INSTITUTE OF TECHNOLOGY AND LEADERSHIP

VINÍCIUS OLIVEIRA FERNANDES

THE UNPREDICTABILITY OF ECONOMIC RECESSIONS AND THE USE OF MONTE CARLO SIMULATION IN INTERPRETING THE RANDOMNESS OF ECONOMIC CRISES

Integration of Probabilistic Models in the Analysis of Recessive Episodes

SÃO PAULO
2025

VINÍCIUS OLIVEIRA FERNANDES

THE UNPREDICTABILITY OF ECONOMIC RECESSIONS AND THE USE OF MONTE CARLO SIMULATION IN INTERPRETING THE RANDOMNESS OF ECONOMIC CRISES

Integration of Probabilistic Models in the Analysis of Recessive Episodes

Final Course Project submitted to the Institute of Technology and Leadership (INTELI), to obtain a bachelor's degree in Information Systems.

Advisor: Prof. Cristina Gramani

SÃO PAULO
2025

Cataloging in Publication
Library and Documentation Service
Instituto de Tecnologia e Liderança (INTELI)
Data entered by the author.
(Cataloging record with international cataloging data, according to NBR 14724. The record will be completed later, after approval and before the final version is deposited. The completion of the cataloging record is the responsibility of the institution's library.)

Sobrenome, Nome
Título do trabalho: subtítulo / Nome Sobrenome do autor; Nome e Sobrenome do orientador. – São Paulo, 2025.
nº de páginas : il.

Trabalho de Conclusão de Curso (Graduação) – Curso de [Ciência da Computação] [Engenharia de Software] [Engenharia de Hardware] [Sistema de Informação] / Instituto de Tecnologia e Liderança.
Bibliografia

1. [Assunto A]. 2. [Assunto B]. 3. [Assunto C].

CDD. 23. ed.

Resumo

Fernandes, Vinícius. **A Imprevisibilidade das Recessões Econômicas e o uso da Simulação de Montecarlo na interpretação da Aleatoriedade das Crises Econômicas.** 2025. 51. TCC (Graduação) - Sistema de Informação, Instituto de Tecnologia e Liderança, São Paulo, 2025.

A imprevisibilidade das recessões econômicas representa um dos maiores desafios da macroeconomia contemporânea. Eventos como a Grande Depressão, a crise financeira de 2008 e a recessão causada pela pandemia de COVID-19 evidenciam as limitações dos modelos tradicionais de previsão. Indicadores macroeconômicos convencionais, como o PIB e índices antecedentes, muitas vezes não são capazes de fornecer alertas tempestivos sobre períodos de contração econômica. Este trabalho tem como objetivo investigar a aplicabilidade da Simulação de Monte Carlo como ferramenta estatística para modelar a incerteza e aprimorar a capacidade preditiva em relação às recessões. A pesquisa adota uma abordagem quantitativa e exploratória, utilizando dados históricos de indicadores compostos de tendência, PIB per capita e índices de produção industrial em dois contextos de crise: a recessão global de 2008 e a recessão pandêmica de 2020. Os resultados demonstram que a Simulação de Monte Carlo é capaz de gerar distribuições probabilísticas que capturam a incerteza sistêmica, evidenciam riscos de cauda e permitem avaliar a robustez dos indicadores agregados utilizados. Conclui-se que a integração entre modelos probabilísticos e simulações estocásticas oferece uma alternativa promissora às abordagens determinísticas, contribuindo para a compreensão da aleatoriedade inerente às crises econômicas e para o desenvolvimento de instrumentos de previsão mais resilientes e adaptáveis.

Palavras-chave: Recessões econômicas; Incerteza; Simulação de Monte Carlo; Indicadores econômicos; Modelagem probabilística.

Abstract

Fernandes, Vinícius. **The Unpredictability of Economic Recessions and the Use of Monte Carlo Simulation in Interpreting the Randomness of Economic Crises.** 2025. 51. Final Course Project (Bachelor) - Sistema de Informação, Instituto de Tecnologia e Liderança, São Paulo, 2025.

The unpredictability of economic recessions represents one of the greatest challenges in contemporary macroeconomics. Events such as the Great Depression, the 2008 financial crisis, and the COVID-19 pandemic recession highlight the limitations of traditional forecasting models. Standard macroeconomic indicators, such as GDP and leading indexes, often fail to provide timely warnings of economic downturns. This study aims to investigate the applicability of Monte Carlo Simulation as a statistical tool to model uncertainty and enhance the predictive capacity of recession analysis. The research adopts a quantitative and exploratory approach, using historical data from composite leading indicators, GDP per capita, and industrial production indexes across two major crises: the 2008 global financial recession and the pandemic-induced recession of 2020. The results demonstrate that Monte Carlo Simulation generates probabilistic distributions capable of capturing systemic uncertainty, identifying tail risks, and revealing the robustness of aggregated indicators. It is concluded that the integration between probabilistic modeling and stochastic simulations provides a promising alternative to deterministic approaches, contributing to a better understanding of the randomness underlying economic crises and to the development of more resilient and adaptive forecasting tools.

Keywords: Economic recessions; Uncertainty; Monte Carlo Simulation; Economic indicators; Probabilistic modeling.

Summary

1. Introduction

Statistics is a very important tool in our lives. To better understand statistics, it involves the collection, analysis, and interpretation of data along with mathematical theorems, which in simpler terms gives us a better understanding of the relationship between the frequency of events and the probability of their recurrence. It is also present in almost everything in our daily lives.

Within this study, one of the most unpredictable and interesting topics for the population is the predictability of economic events. Whether it's stock prices on the stock exchange, the interest rate index, or even the inflation rate, there is a curiosity about whether there is a statistical term that can act in predicting these events, known as "black swans" (Taleb, 2007). In this context, many believe that a stochastic approach is more relevant in predicting economic contexts.

One of the methods used to predict economic events is the Monte Carlo Simulation. Also known as multiple probability simulation, the Monte Carlo Method is a mathematical technique that simulates possible outcomes of events often considered uncertain. Created by John von Neumann and Stanislaw Ulam during World War II, this technique takes decision-making based on conditions very similar to "chance." Consequently, due to this characteristic, the method is widely used in economic contexts.

The problem is that currently, Monte Carlo simulation is heavily associated with stock pricing, such as in the study by Estember and Mañana comparing the use of Monte Carlo simulation with Artificial Neural Networks (ANN), and in Martin Pažický's work: "Stock price simulation using bootstrap and Monte Carlo," which compares the "bootstrap" method with Monte Carlo. Another example is the study by Jeremy Ng Phak Xiang, Shubashini Rathina Velu, and Sotirios Zygiaris: "Monte Carlo Simulation Prediction of Stock Prices," which combines the application of Geometric Brownian Motion with Monte Carlo simulation. In addition to isolating its application to a few financial indicators, these analyses themselves have a superficial nature, as can be evidenced during reading. This is because these indicators are mapped individually and have a daily attribution window, more focused on a general view of the risk involved from multiple controlled scenarios. The complexity of the method allows us to go further. Due to the random premise within its modeling, considering that it includes the element of uncertainty and randomness in prediction (AWS, n.d.), the best application of the Monte Carlo method would be in a type of analysis and modeling that runs periodically. A practical example of this premise is NASA's JSTAR Simulator, which is essentially a program that independently tests the viability of missions according to each scenario using Monte Carlo simulation (NASA, n.d). With so many examples and complex applications, there is a curiosity about the predictability of macroeconomic events, such as recession periods in certain countries, for instance. However, there are few works that use the Monte Carlo method to analyze macroeconomic scenarios. These works, which have great notoriety, also focus on a more closed scope, such as the article by Huub Meijers, Önder Nomaler, and Bart Verspagen, "Demand, credit and macroeconomic dynamics. A micro simulation model," which uses Monte Carlo simulation to highlight macroeconomic cycles, from demand, credit search, and its indicators, in addition to economic dynamics. Another relevant article is "Monte Carlo simulation of macroeconomic risk with a continuum of agents: the symmetric case" by Peter J. Hammond & Yeneng Sun. It consists of an article that uses macroeconomic agents to predict risks using the Monte Carlo method. Other older articles also deserve mention, such as "Stochastic processes depending on a continuous parameter" by Doob J.L. and "Banks, market organization, and macroeconomic performance: an agent-based computational analysis" by Ashraf Q, Gershman B, Howitt P.

As Alan Kirman pointed out, "Economic Crises are a crisis for Economic Theories," due to specific cases where indicators do not necessarily reflect the country/region's context. The unpredictability of recessions is independent of isolated indicators. Nevertheless, indicators have their relevance and should be taken into consideration.

With this, the use of multiple macroeconomic indicators, along with the unpredictability of recessions in Monte Carlo simulation, is expected to provide insights and a deeper discussion about the probability of recession periods in some countries. This will allow us to discover a technique that enables people to visualize another possibility of mapping economic crises, expanding and democratizing the interpretation of everyday events with statistical analysis. Finally, there is a theoretical applicability of the relationship between "black swans" and Monte Carlo simulation for practical reasons related to unpredictable economic crises.

1.2 Problem

The problem and gap that this research aims to fill is the unpredictability of economic crises and how they cannot be defined based on isolated macroeconomic indicators, as they are complex and have a high degree of randomness. Consequently, this problem (unpredictability) is a common factor among most economic crises, especially in developed countries (San-Martín-Albizuri & Rodriguez-Castellanos, 2017). However, this causes a significant impact on determining a country's economic stage, and with indicators being released with considerable delay, there is a very small window for analysis (Sergey V. Smirnov, 2011) and contextualization of the period. This contributes to outdated research and analyses about the country, which often do not reflect the reality at that moment. In light of this scenario, the following central question arises to guide this research: Given a country's macroeconomic and internal context, what is the probability of an economic recession occurring, considering the unpredictability of macroeconomic factors and the limited time window for analyzing its indicators?

