UNIVERSIDADE DE SÃO PAULO

FACULDADE DE ECONOMIA, ADMINISTRAÇÃO, CONTABILIDADE E ATUÁRIA

DEPARTAMENTO DE ADMINISTRAÇÃO

PROGRAMA DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO

JARDEL AUGUSTO GOMES RODRIGUES ALVES

A cross-country analysis on unemployment:
Past, present and a scenario-based foresight into the future

Uma análise *cross-country* sobre o desemprego:
Passado, presente e uma prospecção de futuro baseada em cenários

São Paulo
2024

Prof. Dr. Carlos Gilberto Carlotti Júnior

Reitor da Universidade de São Paulo

Profa. Dra. Maria Dolores Montoya Diaz

Diretora da Faculdade de Economia, Administração, Contabilidade e Atuária

Prof. Dr. João Maurício Gama Boaventura

Chefe do Departamento de Administração

Prof. Dr. Felipe Mendes Borini

Coordenador do Programa de Pós-Graduação em Administração

JARDEL AUGUSTO GOMES RODRIGUES ALVES

A cross-country analysis on unemployment:
Past, present and a scenario-based foresight into the future

Uma análise *cross-country* sobre o desemprego:
Passado, presente e uma prospecção de futuro baseada em cenários

Tese apresentada ao Programa de Pós-Graduação em Administração do Departamento de Administração da Faculdade de Economia, Administração, Contabilidade e Atuária da Universidade de São Paulo, como requisito parcial para a obtenção do título de Doutor em Ciências.

Área de Concentração: Administração.

Orientadora: Profa. Dra. Renata Giovinazzo Spers.

Versão original

São Paulo - SP
2024

Catalogação na Publicação (CIP)
Ficha Catalográfica com dados inseridos pelo autor

Alves, Jardel Augusto Gomes Rodrigues.
A cross-country analysis on unemployment: Past, present and a scenario based foresight into the future / Jardel Augusto Gomes Rodrigues Alves. - São Paulo, 2024.
266 p.

Tese (Doutorado) - Universidade de São Paulo, 2024.
Orientadora: Renata Giovinazzo Spers.

1. Unemployment. 2. Labour Market. 3. Forecast. 4. Foresight. 5. BRICS. I. Universidade de São Paulo. Faculdade de Economia, Administração, Contabilidade e Atuária. II. Título.

ACKNOWLEDGMENTS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

ABSTRACT

Alves, J. A. G. R. (2024). A cross-country analysis on unemployment: Past, present and a scenario-based foresight into the future. (Doctoral Thesis). Faculdade de Economia, Administração, Contabilidade e Atuária, Universidade de São Paulo, São Paulo.

Following multiple studies approach this thesis has as its main objective a proposition of a scenario-based foresight for the unemployment rates in a cross-country analysis, that have as main scope Brazil, Russia, India, China, and South Africa, the BRICS countries. This purpose of research and the scenarios proposed, as well the path to the building of them, will come to answer the research question: How could unemployment rates evolve in a 10-year forecasting horizon for BRICS countries? Two studies precede, our scenario-based foresight. First, a bibliometric analysis is applied to identify the emerging topics on unemployment-related academic literature. From this study we extract some themes that could be used as possible determinants that could explain unemployment rates composition. Second study builds-up from the first one, having some potential determinants for unemployment we use a Vector Error Correction Model (VECM) to identify which of determinants are more influential on unemployment rates configuration. From these studies' findings, we move forward refining our scope of analysis to BRICS nations. Using each country's historical unemployment rates, we apply quantitative forecasting methods, namely: Artificial Neural Networks (ANN), Autoregressive Integrated Moving Average (ARIMA), Exponential Smoothing Technique (ETS), and Seasonal and Trend Decomposing using Loess (STL). Using these methods, we forecast possible future unemployment levels in a 10-years into the future timespan. Results extracted from each method are the basis to the qualitative scenario forecasting technique that is used to present three potential scenarios for unemployment rates in BRICS countries in a 10-years into the future. From these scenarios and throughout all analyses we present and discuss in this thesis, it is hoped that we may advance academic research exploring a not usual scope of unemployment-related studies whereas we may offer to BRICS policy- and decision-makers some anticipation about what might come into their futures regarding unemployment, labour market, and labour-relationships in order to enable a better informed policies proposition that could envision a better future for all those interested economic and socially involved.

Keywords: Unemployment; Unemployment rates; Labour Market; Forecast; Foresight; BRICS.

RESUMO

Alves, J. A. G. R. (2024). Uma análise cross-country sobre o desemprego: Passado, presente e uma prospecção de futuro baseada em cenários. (Tese de Doutorado). Faculdade de Economia, Administração, Contabilidade e Atuária, Universidade de São Paulo, São Paulo.

Seguindo uma abordagem de múltiplos estudos, esta tese tem como objetivo principal a proposição de uma previsão baseada em cenários para as taxas de desemprego em uma análise *cross-country*, que tem como escopo principal o Brasil, a Rússia, a Índia, a China e a África do Sul, países do BRICS. Este propósito de investigação e os cenários propostos, bem como os caminhos que levam até a sua construção, virão a responder à questão de pesquisa: Como poderão evoluir as taxas de desemprego num horizonte de 10 anos no futuro para os países do BRICS? Dois estudos precedem a nossa previsão baseada em cenários. Primeiro, é aplicada uma análise bibliométrica para identificar os temas emergentes na literatura académica relacionada a o desemprego. Deste estudo extraímos alguns temas que poderiam ser utilizados como potenciais determinantes para explicar a composição das taxas de desemprego. O segundo estudo baseia-se no primeiro, tendo alguns potenciais determinantes para explicar o desemprego, utilizamos um Modelo de Correção de Erros Vetoriais (VECM) para identificar quais determinantes são mais ou menos influentes na configuração final das taxas de desemprego. A partir das conclusões destes dois estudos, avançamos refinando o nosso âmbito de análise para as nações do BRICS. Utilizando as taxas históricas de desemprego de cada país, aplicamos métodos quantitativos de *forecast*, nomeadamente: Redes Neurais Artificiais (RNA), Média Móvel Integrada Autorregressiva (ARIMA), Técnica de Suavização Exponencial (ETS) e Decomposição Sazonal e de Tendência usando Loess (STL). Usando estes métodos aplicamos a predição, via *forecast*, sobre o potencial desemprego no futuro dos países analisados para 10 anos adiante. Os resultados extraídos de cada método constituem a base para a técnica qualitativa de cenários preditivos, que utilizamos para apresentar três cenários potenciais para as taxas de desemprego nos países do BRICS daqui a 10 anos. A partir destes cenários, e ao longo de todas as análises que apresentamos e discutimos nesta tese, espera-se que possamos avançar na investigação académica explorando um âmbito não habitual de estudos relacionados com o desemprego, ao mesmo tempo que podemos oferecer aos decisores políticos e gestores dos países do BRICS alguma antecipação sobre o que podem enfrentar em seu futuro no que diz respeito ao desemprego, ao mercado de trabalho e às relações laborais, a fim de permitir uma proposição de políticas mais bem informada que possa assegurar um futuro melhor para todos as partes interessadas.

