuros. Texto para Discussão*, 187. Instituto de Economia da UNICAMP, https://www.eco.unicamp.br/images/arquivos/artigos/1817/texto187.pdf.

[22] Sen, A. (1999), *Development as freedom*. Oxford: Oxford University Press, UK.

[23] Snoek, J., Larochelle, H., Adams, R.P. (2012), *Practical Bayesian optimisation of machine learning algorithms.* In *Advances in neural information processing systems* (pp. 2951-2959). Curran Associates, https://doi.org/10.48550/arXiv.1206.2944.

[24] Sturgeon, T., Van Biesebroeck, J., Gereffi, G. (2008), *Value chains, networks and clusters: Reframing the global automotive industry. Journal of Economic Geography*, 8(3), 297-321, https://doi.org/10.1093/jeg/lbn007.

[25] Vaitsos, C.V. (1974), *Intercountry income distribution and transnational enterprises*. Oxford: Clarendon Press, UK.