1.3 Objectives

1.3.1 Main Objective

This research has as its main objective to analyze the unpredictability of economic recessions and investigate how Monte Carlo Simulation can be used to interpret the randomness associated with these events. To achieve this, it seeks to understand the nature of economic crises, differentiating the concepts and criteria that define a recession, as well as examining economic cycles and the factors that contribute to their occurrence. The unpredictability of recessions is one of the greatest challenges of modern economics, as it involves both measurable variables, such as macroeconomic indicators, and subjective factors, such as market expectations and the behavior of economic agents. In this context, this research aims to evaluate how these variables interact and how they can be modeled to improve the prediction of these events.

1.3.2 Specific Objectives

To achieve this purpose, the research is structured in several stages. First, a detailed analysis of the factors that contribute to the occurrence of economic recessions will be conducted, distinguishing those that are quantifiable from those that have a subjective nature and are therefore more difficult to measure. Next, it will be investigated why recessions are so unpredictable, considering the limitations of traditional economic analysis methods. In light of this scenario, the potential of Monte Carlo Simulation will be explored as an alternative approach to modeling the randomness of economic crises. This includes explaining the functioning of the method, its application in a real context with random variables, and formulating a practical methodology to employ this technique in identifying possible recessions in different countries. Thus, it is expected to contribute to the advancement of predictive methodologies, allowing a better understanding of the risks and uncertainties associated with economic crises.

2.0 Development - Theoretical Framework

The initial validation process of the project consisted of analyzing scientific articles from the literature that discuss the fundamental theories underlying the project itself. This process allows us to contextualize key topics that will be crucial in forming hypotheses and well-founded arguments. Given the theme addressed in the project, it is essential to understand the concepts that precede the project's perspective and serve as a starting point for the reflections and arguments that will be developed in the following sections. Since the project can be considered a combination of two different topics (the Monte Carlo Method and Economic Recessions), understanding these concepts will be important. In this regard, based on the presented themes, the research was divided into two areas of literature: economics and statistics. In the first area, we will explore three concepts that will serve as the foundation for understanding economic cycles and what defines a recession, as well as its unpredictability. In the second part, we will conduct a statistical review of the application of Monte Carlo simulation, its definition, and how its characteristics can be useful for application in macroeconomic contexts, especially in recessions.

2.1 Theory 1 - Economic Cycles

When we hear about the economy on the evening news or see a post on social media, it is common to think that its composition and definition are simple and static. However, this is merely an impression. Our economy is not only complex but also has branches and phases that change over time. Despite its constant fluctuation and variability, it is necessary to capture a "snapshot" or a specific time frame of the economy to give it meaning. These "periods" are called economic cycles (FGV IBRE, 2025). The economy has certain states and phases that define its cycle, specifically four: expansion, peak, recession, and trough.

The time frame that defines economic cycles is also crucial to measure, as they can be divided into short-, medium-, and long-term cycles. In short-term cycles, often referred to as business cycles, the impacts result from monetary and fiscal policies, production conditions, and external, often unpredictable, factors. This cycle typically lasts between 3 and 11 years. An example of this type of cycle and timeframe was the

COVID-19 pandemic, which lasted roughly three years. The medium-term cycle also incorporates external factors, production conditions focused on national supply chains, and other financial indicators such as credit demand and investment in key sectors, particularly infrastructure (a good example is the investment by some countries in infrastructure and research in the semiconductor sector, where the "fruits of the harvest" will only become evident in the medium term) and technological expansion. The medium-term cycle is also known as the "Kuznets Swing" (Kuznets, S., 1930), which was the first economic segmentation analysis based on migration flows and infrastructure investment (mainly real estate projects).

Finally, long-term economic cycles are defined by profound economic changes, technological revolutions, institutional transformations, and demographic shifts. Also known as the Kondratieff cycle (approximately 40-60 years), the impact of these factors is determined by repeated market fluctuations over this period. A relevant example of a factor influencing this cycle is the "waves" of technological advancements that drive long phases of growth and recession (GRININ et al., 2016).

Understanding economic cycles is crucial for analyzing and predicting potential crises or recessions. Consequently, the most "unpredictable" recessions, which have the highest impact, are usually identifiable in short- and medium-term cycles, where random events play a significant role in triggering the recession. Therefore, these time frames are ideal for distinguishing different types of precursor events leading to recessions.

2.2 Theory 2 - Recession Indicators

Economic recession cycles are influenced by many variables (KOLLURU et al., 2021), such as GDP, inflation rates, tax growth, unemployment rates, and their predictability is limited by the time lag—ranging from months to quarters—of the main economic indicators (Smirnov, 2012). This lag hampers real-time interpretation of a country's economic conditions, requiring the use of alternative indicators that can signal economic activity changes before official data is released. A study conducted by the National Bureau of Economic Research (ESTRELLA & MISHKIN, 1998) also points out that incorporating these indicators into models can lead to overfitting in predictions due to this delay. Among the best-performing indicators, the study highlights the "yield curve spread," which is the difference between government bond interest rates with different maturities, as well as stock market indices (ESTRELLA & MISHKIN, 1998). Other useful indicators are identified in the study "Those Unpredictable Recessions" by Smirnov (2012), which outlines three key indicators that may reflect economic cycle changes. The first is The Leading Economic Index (LEI) from The Conference Board, a composite index of variables that tend to shift before overall economic changes. These include weekly working hours in the industry, new orders for consumer goods, building permits, stock prices (for the U.S., the S&P 500 is used as a benchmark), and the difference between 10-year bond yields and the federal funds rate (a key benchmark for the U.S. financial system). The second is The Composite Leading Index (CLI) from the OECD, which aggregates several leading indicators, including business surveys, production data, and financial variables. Its adjusted version is particularly useful as it smooths short-term fluctuations and provides a normalized view of economic trends. Finally, the third one is The Purchasing Managers' Index (PMI) from the Institute for Supply Management, an indicator based on surveys that reflect the economic health of the industrial and service sectors. It is built from responses by supply chain managers regarding new orders, production levels, employment, supplier deliveries, and inventory levels. The index ranges from 0 to 100, with values above 50 indicating expansion and below that signaling contraction.

Smirnov (2012) also complements his study with an analysis of how these indicators "stood the test" in identifying the economic cycle shift during the 2008 U.S. recession. In his assessment, all indicators had been declining since late 2007, signaling an economic slowdown and a potential recession. Notably, the PMI had already been in decline (below 50) for the six months preceding the 2008 crisis. However, it is essential to emphasize that all cyclical indicators are subject to revisions as statistical data is adjusted, leading to discrepancies between real-time perception and historical analysis. To manage this uncertainty, Smirnov mentions empirical methods such as the "5 out of 6" rule, which assumes a trend is genuinely in progress if a cyclical indicator has been rising or falling in at least five of the last six months.

While the indicators presented are valuable tools and offer different approaches to analysis, accurately predicting recessions remains a challenge, as these indicators also fluctuate and may be subjective depending on a country's context. Consequently, real-time analysis should be complemented by historical approaches and robust methodologies to minimize the risk of misinterpretations. Nonetheless, Smirnov's study encourages expanding the use of alternative economic indicators, proving that cycles are not solely defined by widely known metrics but rather by a combination of internal and external factors that highlight the complexity of a country's economic landscape.

Table of Indicators for better understanding					
Indicator	**Abbreviation**	**Frequency**	**Key Components**	**Main Purpose**	**Responsible Institution**
Leading Economic Index	LEI	Monthly	Interest rates, unemployment insurance claims, stock index, new industrial orders	Estimate short-term economic movements	Conference Board (USA)

Table of Indicators for better understanding					
Composite Leading Indicator	CLI	Monthly	Industrial production, consumer confidence, exports, unemployment rate	Understand economic trends across various countries	OECD
Purchasing Managers' Index	PMI	Monthly	Production, new orders, inventories, employment, supplier deliveries	Measure activity in the manufacturing and services sectors	IHS Markit / ISM (USA)
Gross Domestic Product	GDP	Quarterly	Total production of goods and services in a country	Measure economic growth	National governments and international institutions

2.3 Theory 3 - The Unpredictability of Economic Recessions

It is difficult to believe in something we have never seen or do not understand how it works. Our daily lives and routines are shaped by what we know and are accustomed to doing. The economy is no different. Market fluctuations typically follow a pattern based on business performance, technological advancements, macroeconomic contexts, and overall political transformations. This market volatility, although sometimes unpredictable, is directly governed by these everyday contexts. This unpredictability of the economy is captured in the concept of "Black Swans" (Taleb, 2007).

The phenomenon, introduced by mathematician and economist Nassim Nicholas Taleb, draws an analogy to market behavior using the rarity of black swans in a world dominated by white swans. Even though these black swans exist, few believe in them simply because they have never seen one. This metaphor represents rare, catastrophic, and unpredictable events that shake the market, particularly in terms of how extreme events emerge and take people by surprise.

In his book, Taleb differentiates between systems governed by normal distributions ("Mediocristan"), where extreme events are insignificant, and those where a small number of extreme events dominate the outcomes ("Extremistan"). This is a crucial point, as he argues that the Law of Large Numbers, which governs Gaussian distribution behavior, does not necessarily reflect actual market behavior. Based on this concept, Taleb explains that many historical data analyses fail to incorporate the probability of rare events into their samples. This often results from confirmation bias and sampling errors. A classic example is the collapse of several American banks in 1982, which assumed years of growth indicated stability while ignoring latent risks.

For recessions, the issue is even more significant. Taleb argues that crises are fundamentally unpredictable because financial forecasting models frequently overlook fat-tailed distributions and systemic shocks. In identifying potential recessions, there is a need to adopt nonlinear models, as simplified linear models fail to capture the market's true behavior.