Keywords: Desemprego; Taxas de desemprego; Mercado de trabalho; Forecast; Foresight; BRICS.

LIST OF FIGURES

LIST OF TABLES

INDEX

1. INTRODUCTION

1.1. Context.

Unemployment understanding is largely intuitive. It is indeed a reality that almost everyone seems to know or probably has experienced in their own lives. However, as it should be perceived during this initial chapter as well throughout this thesis, unemployment is multifaceted and have many complexities within. As a referential conceptual starting point, this thesis will follow the conceptualization of unemployment proposed by International Labour Organization (ILO, 2022) (Unemployment rates: Definition).

ILO defines unemployment as the share of people on the labour force that is currently without a formal job-allocation but is available and actively seeking for an opportunity. Referring to a share of jobless individuals, unemployment is usually aligned to unemployment rates (UR), probably the best-known labour market indicator. According to International Labour Organization description, unemployment rate is an index that may reflects potential inability of an economy to generate employment opportunities for individuals that desires to work but are not having opportunities to do that.

The World Bank states a very similar definition for unemployment rate, defining it as a percentage of the total labour force (which includes employed people as well) that is currently without a job but seeking for one in a recent period, still wanting to be inserted on a job-opportunity. Important to mention that any rate about the labour force, employed and unemployed, involves all the "active people", which includes cases of people that were relieved of their previous occupation and those who may voluntarily quit from a previous function.

Both ILO and World Bank definitions are useful and the ones we use as guideline for what we assume as unemployment in a broader sense. It is as well from these sources that we extract most of the data information we will later be analysing on this thesis development. Unemployment rate is indeed a useful indicator to measure labour market performance. Nonetheless, it may not be assumed as an absolute factor considering that labour market influences and is influenced by many other variables.

Unemployment, and its measured rates, is potentially the most used indicator of job-search activity (Shimer, 2005). Job-search and the non-fulfilment of people wanting to work and job opportunities is not a recent phenomenon on Western economies, in fact, it has been worryingly high for many years (Clark & Oswald, 1994). Although we acknowledge earlier that unemployment is not the only socio-economic indicator that permeates labour market, it is possibly the common disfunction that exist in all countries of the world, affecting people's living standard and nations' economic status (Germay, 2016).

Some situations may exemplify these potential damaging effects that unemployment may bring to a nation's economy. In United States for example, after COVID-19 was declared a national emergency, the unemployment rate increased to historic levels, peaking in April 2020 at 14.8% (Nguyen et al., 2022). In China, unemployment is one of the most important social and economic problems since the mid-1990s. From 1995 to 1999, a relatively short time span, around 15 million of individuals in the Chinese labour force were relieved from their occupations because their work was assumed as redundant with the technological advances in the country (Meng, 2003).

If unemployment is an emergent and not necessarily new problem for two of the most important and developed countries of the world, it would be naive to not presume an even worse scenario for underdeveloped economies. Brazil, for example, had a historic unemployment rate registered in 2021 at 14.7%. Highest value since IBGE's (Brazilian Institute of Geography and Statistics) unemployment series started to be measured by the institute metrics (McGeever, 2021). Going to other regions, South Africa, after their first democratic elections in 1994 experienced an exponential increase of unemployment, going from an already high rate of 15.6% in 1995 to 30.3% in 2021 (Banerjee et al., 2008).

Most recently, worldwide economy suffered, and in a manner still suffers, with large macroeconomic impacts from COVID-19 pandemic. Within this scenario it is reasonable to assume that alarming number of unemployed people could be worsened by the unforeseen pandemic effects (Chitiga et al., 2021). It is on this context that our research emerges, understanding how unemployment occurred over the years, what is the current landscape and foresight what might occur in the future, is an integral part on how economies, countries and regions could be prepared to deal with this variable.

The achieving of a better understanding of unemployment besides the possible theoretical macroeconomic insights could also implicate to relevant managerial contributions. When identified the main determinants to unemployment rates in the present, companies can perceive in which fronts they must dedicate more effort whereas in parallel could be mitigating problems in future unemployment rates. In this twofold avenue, considering both macroeconomic understanding and potential managerial contribution, it is our expectation that this thesis could be beneficial in as much as possible fields.

1.2. Relevance and expected contribution.

Unemployment is not a recent track for academic studies. Since the mid-1950s, with the seminal study of unemployment and wage rates in the United Kingdom made by Phillips (1958), many authors have been dedicating themselves to better understand the unemployment. Todaro (1969) analysed unemployment as an impulse for migration; Mortensen & Pissarides (1994) tried to explain job-creation and job-destruction cyclical behaviour; even in health studies the phenomenon had been assessed, as in Paul & Moser (2009), relating unemployment and mental health-issues.

Indeed, in recent years, according to a search on Web of Science database, most studies have focused on the health effects that unemployment status may generate for individuals. Impacts of COVID-19 connecting to unemployment, mental health, socioeconomic outcomes among other topics have also been addressed (Kawohl & Nordt, 2020; Galea & Abdalla, 2020; Blustein et al., 2020). Automatization of jobs, as we mentioned happened in China (Meng, 2003), is a growing subject of interest (McClure, 2018; van Roy, Vértesy, & Vivarelli, 2018).

It is possible to say that unemployment has been a persistent theme of research in diversified fields. Discussions on the theme have been evolving over decades and will possibly remain as an important subject of analyses. Rates of unemployed people have been increasing in a constant rhythm and have spiked in the pandemic period all over the world. Even in developed and more structured economies, consider the United States and China as example, unemployment is an urgent problem to be largely discussed and assessed (Meng, 2003; Nguyen et al., 2022).

On these discussions and assessments and considering the relevance of unemployment as a broader theme, that this thesis intends to be inserted and contribute. We intend to perform a supranational analysis of unemployment composition but later we refine our scope to a lesser analysed on unemployment-related studies, which is largely focused on USA and European nations, we move for the BRICS (Brazil, Russia, India, China, and South Africa) countries analyses. We intend to unveil a relatively unexplored group of countries, assessing their historical unemployment rates, their current indexes, and foresight the possible future for jobs BRICS.

There is on the literature studies that assess BRICS countries unemployment individually, as a collective of nations, as far we scanned the literature, we do not find similar study as the one we intend to propose here which is a potential gap that we may resolve whereas contributing to what already exists about these nations. We acknowledge that absence of studies per se does not justify the development of a research. However, we resort to a potential opportunity to contribute on a scope that is underdeveloped in comparison with other more scrutinized nation-states.

With our study proposition the aim is not to be totally innovative and to solve all problematics regarding unemployment. Expectative is to be a useful tool of contribution to a better understanding of this phenomenon for countries that are not usually focused on the theme. Our guiding process is to revisit what already happened about unemployment, which is the current state of this phenomenon and what might potentially happen in a near future.

For delimitation purposes, what we may define as aim for this research is: The proposition of a scenario-based foresight for unemployment rates in a cross-country analysis, that have as main scope Brazil, Russia, India, China, and South Africa. Our premise involves revisit past unemployment, assess current unemployment rates composition, and have a glimpse about how this variable may evolve into a 10-years into the future forecast. Our analyses are at worldwide levels up to the moment we write the scenarios when we refine our scope to BRICS countries.