The unpredictability of economic recessions is both fascinating and terrifying. While we know that predicting the future is impossible, models and analyses must at least acknowledge the existence of unpredictable events by capturing their randomness. Extreme events can emerge without warning, and traditional methods often fail to recognize them. According to Taleb, the key is not to predict crises but to be prepared for their inevitability through more robust approaches.

2.4 Theory 4 - Statistical Methods and Stochastic Variables

It is incorrect to assume that raw data, which we typically have access to, should be considered as absolute truth or, even less, that it is self-explanatory. In any context of data analysis or inference, it is essential to undergo a process of data treatment and transformation (without altering its core content) to facilitate interpretation and usability, ultimately converting data into information. This concept is crucial because, nowadays, it is common practice to make decisions based on data (MIT OPENCOURSEWARE, 2017).

Consequently, based on the premise that data provides us with insights, we utilize certain tools to extract meaningful patterns according to their parameters. These tools are known as statistical methods or models, particularly those relying on inferential statistics, which encompass mathematical techniques that transform sample data into structured information or well-founded hypotheses (Lane, D. et al., 2003). The primary function of these methods is to ensure that, based on our hypotheses, data lead to a clear interpretation of the study being conducted. In other words, they eliminate data redundancy and define the statistical (or mathematical) patterns of a given sample (RISSANEN, 1986). In this way, statistical models equip us with the necessary tools to respond appropriately to the trends and findings derived from the data. It is important to highlight that, when discussing this type of model and inferential statistics, we assume that both the sample and the population are defined randomly.

A fundamental aspect of statistical models is the concept of variables. Variables can be categorized as either qualitative or quantitative. Qualitative variables cannot be numerically or logically ordered (e.g., eye color, favorite sports team), whereas quantitative variables are measurable, represent a quantity or magnitude, and can typically be ordered. Another important characteristic of variables is their variation over the analysis period. Discrete variables assume countable values, usually integers, with no possibility of taking intermediate values. Conversely, continuous variables can assume any value within a given range, including decimal points.

The process of selecting variables to extract information and initiate the analysis phase is often quite challenging. This is because different combinations and definitions of each variable can yield different meanings in the analysis. The complexity increases even further when dealing with stochastic variables—that is, variables that follow a random distribution, making their behavior inherently unpredictable.

During the research development process, the primary goal is to generate insights within a context of random events. In this case, we will employ the Monte Carlo method as a statistical approach to analyze economic recessions, which emerge from environments governed by stochastic (random) variables. As previously mentioned, statistical models play a crucial role in data analysis and real-world phenomena. However, in the context of recession unpredictability, the challenges become even greater, as variable fluctuations are constant and randomness is significantly high. For this reason, traditional statistical methods are not the most suitable for analyzing this behavior. The Monte Carlo method emerges as an alternative approach to model multiple possibilities using probabilistic distributions that incorporate the randomness of variables.

2.5 Theory 5 - Method of Monte Carlo

Defining a statistical analysis model is an extensive and complex task. This choice is typically closely aligned with the available data and the type of inference one intends to make. Therefore, considering the variables and the overall context of recessions, the Monte Carlo method was selected as the most appropriate statistical model for this study. The Monte Carlo method can be defined as a type of simulation that generates random samples based on mathematical functions to reach a specific result. Consequently, the primary purpose of this approach is to improve inference about different possible outcomes, as it allows for an expanded number of simulated scenarios chosen at random (CERÁVOLO & HOCHHEIM, 2022).

In this context, the first step is to assign random values to the model based on its probability distribution function (typically Gaussian). Then, results are calculated using the model. This process is repeated multiple times until a probability distribution function for the expected outcomes can be obtained.

As highlighted by Marcos and Luiz in their work *"Monte Carlo Simulation and Valuation: A Stochastic Approach"* (2012), the process of identifying the probability distributions of input data is crucial for supporting and substantiating the simulation process. Typically, this selection is based on empirical and historical data analyses to fit the data to the appropriate distribution.

2.6 Theory 6 - Method of Monte Carlo in Macroeconomics

Macroeconomic risk modeling faces several challenges due to the inherent complexity of interactions among economic agents. Companies, individuals, and indicators act as independent agents, each representing a specific "risk" within the system, which introduces a high level of randomness and unpredictability in the analysis of the macroeconomic context, especially during recessions (Hammond et al., 2007). To address this uncertainty, Monte Carlo Simulation emerges as an interesting tool, allowing the estimation of probability distributions and understanding the effects of unpredictability on macroeconomic variables. For specific modeling of economic recessions, a common approach is to represent the economy as a "continuum of agents," a mathematical concept used to approximate infinitely large populations. In this model, each agent has an individual risk, and the random variables associated with them are pairwise exchangeable, meaning that swapping two agents i and j does not change the joint distribution of these variables. This characteristic indicates that although individual risks (mentioned earlier) are identically distributed, they can be correlated, reflecting the interdependence among different economic participants.

The application of the Monte Carlo Method in this scenario allows dealing with the difficulties associated with the measurability and dimensionality of variables that have some impact on the economy. By using this method, the intention is to numerically approximate the solution to complex multiple integrals that arise when trying to predict the behavior of macroeconomic variables under uncertainty. The central strategy consists of calculating the integral of a real function by taking the average of the function values evaluated at randomly chosen points. Thus, even with a finite number of samples, convergence to the expected value occurs with high precision, ensuring that the estimates are reliable. In other words, even though the economic system presents individually distributed random risks, the average of the simulations converges to the expected result, allowing solid inferences about macroeconomic dynamics.

3.0 Scientific Methodology

The process of defining the methodology is a cycle that establishes some fundamental points for the development of the research. This is because scientific research methodologies are characterized by the nature of the topic, the approach to the problem, and the technical or more contextual definition of the topic (De Lunetta and Rodrigues Guerra, 2023). All these characteristics must be oriented towards the "core" of the research and aligned with the problem raised or the problematic question. Thus, based on this premise, the process of defining the research methodology allows us to focus more on segmenting the most relevant variables for the study, defining the ideal sampling for the research context, collecting the most relevant data based on what will be analyzed, and so on (Sampaio, 2022). This definition is a critical path for identifying the results and formulating the final theory in the conclusion of the research.

Based on the study and analysis of the themes addressed in the theoretical framework and from the defined problem, the scientific methodology will be an exploratory case study with a quantitative approach. The intention is to conduct an exploration with an emphasis on collecting and analyzing data and indicators during relevant global recessions, mainly the 2008 recession. The purpose of this part is to prove and substantiate the concepts of the theoretical framework and support the next steps of analysis and reflection in the research.

3.1 Methodological Procedures

As presented in the theoretical framework (2.1 and 2.2) to capture macroeconomic contexts, it is crucial to understand the transition cycles within the period and conduct an analysis based on the particularity of the event itself using key indicators. Starting from this premise, the methodological procedure that best fits this research objective is a case study. The case study procedure seeks to understand a specific phenomenon or event through a detailed investigation, gathering extensive knowledge about it (De Toledo Krücken et al., 2009). As highlighted by Leonard-Baxton (1990), a Case Study can be defined as the "history" of a past phenomenon, explained or measured from different sources of evidence.

From its definition, it is possible to segment the study into three different types of research based on its purpose: Intrinsic, Instrumental, and Collective (Stake, 1995). The research process defined here was Instrumental, aiming to generate insights and relevant data for a deeper understanding of recessions.

Thus, the research procedure will focus on studying the particularity and complexity of a specific "case," which is the cycle of economic recession. This procedure is linked to understanding the "behavior" of recessions, their complexity, and their impact according to each studied context (Da Silva et al., 2021).

3.2 Data Collection

The research instrument for data collection corresponds to the process of organizing the information that will be analyzed in the study. At this stage of the research, a systematic collection of information was carried out to support the development of the analysis. This data gathering includes detailed information about the recessions. The selected recessions for analysis were the 2008 financial crisis and the COVID-19 pandemic recession. Based on this, market characteristics, economic indicators, and the contextual elements of each recessionary period were examined.To ensure a comprehensive approach, key economic indicators relevant to the study were also gathered and analyzed. These indicators allow for a broader understanding of the context in which the project is developed. During the analytical phase, monthly time series data were collected for the LEI (Leading Economic Index), CLI (Composite Leading Indicator), and PMI (Purchasing Managers' Index) throughout the full 2008 recession period, including the lead-up and the recovery phase. For the pandemic, the focus was primarily on the LEI index, using monthly windows that covered the full extent of the recession, given the unique characteristics of that crisis and the adverse effects of lockdowns that could compromise the analysis. The GDP growth rate was also analyzed for both recession periods. Thus, this initial phase is crucial for structuring the following stages of the research, ensuring that methodological and analytical decisions are based on solid and reliable information.

3.2.1 The Great Recession

Triggered by the expansion of a plan formulated by Lewis Ranieri, the "father" of securitized mortgage-backed securities (MBS), the 2008 recession was one of the greatest challenges of modern economics. The shock stemmed from the housing market "bubble" and the segmentation of subprime securities (bonds composed of high-risk debt, also associated with CDOs—Collateralized Debt Obligations), which led to a downturn and turmoil for some of the world's largest banks. However, the most interesting aspect of the 2008 recession, or the subprime crisis, is that many considered it unlikely or even impossible, given that the U.S. housing market had historically been very stable. This perception was further reinforced by the performance of the stock market and real estate indices such as TABX and ABX (which track the risk and performance of Mortgage-Backed Securities), as well as the S&P 500, all of which remained bullish even when mortgage default rates were rising significantly. Nonetheless, these indices meant little to those unaware of the housing market bubble—and even to those who were aware, like hedge fund managers with short positions against the market, such as Steve Eisman and Michael Burry, these indices did not necessarily reflect the severity of the problem.

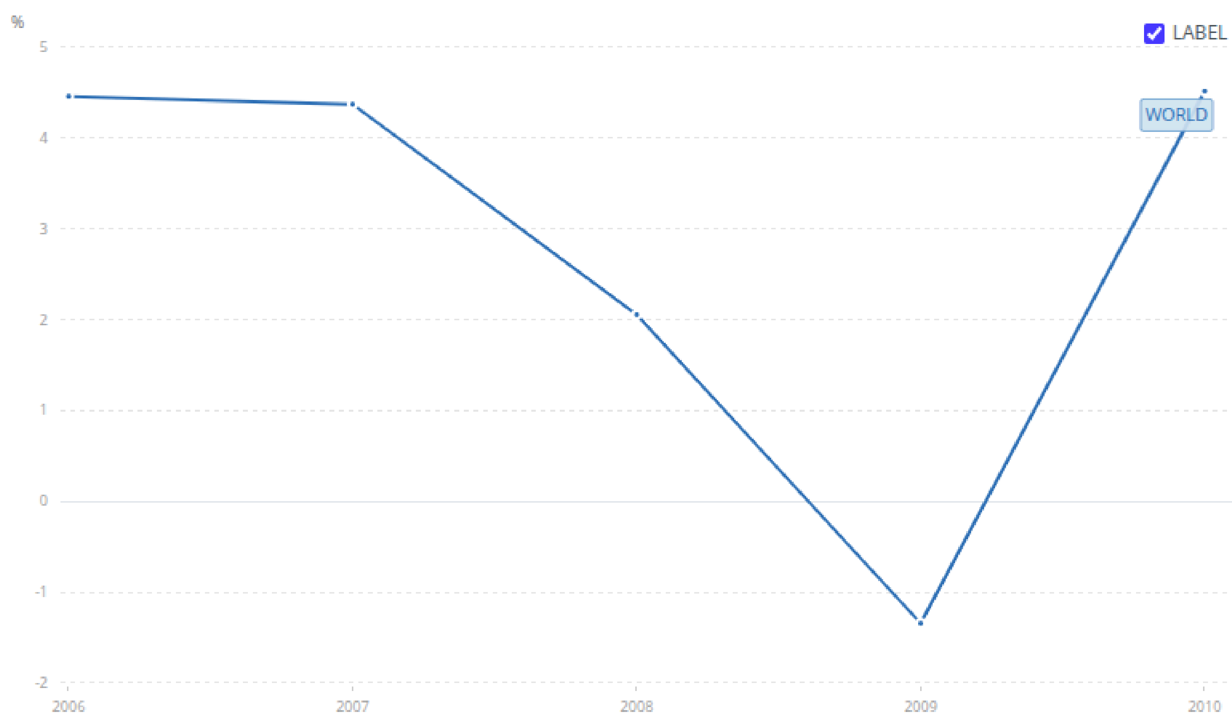
Thus, the main question here is to assess whether any macroeconomic indicator could have identified this bubble, based on the extensive data on default rates and economic slowdown.

3.2.1.1 GDP

As previously mentioned, GDP is a crucial indicator for mapping a country's internal growth and development. It was one of the many indicators affected by the 2008 recession, with caveats for each phase of the seemingly unpredictable economic cycle. To understand the impact of this indicator and its behavior, we will analyze the period before, during, and after the 2008 recession. Before the recession (mid-2007), signs of economic slowdown were evident, with a slight decline in GDP growth. This decline continued until the end of 2007, with growth rates nearing 0%. During the recession, GDP continued to fall steadily. Interestingly, the decline persisted throughout 2008 without abrupt short-term crashes. This behavior can be explained by the market's slow pricing of the crisis (which becomes more evident with other indicators), the optimistic outlook of renowned market experts—who were evidently unaware of the housing market bubble—the uncertainty about "how deep the hole was" (after the crash, estimates and assessments were conducted to determine how leveraged banks were concerning real estate assets, how much they owed in CDS contracts from short sellers betting against the market, and what their total liabilities were), as well as the government's deliberate delay in defining strategies to cover the banking deficit. However, the decline in growth was inevitable, and the U.S. ended 2008 with GDP growth nearing -2.5%. The post-recession period saw significant growth, characterized by various government economic stimulus packages and Federal Reserve incentives, such as interest rates approaching zero to revive the economy. The figure below (World Bank Group, 2008) shows GDP growth percentages from 2007 to 2010, highlighting the key points mentioned above:

Source: World Bank - GDP growth (annual %) for the United States

The U.S. housing market was crucial to other countries, especially European nations holding positions in real estate assets previously considered low risk. Consequently, they also suffered from the crash during the recession. Although the impact was somewhat smaller than in the U.S., it was still significant. Before the recession, the global economy was expanding, driven by the increased participation of emerging markets in commodity control and the expansion of their internal markets, such as China (Wentzel, 2008). During the crisis, global GDP contracted by approximately -1.7% from 2008 to 2009. Advanced economies, including the Eurozone and Japan, experienced severe recessions, while emerging markets were significantly impacted, though less intensely. However, the most affected countries during this period, besides the U.S., were the United Kingdom and Iceland. In the UK, a notable case involved the insurance company AIG. The renowned insurer was one of the largest issuers of CDS (Credit Default Swaps, previously mentioned), accumulating significant exposure to these products, with an estimated $441 billion in CDS contracts by June 2008 (Silva, 2019). This meant that if widespread mortgage defaults occurred, AIG would have to honor these contracts, which was financially unsustainable, ultimately leading to the company's collapse and a major shock in the UK. Another significant example of the recession's impact on a small economy was Iceland. The financial crisis stemmed from the heavy dependence of the country's three main banks: Glitnir, Landsbanki, and Kaupthing. These banks had expanded rapidly, accumulating debt that exceeded the country's GDP. The collapse of Lehman Brothers triggered the crisis, as these banks were leveraged with short-term loans and could no longer expand, in addition to their exposure to U.S. assets. This led to an unprecedented economic shock, causing a sharp depreciation of the Icelandic krona and high inflation. Other countries, such as China and India, continued growing but at reduced rates. After the recession, following U.S. monetary policy, global GDP grew by around 4.5%, driven by fiscal and monetary stimulus in various economies. Emerging markets, led by China, recovered quickly, while the Eurozone struggled due to the sovereign debt crisis. The supporting chart below shows global GDP growth percentages:



Source: World Bank - GDP growth (annual %) for the United States

3.2.1.2 LEI, CLI, PMI

Inspired by Smirnov’s (2011) research and findings, a strong correlation can be identified between some indicators and the shifting phases within the 2008 recession, particularly the LEI, CLI, and PMI. These indicators are primarily limited to the country’s context, and in this sense, their interpretation in the American scenario is highly relevant to the "Great Recession." It is important to understand that these indicators encompass a nearly complete view of the country’s economy and infrastructure policy, ranging from stock market and government bond evaluations to industrial stimulation and an overview of the service sector, with the latter being particularly considered by the PMI. Given its more precise forecasting performance during the 2008 recession, it can be inferred that this index successfully captured the slowdown in the service sector due to rising default rates and decreasing domestic market participation in consumer goods acquisition.

The adaptability of these indicators is crucial as they are relatively insulated from fiscal incentives and short-term stimuli, which are typically the government’s primary responses to the most drastic signs of a recession.

Index	**Period**	**Initial Index (score) ***	**Year to Year Change ***	**Hypothesis**
LEI (TCB)	December 2007	-4	-4	Fall below two-year support level.
CLI (OECD)	November 2007	-4	-4	Negative trend despite limited data.

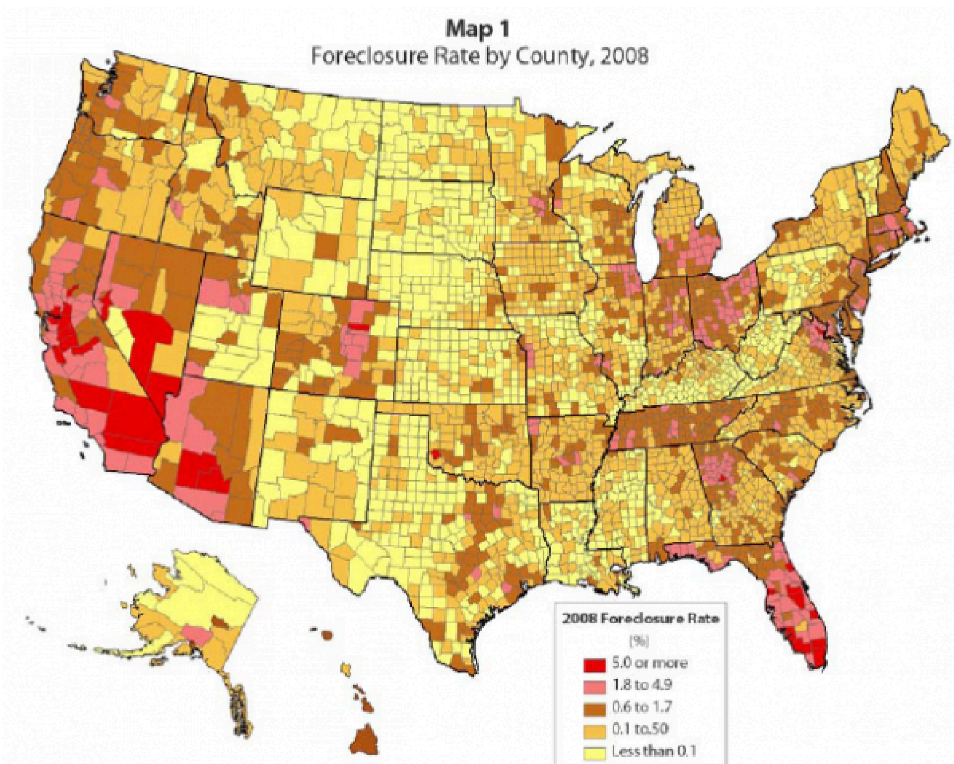
PMI (ISM)	December 2007	-6	0	Strong signal for the start of the 2008 recession.
LEI (TCB)	June 2009	0	+2	Increase in the last three months.
CLI (OECD)	June 2009	-2	0	Low growth with limited data.
PMI (ISM)	July 2009	+6	+6	Strong signal for the end of the 2008 recession.