Although the scenarios to be presented are directly related to Brazil, Russia, India, China, and South Africa, many of their problems are supranational and alike other countries. We believe that the scenarios, and the extensive process that leads to their writing process, could be helpful and replicable for different country’s realities. General purpose is to offer solid contributions for macroeconomic understanding of unemployment as well managerial insights to improve organizational and public policies to better cope with this phenomenon.

Not necessarily we intend to propose here a new concept or theory by itself. Our expectation is that from what we are presenting on this thesis will emerge a somewhat consistent comprehension about a complex and important phenomenon as unemployment. Understanding the past, ascertaining the present and foreseeing the future of unemployment, we believe that contributions by this thesis could go beyond academic purposes, suggesting information for "real-life" situations and serving as a guidance to the interested-on labour and unemployment themes.

1.3. Research question.

Considering the introduced up to this point, this thesis aims to answer the following research question: **How could unemployment rates evolve in a 10-year forecasting horizon for BRICS countries?** In response to this question, we are following a scenario forecasting proceeding that has as main purpose: **The proposition of a scenario-based foresight for unemployment rates in a cross-country analysis, that have as main scope Brazil, Russia, India, China, and South Africa**.

Our scenario-based proposition is the final product and is built upon other within this thesis studies. In other words, to propose feasible and reliable scenarios, which resolves our defined research question, we proceed first with an extensive unemployment-related literature bibliometrics analysis; from where we extract some emergent topics discussed on the field.

Secondly, we assess determinants that composes unemployment rates, some of these determinants are directly extracted from the first study some are well-anchored on the literature; overall we understand which factors influences unemployment and in which direction (if positively or negatively). After these two studies we proceed to forecasting techniques to project 10-years into the future unemployment and later we write scenarios forecasting from all results we have in our hands.

To summarize, this thesis is organized in a multiple studies approach, composed with three complementary chapters all dealing with the broader theme of interest in our study, unemployment, and all helping to address the main purpose of this work and resolve the proposed research question. Although the reading of the studies in sequence contributes to a better understanding of our main objective, the idea is that they could be read independently.

In addition to that, it is worth of mentioning that the soon to be presented three studies division, eventually implicate on a few redundancies on texts, citations, or references. As we reinforce, all chapters and studies are dealing with the larger and extensive theme unemployment is. Nonetheless, it could be perceived among the chapters, notably in the introductory sections, some recurrency of concepts and definitions.

Chapter that follows this introduction section aims to present an overview of the unemployment-related academic literature. Applying a bibliometric analysis, the proposition is to analyse studies that mentions unemployment or unemployment rate covering all the timespan available using two of the most reliable and used databases of scientific production: Scopus and Web of Science.

As a result of this bibliometric analysis the purpose is to have an identification about main topics and themes related to unemployment that are currently discussed (emergent) and the ones that endures on this research field. By doing this, we believe to be possible to extract some determinants that will be used in the third chapter as potential determinants that composes unemployment rates.

Research question that guides our second chapter is: *What are the main topics discussed in the unemployment-related academic literature?* Complementary questions that therefore will lead to secondary objectives are: Where are, geographically, the most frequently cited documents and authors discussing unemployment? What the most relevant sources receiving unemployment studies?

One of the concerns that lead to this study proposition is the belief that unemployment is a tricky theme although at the same time is a common-sense concept. We believe that observing some tendencies on the unemployment-related literature it would be possible to consolidate a reasonable thematic starting point to proceed on our further studies assessments.

After unveiling the main topics related to unemployment extracted from the bibliometric analysis, the following study has as purpose the identification of which factors are influential on the unemployment rates composition. Using a Vector Error Correction Model (VECM) application the aim is to ascertain what are the main determinants composing unemployment. VECM proposes an equation with a dependent variable, in our case the unemployment rate, and a series of independent factors to identify which of them are most influential.

It is noticeable that this third chapter derives from the second and have the following main research question: *What are the main determinants that composes the unemployment rates?* Using VECM and a clustering technique to better assess data available, the premise on this chapter is to develop a cross-country analysis to identify main determinants composing unemployment rates considering within a sample of 154 countries in all regions of the world.

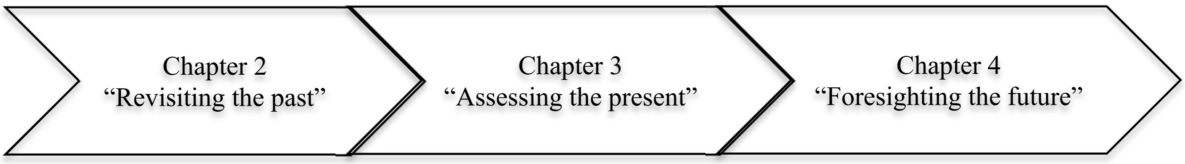
From this macro-level and supranational analyses, we finally refine our scope for Brazil, Russia, India, China, and South Africa (BRICS countries) on the fourth chapter of our thesis. On chapter number four and study number three we proceed to effectively forecast and project potential unemployment levels in a 10-years into the future horizon. Characteristics of BRICS nations will be further depicted and the reasoning about this restraining on country's are to be better explained.

Expected objective on this essay being the foresight of unemployment in the BRICS, the fourth chapter are presented to resolve the following research question: *What are the foreseeable tendencies for unemployment in the BRICS countries?* To answer this question, we apply quantitative forecasting methods: Artificial Neural Networks (ANN), Autoregressive Integrated Moving Average (ARIMA), Exponential Smoothing Technique (ETS), and Seasonal and Trend Decomposing using Loess (STL). Results from each method are the basis to the scenario forecasting writing and presentation.

Considering the presented regarding the chapter four, complementary question that could also be responded by this chapter is: What are the possible paths for BRICS stakeholders to pursue to be better prepared and to deal with unemployment rates in the future? Although chapter four, in isolation, could offer the answer to the broader thesis' research question (How could unemployment rates evolve in a 10-year forecasting horizon for BRICS countries?) we reinforce how all studies are connected. Inputs from one are primordial for the following and the three, collectively, presents the deepening on understand unemployment we aim to present.

For a general illustration about the chapter underpinning themes and how these topics and chapters interact in a way to conclude about unemployment complexities, up to the glimpse into the future. Sequential relationship between the chapters is depicted in figure 1.

Figure 1
Relation of the articles presented in the thesis.
(Elaborated by the author).



Note: The sequential relationship illustrated in the figure is an assumption. As mentioned earlier, the reading of the chapters is presumed to be functional separately. Names presented on the figure is illustrative of chapter's idea not necessarily the names that will be used.

1.4. Research shortcomings.

While this research aims to contribute to the understanding of a broader perspective about unemployment and at the end to anticipate how this phenomenon could evolve into the future of BRICS countries, it is crucial to recognize some potential shortcomings. First and foremost, the reliance on unemployment rates as the primary indicator might oversimplify the complex nature of unemployment. Unemployment is a multifaceted phenomenon influenced by various socio-economic factors, and a sole focus on rates may neglect the nuances of underemployment, informal employment, and the quality of jobs.

Another limitation pertains to the availability and reliability of data. Despite ILO and World Bank data are widely used, variations in data collection methodologies among different countries could introduce inconsistencies. Additionally, the timeframe of this study's forecast, which spans a decade, may pose challenges in anticipate accurate and up-to-date information. Although if we expand our forecasting in a large time span, results would be potentially deceitful.

Furthermore, the research's scope, which initially encompasses a broad supranational analysis and later narrows down to BRICS nations, may overlook unique national intricacies within each member country. We decide to delimit our scope using a diversified group of countries that still share some similarities whereas if we maintain the worldwide analysis into scenario forecasting writing we could have a potential unmanageable task. Even restraining to five countries we acknowledge that we cannot be able to fully capture regional challenges and policy implications inside each BRICS nation.

The forecasting methods employed in the research, including ANN, ARIMA, ETS, and STL, while advanced, come with inherent uncertainties. Predicting unemployment over a ten-year horizon involves numerous variables and assumptions, and unexpected events or policy shifts could significantly impact the accuracy of the forecasts we present. We are focusing on unemployment during our quantitative proceedings however at the subjective scenario-based forecasting we resort to other factors and variables previously discussed, particularly on the chapter 3.

Additionally, the study's emphasis on quantitative forecasting methods may inadvertently neglect qualitative aspects of unemployment, such as the cultural, political, or institutional factors influencing labour markets. A more holistic approach, combining quantitative and qualitative analyses, could offer a more comprehensive understanding of the complexities surrounding unemployment. However, if we insert ourselves on these holistic discussions, we would potentially apart our research from the economic and managerial perspective that are primordial for the locus we are writing this essay.

In conclusion, while this thesis strives to contribute valuable insights about past, present and potential future of unemployment, the latter focusing on BRICS nations reality, it is essential to approach our findings with a recognition of these limitations we present and potential others that escapes from this topic. The dynamic and multifaceted nature of unemployment demands an awareness of inherent uncertainties and challenges associated with forecasting and analysing such a complex socio-economic phenomenon.

1.5. Research connectivity.

Table 1 presents the binding between the following three chapters and how they isolated are supposed to be conducted and concluded. This table is a complement and further assessment by the early presented figure 1. Our final intention is that the chapters 2, 3, and 4 connectivity may satisfactorily respond our main research question and fulfil this thesis' main objective.

Table 1
Connections between studies and the thesis.
(Elaborated by the author).

RESEARCH QUESTION	THESIS OBJECTIVE	CHAPTERS R.Q.	CHAPTERS OBJECTIVE	METHOD	DATA SOURCE AND ANALYSIS
How could unemployment rates evolve in a 10-year forecasting horizon for BRICS countries?	Proposition of a scenario-based foresight for unemployment rates in a cross-country analysis, that have as main scope Brazil, Russia, India, China, and South Africa.	What are the main topics discussed in the unemployment-related academic literature?	To have an identification of main topics and themes related to unemployment currently discussed (emergent) and the ones that endures on the field.	Bibliometric Analysis	Retrieved data from Scopus and Web of Science databases and analysed using Software R-Studio
		What are the main determinants that composes the unemployment rates?	Develop a supranational cross-country analysis to identify main determinants composing unemployment rates	VECM	Secondary data is available on World Bank and analysed using Software R-Studio.
		What are the foreseeable tendencies for unemployment in the BRICS countries?	Forecast future unemployment rates and from these the proposition of potential future scenarios for foresight unemployment rates in the BRICS countries.	ANN ARIMA ETS STL Scenario forecasting	World Bank and ILO repositories information about unemployment rates analysed using software R-Studio.

As table 1 illustrates, each chapter that will be presented on the following have their own research problem and objectives while they also rely on different methodological techniques. Nonetheless, the idea is that although they could be read and understand on isolation exists a converging point to respond what this thesis intends. We build our thesis in this connectivity believing that all chapters together offer the most on better understanding some of the multifaced unemployment phenomenon.

2. REVIEWING UNEMPLOYMENT AND ITS DRIVING THEMES: BIBLIOMETRIC ANALYSIS ON THE SCOPUS AND WEB OF SCIENCE

2.1. Introduction.

Understanding of macroeconomic variables could represent one of the most reliable sources to have an overview about the status of a given economy or country in a period. Among these macroeconomic indicators, employment and unemployment rates are one of the main factors that proxies the wellbeing and economic momentum for a population (Simionescu, 2020). On regular basis it is possible to see in TV journals or read in the news reports of the amount of people searching for a job and dealing with the eminent problems related to unemployment.

In mid-1950s, Phillips (1958) made one of the seminal studies on unemployment theme, relating the absence of jobs and wage rates in the United Kingdom. In an essential economic-based analysis, author found evidence that changes on money wage rates varies accordingly with the highs and lows on the levels of unemployment (Phillips, 1958). Earlier than the proposition of the known as “Phillips Curve” (Phillips, 1958; Friedman, 1968), Joseph Alois Schumpeter as well discussed unemployment as a multifaceted phenomenon that could be frictional, cyclical, or structural (Schumpeter, 1939; Boianovsky & Trautwein, 2010).

More recently, other topics of unemployment have been assessing beyond the economics, bringing to table more societal lenses for discussions. Unemployment of young people is gaining more and more attention; questions of gender, considering the differences between the rates of men and women without a job is another example of a topic demanding recent attention of researchers.

According to the International Labour Organization (ILO) almost 74 million people aged 15 to 24 were looking for a job vacancy in 2014. Youth unemployment rates remained consistently high across all regions of the world and a tendency to these numbers increasing even more were already envisioned (ILO, 2015; Schmillen & Umkehrer, 2017). Taking into consideration the outcomes from COVID-19 pandemic outbreak, ILO stated that approximately 26 million jobs were lost only in Latin America and Caribbean countries during the initial years of pandemic effects (ILO, 2021).

A “Gender Gap” exists considering almost all social or economic indicators and unemployment is not an exception. The gender unemployment gap, which represents the difference between female and male without an active job, was regularly positive until the early 1980s and lowered after 1983, being again elevated during recessions periods, when usually more men lost their jobs according to Albanesi & Sahin (2018) study. The reason why this distinction by gender happens and why this gap has been relatively reduced has already been discussed in other studies but how the pandemic might have acted on this gap remains an issue to be observed.

Going back to Phillips (1958) and his studies on job opportunities and wages; considering the unemployment may occur in a frictional, a cyclical or a structural manner (Schumpeter, 1939); the growing youth unemployment rates (ILO, 2015) and the gender questions surrounding the employed and unemployed persons (Albanesi & Sahin, 2018), are just some examples about how wide the unemployment research field is and how far from being exhausted this phenomenon remains.

The assumption guiding this chapter is the consideration of complex and multifaceted phenomenon unemployment is. Furthermore, we may perceive that this is a theme recurrently discussed throughout many years and, in this sense, it seems reasonable to revisit studies that has been already done by many other researchers who contributed to this research field. Therefore, and considering the relevance and endurance of unemployment studies, the main purpose of this chapter is: To have an identification of main topics and themes related to unemployment currently discussed (emergent) and the ones that endures on the field. Aiming to respond the main chapter's research question: What are the main topics discussed in the unemployment-related academic literature?