Source: Table created in the study by Smirnov (2011) with indicators extracted from "The Conference Board" organization.

- – Initial Index and Year-over-Year Change (January 2008 or July 2009).
  A negative net score means that the number of declines (during a six-month period before a turning point, or change in the economic cycle) is greater than the number of increases; for a positive score, the opposite is the case.

3.2.1.3 Foreclosure Rate

Each recession has its own characteristics. The particularity of the 2008 recession was the bubble created in the U.S. real estate system. This bubble was exacerbated by the abrupt leveraging of some banks, as well as the aggregation of multiple high-risk loan contracts into securities sold on the market (subprimes). Even so, what's interesting is that before this bubble "burst," there were already indicators predicting a potential "collapse" of this system, such as the ABX, an index that tracks the closing values of CDOs related to mortgages and real estate loans (Longstaff, 2010), and the percentage of foreclosures, which essentially reflects the default rate, considering that the foreclosure process involves the creditor recovering the asset that was lent. Of the two indicators, foreclosure was not only noticeable in the financial sector, but also part of the daily lives of Americans, many of whom lost their homes, knew someone who did, or faced a dramatic increase in their mortgage rates. This perception, beyond being part of daily life, could be inferred from the high foreclosure rate distribution across the United States in 2008:



Source: Realty Trac

With the average foreclosure rate at 1.8% (high in the U.S. context, across 50 states), it becomes clear how widespread the real estate "bubble" was throughout the country and how it posed an imminent risk during 2008. However, any movement at that point would have been "too late," and few predictions accounted for this surge in their analyses. Those that did, either failed to assign the necessary relevance to the issue or dismissed any risks arising from the rise in rates.

3.2.1 The Covid-19 Pandemic and Its Economic Impact

The Covid-19 pandemic, which began in 2019, was a global and unpredictable event caused by the spread of the SARS-CoV-2 virus. In addition to being a severe public health crisis, resulting in millions of deaths and overwhelming healthcare systems, the pandemic triggered a deep global economic crisis.

To control the spread of the virus, governments implemented various measures, including mask mandates and restrictions on gatherings. However, as cases and deaths surged exponentially, these measures became more drastic. The most significant was the lockdown, a government-imposed restriction on movement. This led to a substantial interruption in economic activities, causing simultaneous supply and demand shocks. This scenario resulted in the brief Covid-19 recession, considered one of the most severe global recessions since 2008. In 2020, global GDP contracted by 5.2%, with even steeper declines in major world economies (7%) and severe impacts on emerging and developing markets (World Bank Group, 2023). Unemployment skyrocketed in various countries, with millions losing their jobs or facing reduced working hours. Sectors such as tourism, hospitality, and manufacturing were particularly affected.

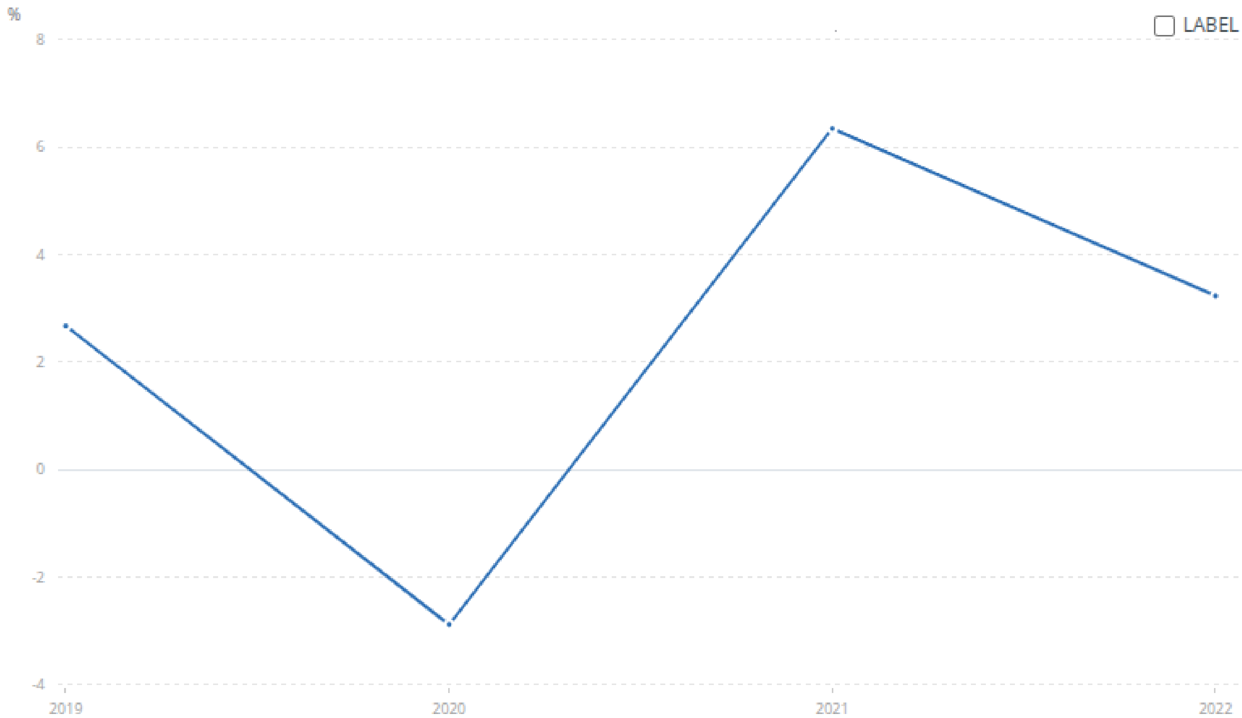
Additionally, the pandemic led to a financial market collapse, including stock market crashes in March 2020 and a sharp decline in oil prices, exacerbated by a price war between Russia and Saudi Arabia. Despite swift government responses with historic fiscal and monetary stimulus, the economic effects were extensive and prolonged.

Ultimately, while some economies began recovering in 2021 due to vaccination efforts and gradual reopening, the economic legacies of the pandemic include elevated public debt levels, rising inflation, and challenges in global supply chains.

3.2.2.1 GDP

Global GDP growth from 2019 to 2022, marked by the Covid-19 pandemic, exhibited strong fluctuations. In 2019, before the pandemic, growth was around 2.5%, reflecting a moderately expanding global economy. However, at the beginning of 2020, the economy witnessed a drastic decline, with global GDP growth reaching approximately -3%, reflecting the impact of containment measures and the health crisis. This decline highlights the consequences of lockdowns, mobility restrictions, and widespread uncertainty, leading to a significant contraction across various sectors and characterizing this period as a global recession.

In 2021, following the pandemic's most critical phase and driven by economic reopenings and government stimulus, there was a rebound, with growth jumping to about 6.5%. However, in 2022, growth slowed to approximately 3.2%, possibly due to factors such as persistent new variants, increased public debt following multiple stimulus measures, and rising inflation. The chart below (World Bank Group, 2023) illustrates the impact of the pandemic on the global economy (represented by GDP growth percentage), depicting an initial decline, followed by recovery and subsequent deceleration.

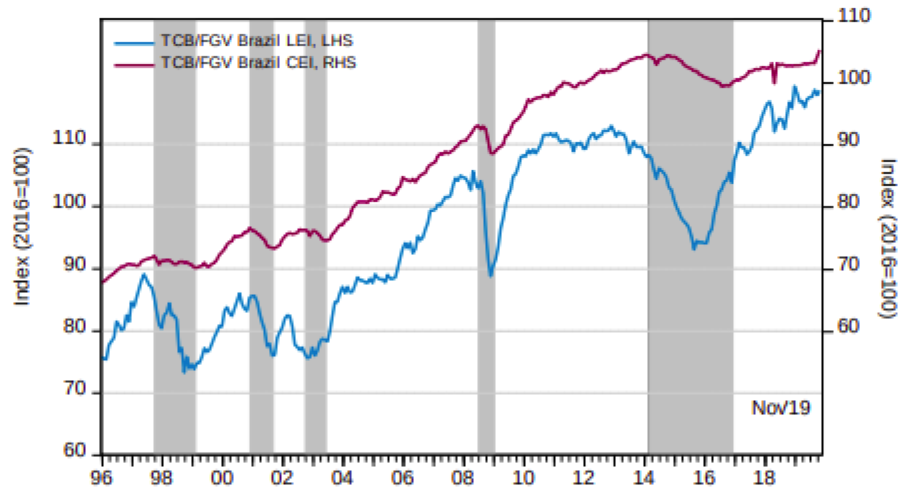


Source: World Bank - GDP growth (annual %) for the World

3.2.2.2 LEI, CLI, PMI

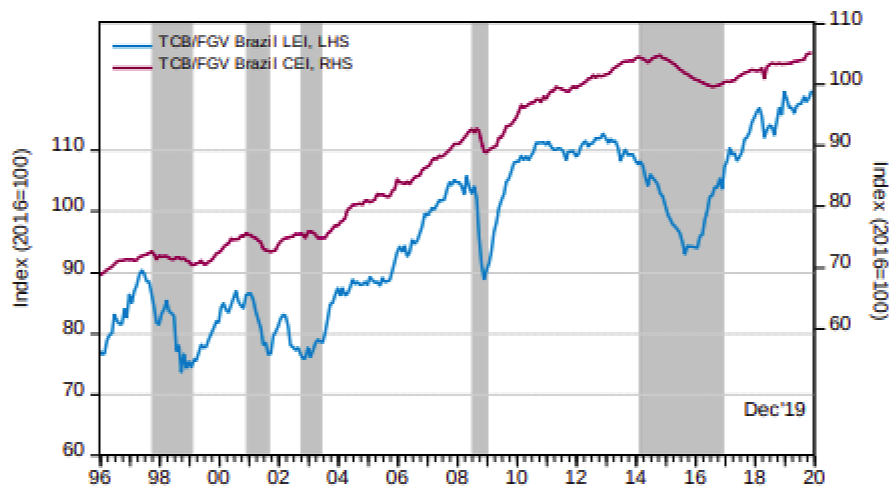
To analyze the economic impact at a national level, some indicators highlighted by Smirnov (2011) can be used to understand the pandemic period. Specifically, the Leading Economic Index (LEI) can be selected as the main indicator for analysis, and Brazil can be chosen as a case study to observe the pandemic's economic impact and its predictability:

LEI and CEI rose in November



Source: FGV IBRE e The Conference Board

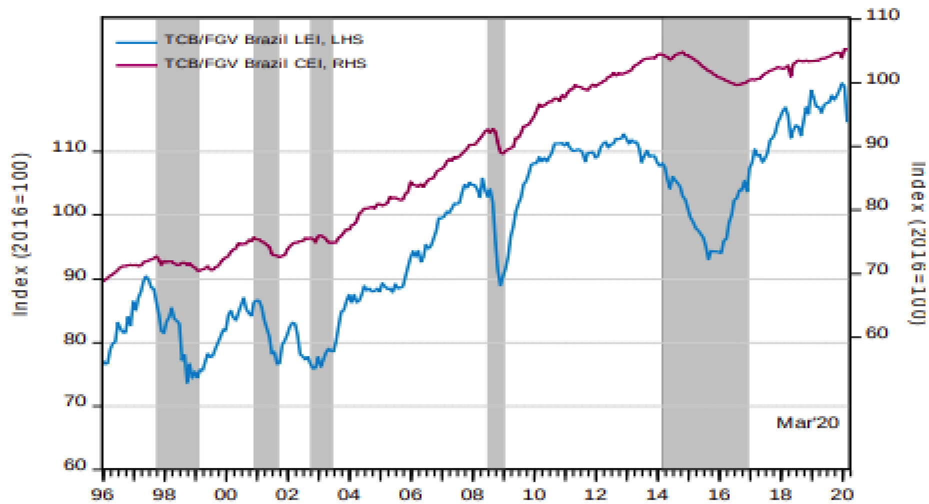
LEI and CEI rose in December



Source: FGV IBRE e The Conference Board

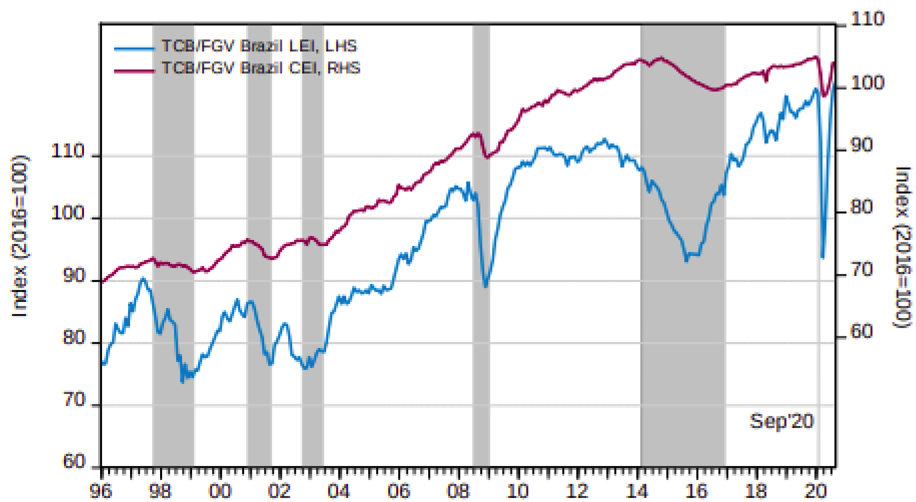
An increase in the LEI index, albeit minor in December, suggests economic growth and development in Brazil towards the end of 2019 and early 2020. However, with the arrival of 2020 and the virus outbreak, the scenario quickly changed:

LEI and CEI decreased in March



Source: FGV IBRE e The Conference Board

LEI rose in September



Source: FGV IBRE e The Conference Board

From March to September, a significant decline in the index is evident during the peak of the pandemic, marked by the months of March, April, and May, when uncertainty regarding the virus and its treatment was at its highest. A subsequent recovery is observed until early 2021, after which the index follows normal market fluctuations without abrupt volatility:

LEI and CEI declined in March

Source: FGV IBRE e The Conference Board

Even so, the indicator's behavior remained abnormal, though more controlled and influenced by internal policies rather than the unpredictable external events that initially triggered the pandemic.

This observation highlights a gap. There is a limit to how far recession predictions can go, restricted to economic cycle changes not caused by financial crises or industrial crashes but by random events that are not detectable through the overlap of economic indicators.

3.3 Inferring the Data Collected

Each point of discussion and analysis regarding the selected recessions leads us to reflect on their interpretability. By not only looking at the two recessions in isolation, but also understanding the context of past recessions, it becomes clear that their causality differs. Additionally, we can group them into different segments, with the 2008 recession falling into the category of financial crises, and the COVID-19 recession being more related to the advent of the pandemic/public health crisis, for example. By making this distinction, it is evident that each recession has different precursors, levels of predictability, and impacts on society. Therefore, the impacts of the cause and, later, the recovery will also differ (PERRY et al., 1993).

The choice of the 2007 to 2009 interval for the 2008 recession is justified by its coverage of the full cycle of the global financial crisis. This timeframe allows for the analysis of economic effects in the pre-crisis stage (2007), during the crisis (2008–2009), and post-crisis recovery phase (starting in 2010). The sharp drop in GDP in 2009, as previously mentioned, highlights the severe impact of the crisis on the U.S. economy, reflecting a significant contraction triggered by the housing market collapse and the failure of major financial institutions. As for the pandemic, the interval from 2019 to 2021 is commonly used to assess the economic impacts of the COVID-19 health crisis on the global economy. The justification for this timeframe lies in the observation of economic activity levels immediately before the pandemic (2019), during the peak of the crisis (2020), and during the early stages of recovery (2021).

What is particularly interesting about this analysis is the understanding that each recession carries a high level of unpredictability within its specificities. However, after the economic cycles have moved, it is possible to identify even more specific niches within each recession based on the moment a country, for example, is experiencing. This helps in determining more precise indicators in the post-event analysis process. Still, the particularity of each recession also aids us in inferring the degree of predictability and research on its "root cause." It is evident that the predictability of the 1929 crisis, for example, is not the same as the Great Depression, and certainly not the same as the pandemic. This is because the context and monitoring of the cycle change for some recessions are explained and measured through deviations, indices, rates, growth curves of internal or external variables, and other quantifiable indicators. In situations like the one we faced during the pandemic, it is very difficult to find any indicator that could have helped us anticipate or at least explain the crisis itself. Therefore, the field of predictability for most recessions is more relevant for financial crises, environmental crises, and other areas that are unpredictable in their specific causes but are measurable and tangible for statistical calculations and inferences.

Finally, more interesting than this is the ability to map these indicators before a recession fully unfolds.

3.3.1 Data Description

This section presents the data used in the analysis, detailing the sources, the types of variables involved (aggregated, continuous, categorical, or normalized), and justifying the choice of the selected time intervals, particularly regarding the 2008 economic crisis and the COVID-19 pandemic.

For the GDP growth data—both for the United States (specifically regarding the subprime recession from 2007 to 2010) and for global data—these refer to the annual growth rate of Gross Domestic Product (GDP), as presented by the World Bank. These are aggregated data, as they represent the variation in the total economic activity of a country over time. The values are not normalized, being expressed in real annual percentage terms. For the foreclosure rates in the U.S. (2008), the data are aggregated by geographic region (county) and classified into ranges, which characterizes a categorical variable. The rates are shown in raw percentage values, meaning they are not normalized. Counties highlighted in red and orange indicate regions with foreclosure rates above 1.8%—reaching over 5% in some areas—especially in states such as California, Nevada, Arizona, and Florida. These data complement the macroeconomic analysis of GDP by offering a geographic perspective on the financial vulnerability experienced by thousands of American households during the crisis.

Regarding the COVID-19 pandemic, in addition to GDP analysis (already discussed above), the other dataset includes the LEI (Leading Economic Index) for Brazil, developed by FGV/IBRE in partnership with The Conference Board. These values can be considered continuous, as they fluctuate over time, and the data have been normalized, with the base year set to 2016 (index = 100), facilitating standardized temporal comparisons.

4.0 Aggregated Models

The Monte Carlo Simulation was adopted in this study as the central tool for the probabilistic estimation of economic recessions, based on a set of aggregated indicators (such as Composite Leading Indicators – CLI, GDP per capita, and industrial production indices).

The method combines statistical and probabilistic models to generate probability distributions over future events, taking into account the inherent uncertainty of parameters and the temporal dynamics of economic variables.
In practice, the proposed structure was composed of three fundamental blocks:

- Probit Model – responsible for estimating the conditional probability of a recession occurring based on explanatory variables at a given point in time;
- VAR(1) Model – captures the temporal dependence among the explanatory variables, simulating their evolution over the future horizon;
- Monte Carlo Simulation – generates multiple possible scenarios for economic indicators and calculates, in each scenario, the probability of a recession, obtaining averages, confidence intervals, and cumulative probabilities.