Besides answering to the described research question, secondary objectives that should be addressed during this process includes see where are, geographically, the most frequently cited documents and authors discussing unemployment and what the most recurrent sources of scientific promotion receiving unemployment studies. Responding the main purpose, fulfilling the research question, and passing for the secondary objectives, the idea is to have as a final product for the chapter a solid number of potential determinants that could explain unemployment rates, a useful starting point for further developments aimed on this thesis.

To achieve this chapter intentions, the idea is to proceed with an application of a bibliometric analysis, providing a consolidated representation about studies that mentions unemployment or unemployment rates considering all available documents in the Scopus and Web of Science repositories due to their reliability and ease access to information. Further details about the selection of these databases and methodological proceedings will be presented in the following sections. Besides this introductory section, the chapter is structured with a literature review about unemployment; a methodological section depicting the bibliometric analysis procedure; a results and discussions topic presenting the bibliometric outputs and its findings and, finally, a conclusion topic.

2.2. Literature review.

Unemployment has been a persistent topic of discussion for a while. Considering the academic literature a basic search on the Web of Science (WoS) database filtering for the earliest mentions about the theme, it could be retrieved studies dating from the first years of 1900's. Using Scopus, the earliest documents are not far away from this same period. Macgregor (1907) is one of these early studies, discussing labour exchanges and unemployment, the author brings to debate not only economic effects of the non-existence of job opportunities, but also brings to discussion ethical problems within the job relations (Macgregor, 1907).

An initial observation of titles and abstracts in the first papers returned in both Web of Science and Scopus databases demonstrates that many of the subjects approached in the earliest unemployment studies remain recurrent nowadays. Johnson (1917), for example, way before more recent studies that relate unemployment and health issues (e.g., Paul & Moser, 2009; Kawohl & Nordt, 2020), already discussed differences on health statuses among individuals with and without jobs. Furthermore, a reasonable number of studies have been developed to better understand the unemployment problem in different countries and contexts, more intensively in Germany, United Kingdom, and United States.

These initial insights of unemployment-related literature are just an illustration that unemployment studies are persistent over the years. Within this context, we believe it is reasonable to further explore unemployment theme considering this chapter purposes as well the entirety of this thesis. Chapter continues with the following sections highlighting some topics that does not exhausts the unemployment themes but shed some light about important guiding topics.

2.2.1. Understanding Unemployment: Early and recent approaches.

Long (1942), argues that a lack of a commonly accepted concept it was probably the most difficult barrier surrounding the unemployment. Misconceptions on about the theme were largely diffused even by the most prominent economists, statistical practitioners, politics, and public. Long (1942) believed that probably the most egregious mislead was that unemployment could be solely a mathematic problem, a matter of quantities on unused labour forces and time.

Still discussing the complexity of unemployment conceptualization, the same author stated that "*It is not often fully realized that conceptual limits of unemployment are not definite boundaries, but rather wide battlefields over which economic and social philosophies are still fighting*" (Long, 1942, p. 2). Although we present the Long's argument, it is not the proposition for this chapter to present a definitive concept for unemployment; in fact, this is probably not an achievable task to meet.

Considering that and assuming the existence of more than one possible definition for unemployment, it is useful to revisit some of the more usually accepted propositions about the theme. Again, according to Long (1942), at the time of his study three main research tracks about the theme were existing. First group is composed of studies made before 1908 and that presumed unemployed people will always tend to be unemployed due to the strict wages they demand (Long, 1942).

Second group defined by Long (1942) involves the ideas on the 1910, 1920, 1930 decades. Main characteristic was an attempt to confine unemployed people as persons willing and able to work but not having the opportunity. Third groups were probably the most conflicted within the definitions of unemployment. Recognizing unemployment as a disequilibrium between supply and demand for labour, unemployment would be solved by following mathematical functions (Long, 1942).

Bearing in mind the leading thought on the previous groups of studies, the author believed that any endeavours to define unemployment should be simultaneously narrowed and broadened depending on the interest of measurement or analysis. Conceptualizations should be narrow to exclude or place in perspective abnormal economic or social conditions; while it should be broadened when observing part-time jobs, strikers and people willing to work but only in some specific circumstances (Long, 1942).

Conceptualisation evolved from the initial propositions by Macgregor (1907), Johnson (1917) and Long (1942); Jahoda (1981) found that at her study's time there was an immense number of studies on unemployment, employment, and work, emerging the necessity to consolidate all of those. This presumed chaos on concepts and approaches in work-related studies are in a way summarized by Jahoda (1981) in three basic perspectives: conservative, reformist and radical (Jahoda, 1981) lines on unemployment theme.

Conservative perception understands work, and especially unemployment, as a problem on individuals' morality. Radical's perception considers that unemployment is a systemic problem, producing alienation and exploitation of labour force, creating unemployment intentionally. Differing from the other two, Reformists understand alienation as a fact but not as intentionally imposed; rather, unemployment would be a humanitarian problem that must be dealt with public policies (Jahoda, 1981).

Overall, the understanding about the nature of work, employment and unemployment are inevitably influenced by individualistic beliefs and values. Therefore, a wide range of theories are acceptable and feasible offering to researchers interest their ow effort on select the idea that is more in line with own premises and judgments. On this thesis scope our idea is to assess what already exists in the unemployment-related literature. We do not intend to take for ourselves a definitive premise for unemployment but to recognize and identify what is emergent on this rich research field.

From what has been presented thus far I believe that at least one point seems to be relatively clear about the unemployment literature: this has been a subject of studies for a considerable long time. Studies about this topic dated back from early 1900 up to nowadays. Furthermore, recent topics being discussed have not been less complex than the earliest approaches. Technological advances on manufacture and service offers, climate change and demographic problems within a continually accelerating globalization are some of the potential key pieces within the context of the work relations in the future (Balliester & Elsheiki, 2018).

Not only these “tangible” factors could influence the future state of jobs, but a more recent phenomenon suggests that many individuals are opting to not have a job even having the opportunity. Meaning that an apparent downward trend in unemployment rates over the past 25 years considering historical data could be misleading (Barnichon & Figura, 2015). Taking this most recent complexity and observing what already is a long-standing complex phenomenon, we believe that an endeavour to achieve a consistent understanding about unemployment had to be made. Envisioning studies from the past by the bibliometric analysis is the first step we intend to perform to achieve this on the final of our thesis.

2.2.2. Bibliometric Analysis: Concepts and applications.

Bibliometric analysis, although further discussed on the methodological section of this chapter, has a short topic here on the literature review to discuss more specifically about the theoretical behind the methodological proceeding. The “bibliometrics”, as far as we known, was first used, in the Journal of Documentation as terminology proposed by Alan Pritchard. The author defined it as the application of quantitative procedures in a dataset to identify patterns in scientific discussions on a subject or area of knowledge (Pritchard, 1969; Coates, 2000).

Later, Broadus (1987) revisited some studies using the term bibliometrics and believed that previous definitions were too broad, proposed his own “*Bibliometrics is the quantitative study of physical published units, or of bibliographic units, or of the surrogates for either*” (Broadus, 1987, p. 376). Although in this thesis the Broadus’ definition is assumed as the main reference for the method, it would be noteworthy to consider some limitations highlighted by the author himself.

Broadus recognizes that his definition could be too narrow, since it does not consider George K. Zipf’s law of word occurrence (Bookstein, 1980). Still according to Broadus (1987), the counting of words is more important when the texts analysed are of great importance to the researcher’s objective, which seems to be more aligned with Zipf’s Law than with the bibliometric studies following the Broadus’ definition necessarily. Therefore, we use both concepts on the later application developed by ourselves.

Wang, Reinhilde & Stephan (2017) and Linnenluecke, Marrone & Singh (2020) proposed similar approaches as the one we intended to perform on this chapter, believing that the assessment of past literature may be a basal step on collecting existing knowledge and to examine the state of a given research stream as we are positioning this study in relating with unemployment-related literature. Donthu et al. (2021) suggests that although the bibliometric analysis has been applied with constancy in other areas, its application in management and business fields is relatively new and sometimes applied with less scientific rigor, a gap we intend to respond when this chapter is concluded.

The bibliometric analysis applied here, follows the premises of recent studies such as Verma & Gustafsson (2020), Donthu et al., (2021) and classical concepts like Pritchard (1969) and Broadus (1987). More about the methodological procedure we use is depicted on the next section of this chapter.

2.3. Methodological procedures.

2.3.1. Operating the Bibliometric Analysis.

To put in practice the analysis proposed here I relied essentially on the utilization of `bibliometrix`, a package available on a free software for statistical computing and production of graphics, the R-Studio. `bibliometrix` is an interface able to generate comprehensive analysis on science mapping while providing a set of tools for quantitative research in bibliometrics and scientometrics research developed by Aria & Cuccurullo (2017). Although the option for `bibliometrix`, it is worth of briefly mentioning the existence of other software's and proceedings to apply bibliometric analysis with their own pros and cons.

CitNetExplorer and VOSviewer, both developed by van Eck & Waltman are Java designed applications for analysis and visual representation of citations networks (van Eck & Waltman, 2010; 2014). Also in Java, CiteSpace is another option for visualizing trends and patterns in literature (Chen, 2006). SciMAT is an open-source software able to perform scientific mapping analysis in a longitudinal type of framework using a three-way workflow: Management of a knowledge base; science mapping analysis and visualization of results (Cobo et al., 2012).

Even on R-Studio environment there are other packages besides `bibliometrix`. Packages like CITAN (Gagolewski, 2011), ScientoText, H-index Calculator and Scholar have their own attributes and specific analysis functions; however, none addresses the entire workflow for science mapping suggested as in Aria & Cuccurullo (2017). `bibliometrix` (http://www.bibliometrix.org) is written in the R language and due the existence of a substantial amount of effective statistical algorithms inside it provides high-quality numerical routines, and integrated data visualization tools.

Aiming to portray an extensive overview on unemployment-related studies, the idea is to embrace an extensive timespan of published documents. This is particularly relevant since as we presented earlier, labour and labour studies have been constantly proposed since the early 1900's until nowadays. Therefore, we decided to initially work with a timespan from 1960 to 2021. This 61 years span will give the opportunity to identify from the earliest topics surrounding the unemployment until the most recent ones covering the COVID-19 pandemic outbreak.

Decision to start in the 1960s is to coordinate both databases (Scopus and Web of Science) that will be used to retrieve documents. Having a timespan defined the next step is to proceed for the data collection considering the selected time interval. Retrieving of raw data for bibliometric analysis following Aria & Cuccurullo (2017) is divided into three steps: Data retrieval, data loading and data converting; and we follow all these three.

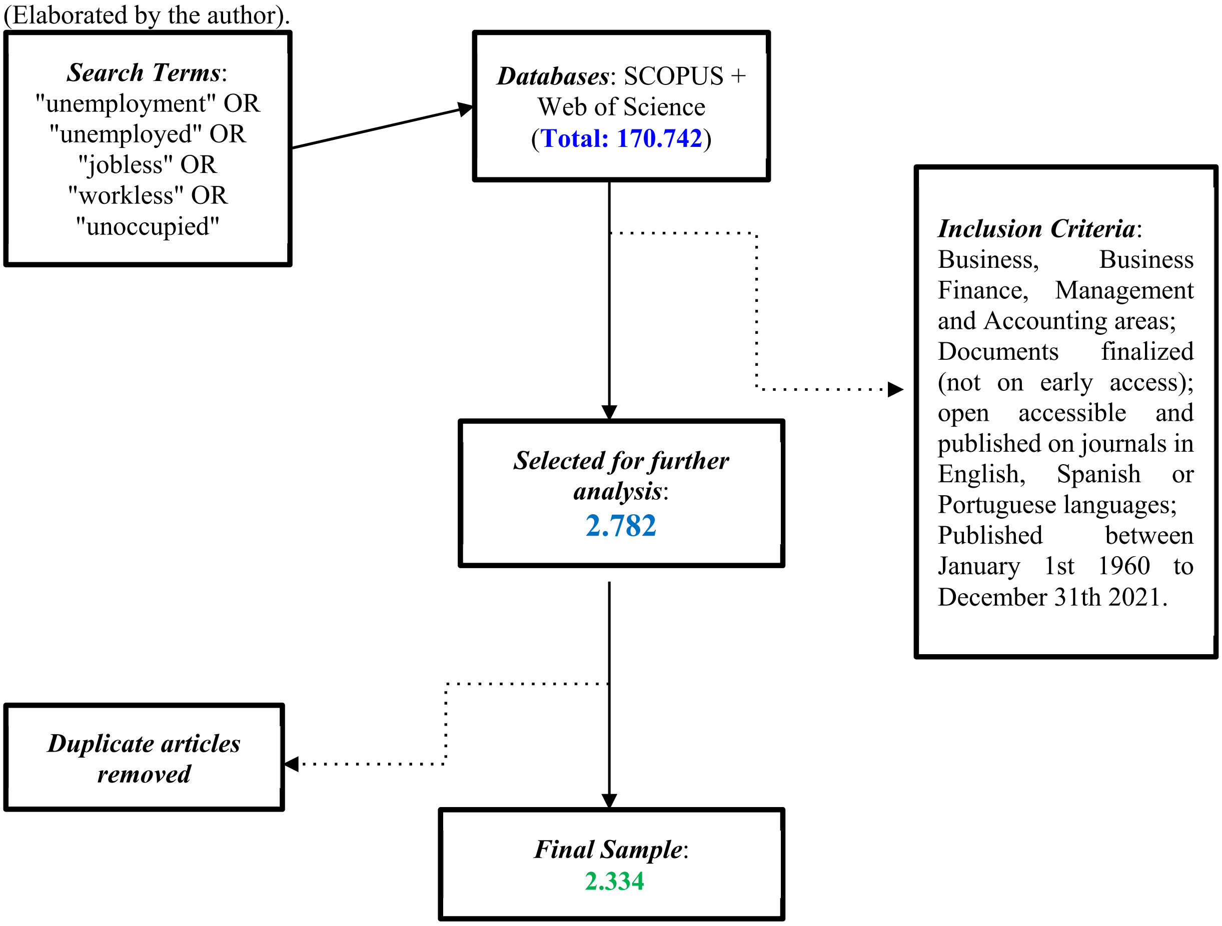
About the selection of databases, nowadays there are many online repositories of bibliographic databases with metadata about scientific studies. Web of Science database (WoS at http://www.webofknowledge.com), Scopus (http://www.scopus.com), Google Scholar (http://scholar.google.com), and ScienceDirect (http://www.sciencedirect.com/) are the most used in bibliographic studies (Cobo et al., 2011; Aria & Cuccurullo, 2017).