This integration between probabilistic regression and temporal dynamics makes the model more realistic in an economic context, allowing analyses that go beyond point forecasting and enabling the quantification of recession risk as a probability distribution.

The modeling was fully implemented in Python, employing numerical optimization routines (the Newton-Raphson method for Maximum Likelihood Estimation in the Probit model), Ordinary Least Squares (OLS) estimation for the VAR(1), and custom Monte Carlo sampling and aggregation routines.

The final result provides not only average probabilities of recession over the forecast horizon but also confidence intervals (5% and 95%) and Average Marginal Effects (AME) for each variable, facilitating economic interpretation.

4.1 Logistic Regression and Probit Regression

During development, different binary models were evaluated to estimate the probability of recession. The main ones were:

- Logistic Regression, which estimates the probability of a binary event through the logistic function. This model is widely used due to its interpretability and numerical stability. However, its main limitation in this case lies in the assumption of temporal independence and fixed functional form, which may not adequately represent economic regimes with abrupt transitions.

- Probit Regression was implemented as an alternative, whose formulation uses the cumulative distribution function of the standard normal. This modeling is more suitable in contexts of aggregate uncertainty, as it allows smoother variations in probabilities and can be directly integrated into the stochastic process of the Monte Carlo Simulation.

The estimation was performed via Maximum Likelihood Estimation (MLE), using first- and second-order derivatives for optimization (gradient and Hessian). The model showed stable convergence and interpretable coefficients, allowing inference of the marginal impact of each variable on recession risk.

From this structure, the instantaneous probability of recession was integrated into a dynamic process, in which the predictors evolve according to the VAR(1) model, thereby providing a consistent chain of probabilistic forecasts over time.

4.2 Counterproof Models vs Montecarlo

To assess robustness and compare the performance of the Monte Carlo Simulation against other methods, a series of experiments was conducted using different statistical and probabilistic models.

These experiments formed what was called the **"Model Counterproof"**. The tested models included:

Model	Type	Considers Temporal Evolution	Accuracy	Precision	Recall
Multiple Regression	Static Linear	No	N/A	N/A	N/A
Logistic Bootstrap	Nonlinear Statistical	Yes	1.00	0.00	0.00
Markov Chain	Temporal Probabilistic	Yes	0.556	0.00	0.00
Threshold VAR (TVAR)	Temporal Statistical	Yes	1.00	1.00	1.00
Monte Carlo Simulation	Temporal Probabilistic	Yes	0.875	1.00	0.875

5.0 Results

The consolidated results showed significant progress in the maturity of the proposed model and in the understanding of its practical limitations.

The final version of the model integrated with the Monte Carlo Simulation presented the following results:

- **Average Accuracy:** 0.875
- **Precision:** 1.000
- **Recall:** 0.875
- **F1-Score:** 0.933
- **Brier Score:** 0.173

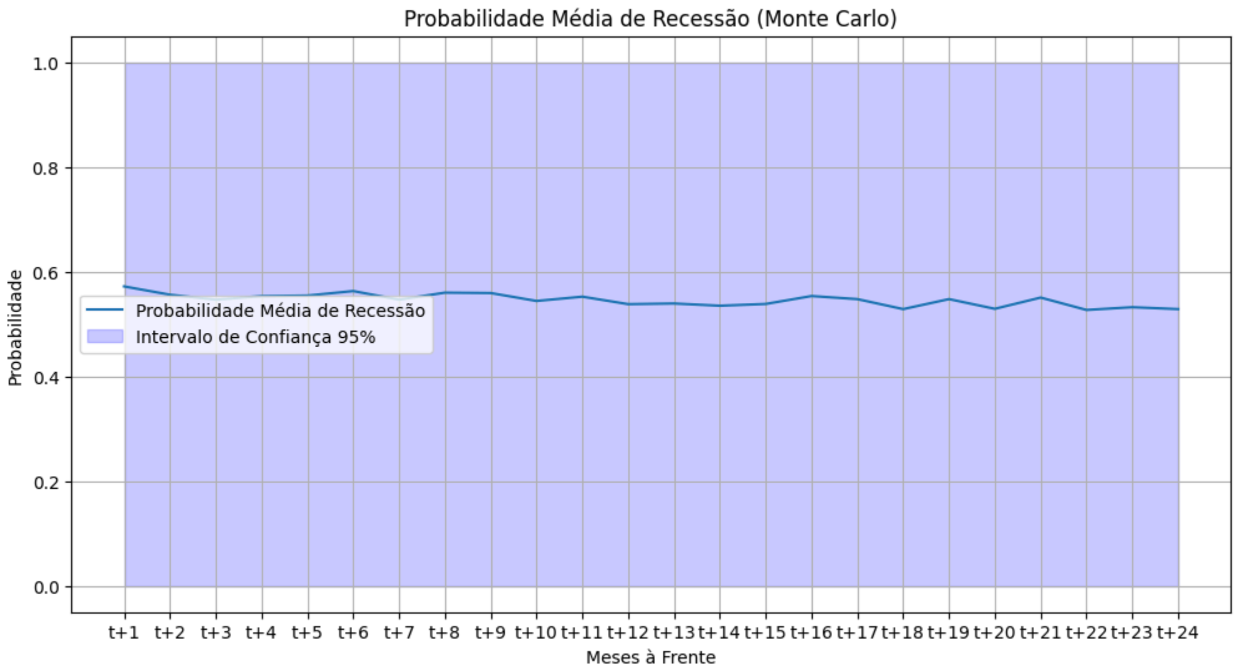
These metrics indicate that the model correctly identifies most recession episodes, with well-calibrated predictions and consistent probabilities. The main advantage lies in its ability to generate probabilistic future scenarios instead of a single deterministic forecast, providing a more realistic risk analysis tool.

In [19]:
```python
from monte_carlo_recession import monte_carlo_recession
out = monte_carlo_recession(
    resampled_data,
    target_col="recession_flag",
    feature_cols=["Point_CLI", "G0M910", "PCM910", "USSLIND", "GDP per capita (current US$)"],
    horizon=24,
    n_sim=2000
)

print(out.prob_mean)
print("P(any recession in horizon):", out.prob_any_recession)
```

```
horizon
t+1     0.571903
t+2     0.556056
t+3     0.546435
t+4     0.553388
t+5     0.554373
t+6     0.563031
t+7     0.546500
t+8     0.559971
t+9     0.559232
t+10    0.544104
t+11    0.552294
t+12    0.538001
t+13    0.539247
t+14    0.534991
t+15    0.538376
t+16    0.553499
t+17    0.547610
t+18    0.528538
t+19    0.547641
t+20    0.529095
t+21    0.550561
t+22    0.526943
t+23    0.532167
t+24    0.528537
Name: P(recession), dtype: float64
P(any recession in horizon): 1.0
```

Source: Elaborated by the author (using matplotlib.pyplot)



Source: Elaboration by the author (using matplotlib.pyplot)

As illustrated in the graph (and in the dispersion of the estimated probabilities) for the month-ahead recession forecasts from t+1 to t+24, the values remain at a relatively stable level—ranging approximately between 0.52 and 0.57—even as the forecasting horizon increases. This slight oscillation suggests that the model maintains persistently high predictions, which indicates stability but also a potential dynamic rigidity, since the temporal behavior does not respond strongly to different projections of macroeconomic variables.

This pattern is further reflected in the confidence interval, which exhibits a wide range, signaling substantial uncertainty in the projections. This characteristic is directly linked to a key challenge: probabilistic calibration under stress scenarios. The amplitude of the interval, combined with the consistently elevated probabilities across all months, indicates that the model tends to overestimate the persistence of recession risk over longer horizons. It is highly sensitive to the initial conditions used in each simulation—particularly when the historical dataset includes episodes of severe contraction, such as those observed in 2008 or 2020. Thus, even with thousands of simulations (for example, n = 2000, as used in this illustrative experiment), part of the variability remains unexplained.

Taken together, the results reinforce that although the model has practical value for estimating recession probabilities, there remains room for critical evaluation, especially regarding temporal stability, calibration under shocks (the main difference between the two recessions), and discrimination between distinct macroeconomic regimes and cyclical dynamics.

5.1 Validation and Robustness Tests

To verify stability and avoid overfitting, three main tests were applied: **Cross-Validation**, **Stress Testing**, and **Rolling Validation**.

In the **Cross-Validation**, using temporal splits with *TimeSeriesSplit*, an average **Brier Score** of 0.423 was obtained, indicating moderate calibration of predictions.
In this specific case, the **AUC** could not be computed in some partitions due to the lack of class variation, revealing the model's difficulty in discriminating against rare recessions.

In **Stress Testing**, three scenarios were simulated:

- **Mild Recession:** probability of recession = 1.000
- **Severe Recession:** probability of recession = 1.000
- **Boom:** probability of recession = 1.000

The deterministic behavior (total probability in all cases) revealed a lack of sensitivity to extreme variations and possible overfitting to historical patterns.

Finally, the **Rolling Validation** method was also applied. In the rolling validation, the average **Brier Score** was 0.4338, with strong variability (0.0 to 1.0), demonstrating temporal instability.

These results indicate that the model performs well in some time windows but loses performance in others—suggesting that its calibration must be improved for greater consistency over time.

5.2 Overall Perspective

The experiments showed that the Monte Carlo model is effective in generating recession probabilities, but still faces challenges regarding:

- Temporal stability;
- Probabilistic calibration under stress scenarios;
- Discrimination between distinct economic regimes.

However, its overall performance is superior to the linear and non-temporal models tested, reinforcing its practical utility as a tool for macroeconomic simulation and risk analysis.

Thus, the integration of Monte Carlo Simulation with probabilistic econometric models has proven to be a promising approach for predicting recessions based on the real uncertainties of the economic system.