Evidently those three databases do not exhaust the entirety of studies nor all the research fields but, considering the R-Studio package `bibliometrix`, which will be the tool used here, extracts data from WoS and Scopus bases directly, these are the ones selected to retrieve information for the development of this chapter. After retrieving raw data, the next step is to load and convert this information to an adequate format. Again, as the software R-Studio and `bibliometrix` is the tools to be used, the conversion will be for a R-readable format.

Final stage before proceeding to the analysis is the data cleaning process. Several pre-processing must be applied to detect duplicate documents or misspelled elements on the databases. This is important due to although most databases, specially WoS and Scopus, sometimes have multiple versions of a same study, different spelling on author's names and so on. All these eventual corrections were taken into consideration and when needed will be detailed in sequential steps in this section.

Following figure 2 summarizes the research design that serve as guidance for this study. Data loading, data collection, and data conversion are all parts on the depicted in figure 2. Our adopted research design is aligned and adapting steps described in Börner, Chen & Boyack (2003) and Zupic and Cater (2015) to the mapping of science studies.

Figure 2
Research design for the bibliometric study.
(Elaborated by the author).



Note: Queries proceeded simultaneously on both Scopus and Web of Science databases on May 25th, 2022.

We worked on both Scopus and Web of Science individually as depicted in figure 2 summing initially 170.742 documents. This number of documents were reduced by our delimited inclusion criteria resulting on 2.782 documents later reduced one more time due the removing of documents found on both databases. Finally, we have a first final sample by 2.334 documents, those depending on the availability of access and our intended analysis being the unemployment-related literature. Further details about this will be presented in the results and discussion section.

Therefore, we will be using a merged sample composed by Scopus and WoS documents. There exist a few caveats considering this approach are worth of mention. Some studies defend that Web of Science have the advantage on offering a large amount of information across distinct research fields in comparison with Scopus (Garfield, 1971; Chirici, 2012; Echchakoui, 2020). Scopus has some advantages by itself, such as a high number of publications in some areas as administration and business management, for example (Mingers & Lipitakis; 2010).

Although acknowledging singularities for each database, still exists a significant level of correlation between them. Sánchez, Rama & García (2017) found a high correlation between Web of Science and Scopus repositories not only considering the availability of documents in a broader sense but particularly high on social sciences, health-related and physical studies. In line with Sánchez, Rama & García (2017) and Echchakoui (2020) we do not presume a superiority of one database over another. Premise is that both are important and to conduct a reliable bibliometric analysis, it is for the best to use both considering that one tends to complete the other.

For this chapter the retrieving process in the existing literature related to unemployment was performed on both Scopus and Web of Science repositories considering the business, business finance, management and accounting domains, using keywords related to unemployment or synonyms to the term such as jobless, workless and others presented in figure 2. Still on figure 2, same terms ("unemployment" OR "unemployed" OR "jobless" OR "workless" OR "unoccupied") were searched on both databases.

To satisfactorily respond the question and purpose delimited for this chapter, our developed science mapping, their results, and the steps that lead to our outputs will be considering a reasonable number of documents (2.334), allowing the identification of emergent themes related to unemployment and main topics discussed in this field. To not overextend in detailed explanations and avoiding being repetitive on what will be presented in the following section, further details adopted in each stage of the bibliometric analysis will be explained on the Results and Discussions section in the following.

2.4. Results and discussion.

This section presents the bibliometric analysis results considering the merged database from Scopus and Web of Science documents. Following topics are covered in this section: Descriptive information about the dataset, unemployment-related research growth, most productive authors, most relevant sources, frequently referenced papers and authors, most productive countries, and the keywords frequencies.

2.4.1. Bibliometric Analysis Results – Descriptive information about data and authorship.

Within the already described usage of `bibliometrix` R-Studio interface the function `biblioAnalysis` enables the assessment of main bibliometric measures of the merged database whereas `summary` function presents the main information about the working data. Table 2 illustrates the outputs obtained by these functions.

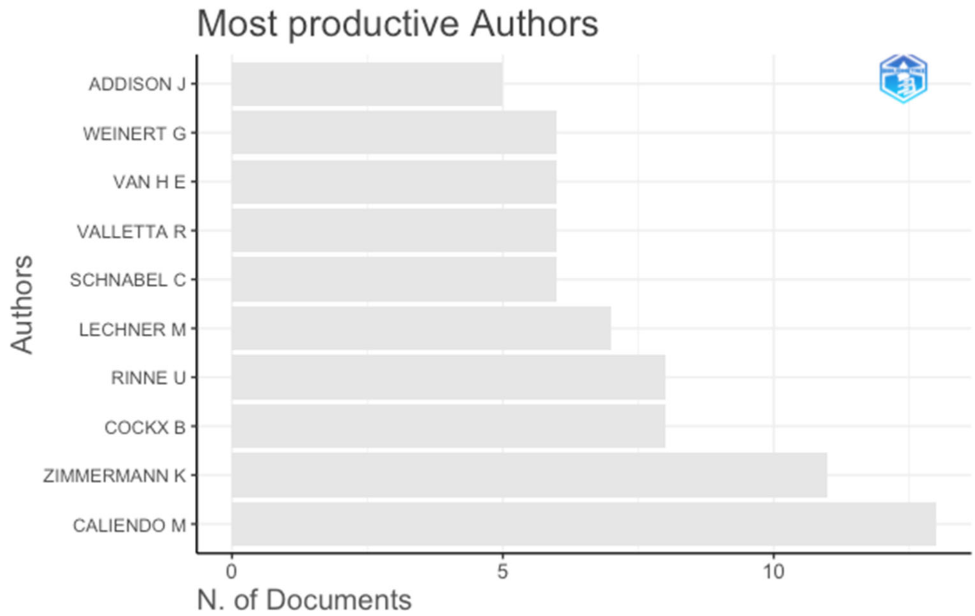
Table 2
Main information about the database.
(Elaborated by the author).

Merged database (Scopus + Web of Science)	
Documents (Total)	(2.334)
Single-authored documents	599
Multi-authored documents	1697
Sources (Journals, Books, etc.)	483
Keyword Plus (ID)	1721
Author's Keywords (DE)	5114
Timespan	1971 - 2022
Average citations per documents	15.94
Authors (Total)	(4693)
Author appearances	5468
Authors of single-authored documents	549
Authors of multi-authored documents	4144
Documents per author	0.489
Authors per document index	2.04
Co-authors per document index	2.38
Collaboration Index	2.44

Some results presented on table 2 are further refining our data beyond the initial trimming process we applied following the described-on figure 2. Although in inclusion criteria (presented in figure 2) were established as starting point on searching documents the year of 1960, the database has in its reach a first document dated from 1971; on the other end, limiting the search until the end of 2021, some documents are dated for 2022, possibly due to some journals acceptance being from 2021, but the publication occurring in the following year. Therefore, we have covered a 51-years timespan (1971-2022) and 2.234 documents to proceed.