Despite the observed limitations in terms of stability and extreme sensitivity, the model achieved consistent and interpretable results, confirming its relevance as a robust analytical tool and as a solid methodological foundation for future applications in macroeconomic forecasting and risk analysis.

6.0 Conclusion

The purpose of this research was to investigate the unpredictability of economic recessions through the application of Monte Carlo Simulation, integrating macroeconomic variables, probabilistic models, and the interpretation of two distinct recession episodes. The study began with the premise that recessions are complex and nonlinear phenomena influenced by stochastic factors, making it inadequate to rely exclusively on isolated indicators or deterministic forecasting methods. Based on this understanding, the research sought to determine whether the random nature of economic crises could be captured through a probabilistic and dynamic approach—an aspect still seldom explored in the current scientific literature.

Throughout the development of the study, central concepts from economic literature were examined, especially business cycles, leading indicators, and the inherently unpredictable nature of crises, as well as essential statistical foundations, including stochastic variables, probabilistic models, and the logic behind Monte Carlo Simulation. The theoretical framework, supported by authors such as Taleb (2007), Smirnov (2011), Estrella and Mishkin (1998), among others, made it possible to conclude that recessions are often characterized—and frequently triggered—by external shocks that exceed the explanatory capacity of traditional indicators.

Based on this premise and the literature, the modeling developed in this research integrated three main components: (i) a Probit model to estimate instantaneous recession probabilities, (ii) a VAR(1) model to capture temporal dependence among indicators, and (iii) Monte Carlo Simulation as an essential tool to generate future scenarios and evaluate aggregate risks through multiple “rounds” (simulations). This integration proved adequate for representing macroeconomic dynamics, enabling the generation of probability distributions and more comprehensive risk measures than conventional point forecasts. It also provided a distinct analytical perspective on recession events.

The results revealed that the model showed strong initial performance, suggesting effectiveness in identifying recessionary episodes. However, robustness tests (cross-validation, stress testing, and rolling validation) highlighted important limitations, including difficulty in distinguishing between different types of recessions and a deterministic behavior under key scenarios, such as turning points in the economic cycle. These findings suggest the presence of overfitting, demonstrating that while the model is efficient at reproducing historical patterns, it has limited ability to generalize to unobserved contexts. This reinforces the study’s initial premise: economic recessions possess a deeply stochastic nature and depend on multiple combinations of internal and external factors, making precise prediction inherently difficult.

Despite these limitations, Monte Carlo Simulation proved to be a promising methodological alternative, capable of incorporating real uncertainties of the economic system and offering a probabilistic analysis more consistent with the complexity of macroeconomic cycles. Overall, this study contributes to expanding the understanding of the unpredictability of recessions and empirically demonstrates that dynamic probabilistic methods can complement—and in some cases outperform—conventional economic forecasting models. The research also shows that stochastic models require continuous refinement, especially in terms of probabilistic calibration, treatment of extreme events, and temporal stability of simulations.

Looking ahead, there are several promising avenues for future research:

- expanding the dataset with higher-frequency and broader indicators;

- adopting hybrid models integrated with Monte Carlo Simulation, such as Bayesian VAR, Dynamic Probit, and regime-switching models;
- applying more regularization techniques to reduce overfitting;
- incorporating deep learning or more computationally intensive methods to enhance the simulation process.

In summary, although no model can forecast recessions with absolute certainty—given the very nature of the “black swans” that permeate economic activity—Monte Carlo Simulation offers a solid and potentially transformative methodological pathway for the probabilistic analysis and interpretation of economic crises. This study therefore contributes both to academic and scientific advancement and to practical, day-to-day analytical efforts aimed at understanding contemporary macroeconomic risks.

7.0 References

TALEB, Nassim Nicholas . **The Black Swan**. United States: Random House, 2007. ISBN 978-1400063512 (U.S.).

SMIRNOV, Sergey V. **Those unpredictable recessions.** SSRN Electronic Journal, 1 Jan 2011. Available at: <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1996991>.

KOLLURU, Mythili e HYAMS-SSEKASI, Denis e RAO, K.V.Ch.Madhu Sudhana. **A study of global recession recovery strategies in highly ranked GDP EU countries**. Economics, v. 9, n. 1, p. 85–105, 1 Jun 2021. Available at: <https://doi.org/10.2478/eoik-2021-0011>.

ESTRELLA, Arturo e MISHKIN, Frederic S. **Predicting U.S. recessions: Financial variables as leading indicators**. The Review of Economics and Statistics, v. 80, n. 1, p. 45–61, 1 Fev 1998. Available at: <https://doi.org/10.1162/003465398557320>.

GRININ, Leonid e KOROTAYEV, Andrey e TAUSCH, Arno. **Economic cycles, crises, and the global periphery**. [S.l: s.n.], 2016. Available at: <https://doi.org/10.1007/978-3-319-41262-7>.

Kuznets, S. **Secular movements in production and prices. Their nature and their bearing upon cyclical fluctuations**.Houghton Mifflin, Boston, MA. 1930

Sobre ciclos econômicos | IBRE. Available at: <https://portalibre.fgv.br/sobre-ciclos-economicos>.

MIT OPENCOURSEWARE. **1. Introduction to statistics**. YouTube. [S.l: s.n.]. Available at: <https://www.youtube.com/watch?v=VPZD_aij8H0>. Acesso em: 9 mar 2025. , 30 Out 2017

RISSANEN, Jorma. **Stochastic complexity and modeling**. The Annals of Statistics, v. 14, n. 3, p. 1080–1100, 1986. Available at: <https://www.jstor.org/stable/3035559>.

Lane, D., Scott, D., Hebl, M., Guerra, R., Osherson, D., & Zimmer, H. (2003). **Introduction to statistics** (pp. 273-284). David Lane.

CERÁVOLO, Eduardo Fujiwara e HOCHHEIM, Norberto. **Simulação de Monte Carlo considerando a dependência entre variáveis em análise de investimentos imobiliários**. Revista Valorem, v. 1, n. 1, p. 92–107, 1 Jan 2022. Available at: <https://doi.org/10.29327/2290393.1.1-8>.

DE OLIVEIRA, Marcos Roberto Gois e DE MEDEIROS NETO, Luiz Borges. **SIMULAÇÃO DE MONTE CARLO e VALUATION: UMA ABORDAGEM ESTOCÁSTICA**. Revista De Gestão, v. 19, n. 3, p. 493–512, 1 Jan 2012. Available at: <https://doi.org/10.5700/rege474>.

Leonard-Baxton, D. (1990). **A dual methodology for case studies: Synergistic use of a longitudinal single site with replicated multiple sites**. Organization Science, 1(3), 248-266.

DE TOLEDO KRÜCKEN PEREIRA, Laïs e GODOY, Dalva Maria Alves e TERÇARIOL, Denise. **Estudo de caso como procedimento de pesquisa científica: reflexão a partir da clínica fonoaudiológica**. Psicologia Reflexão E Crítica, v. 22, n. 3, p. 422–429, 1 Jan 2009. Available at: <https://doi.org/10.1590/s0102-79722009000300013>.

DA SILVA, Glênio Oliveira e DE OLIVEIRA, Guilherme Saramago e DA SILVA, Michele Maria. **ESTUDO DE CASO ÚNICO: UMA ESTRATÉGIA DE PESQUISA**. Available at: <https://revistaprisma.emnuvens.com.br/prisma/article/view/44>

DE LUNETTA E RODRIGUES GUERRA, Avaetê. **METODOLOGIA DA PESQUISA CIENTÍFICA e ACADÊMICA**. Zenodo (CERN European Organization for Nuclear Research), 12 Ago 2023. Available at: <https://zenodo.org/record/8240361>.

SAMPAIO, Tuane Bazanella. **Metodologia da Pesquisa**. 2022. Available at: <http://repositorio.ufsm.br/handle/1/26138>.

STAKE, R. **The art of case study research**. Thousand Oaks, CA: Sage Publications, 1995

WENTZEL, Marina. **China teve em 2007 maior crescimento em 13 anos**. Available at: <https://www.bbc.com/portuguese/reporterbbc/story/2008/01/080124_chinacrescimento_mw?>.

SILVA, João Pedro. **Análise da Crise Financeira de 2008**. Revista Brasileira de Economia, Rio de Janeiro, v. 70, n. 2, p. 20-35, abr. 2019.

WORLD BANK GROUP. **The Global Economic outlook During the COVID-19 Pandemic: A changed world**. Available at: <https://www.worldbank.org/en/news/feature/2020/06/08/the-global-economic-outlook-during-the-covid-19-pandemic-a-changed-world>.

ESTEMBER, Rene D. e MARAÑA, Michael John R. **Forecasting of Stock Prices Using Brownian Motion – Monte Carlo Simulation.** https://ieomsociety.org/ieom_2016/pdfs/192.pdf. [S.I.]: IEOM Society International, [S.d.]. Available at: <https://ieomsociety.org/ieom_2016/pdfs/192.pdf>.

PAŽICK, Martin. **Stock price simulation using bootstrap and Monte Carlo**. Available at: <https://www.ceeol.com/search/article-detail?id=546025>.

XIANG, Jeremy Ng Phak e VELU, Shubashini Rathina e ZYGIARIS, Sotirios. **Monte Carlo simulation prediction of stock prices**. 2021 14th International Conference on Developments in eSystems Engineering (DeSE), p. 212–216, 7 Dez 2021. Available at: <https://doi.org/10.1109/dese54285.2021.9719349>.

US leading indicators. Available at: <https://www.conference-board.org/topics/us-leading-indicators>.

LEI and CEI March 2020 | IBRE. Available at: <https://portalibre.fgv.br/en/press-releases/lei-and-cei-march-2020>.