Continuing with some initial descriptive analysis about our merged database, `bibliometrix` on R-Studio allows the presentation of some visual information that deepens some information already presented in table 2. Following figure 3 presents the 10 most productive authors within our database of documents. If needed or wanted it would be possible to have as output more than 10 authors but to not overextend the results we will be limiting most of our presentation up to 10 from hereafter.

Figure 3
Most productive authors.
(Elaborated by the author).



With 13 documents among the total by 2.334, Marco Caliendo is the author with most documents and one of them (*"Risk attitudes of nascent entrepreneurs–new evidence from an experimentally validated survey*") is also in the top-10 most cited documents. Second most productive author in the sample is Klaus F. Zimmerman, with authorship in 11 documents which represents by itself around 0.48% on the entirety of the merged database (whereas Caliendo has 0.56%). Figure 4 illustrates most productive countries.

Figure 4
Most productive countries.
(Elaborated by the author).

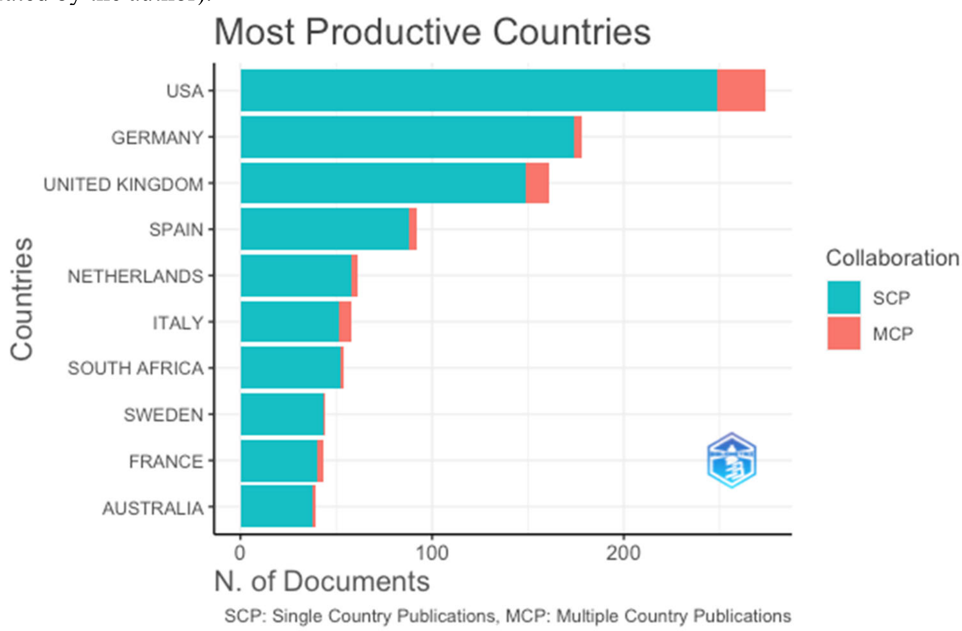
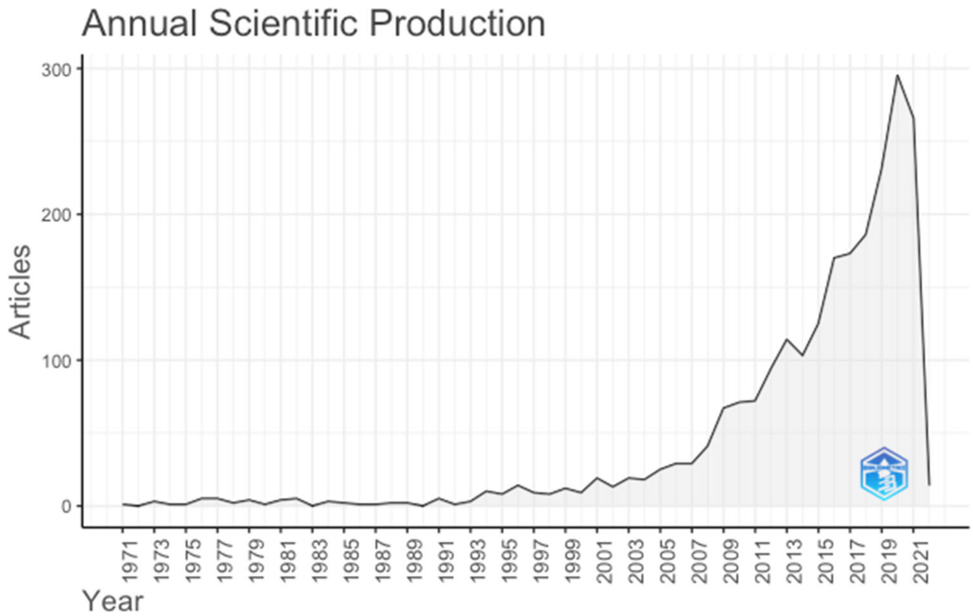


Figure 4 depict a clear representation that most of scientific production regarding unemployment-related literature are being produced in most developed economies and countries; USA, Germany and United Kingdom leads the way very comfortably. Considering absolute numbers, United States are even far away from their pairs on the top of productivity, having a total of 274 documents within the sample while Germany has 178 and United Kingdom having 161.

Numbers presented in figure 4 are considering studies with authors writing from a same country as well multiple countries' authorship (SCP and MCP collaborations as distinguished in the figure). This result presents the first alert that later guide our research efforts to move away from these advanced economies, where scientific production is more robust, to turn attention for realities less examined outside the European and USA-centred studies.

About overall production of studies during the timespan covered by our merged database of documents retrieved from Scopus and Web of Science, `bibliometrix` shows a slow start at the early period, few studies were produced from 1970s up to 1990s years. After 2000 years the number of studies starts to ascend and ascending constantly up to the peak on after 2020. Figure 5 illustrates this finding.

Figure 5
Annual production of documents.
(Elaborated by the author).



Growing on production starts a more accentuate ascending between 2008 and 2009 with 41 and 67 papers respectively. There could be a diversity of explanations for this phenomenon but probably the most influential is the global economic crisis and recession that starts around this same period and inevitably spilled effects on labour market and unemployment levels. Indeed, Choudhry, Marelli & Signorelli (2012) found that this financial crisis has a significant impact on unemployment rates, especially on youth people at the time. The correlate effect being more studies being produced about youth unemployment and the absence of jobs in a wider scope.

Refining our analyses for documents in our working sample, the following table 3 presents the 10 studies that have received most citations and the total of citations per year considering the timespan our database is covering.

Table 3
Top manuscripts per citation.
(Elaborated by the author).

Title of the Paper	Total Citations	Total Citations Per Year
Economic growth in a cross-section of cities	544	19.4
Self-employment in OECD countries	536	23.3
The sum of all FEARS investor sentiment and asset prices	488	61.0
Does self-employment reduce unemployment?	367	24.5
Real wage rigidities and the new Keynesian Model	339	21.2
Explaining female and male entrepreneurship at the country level	306	18.0
The stock market's reaction to unemployment news: Why bad news is usually good for stocks	294	16.3
Fiscal policy in a depressed economy	290	26.4
Entrepreneurship and the process of firms' entry, survival, and growth	270	16.9
Risk attitudes of nascent entrepreneurs–new evidence from an experimentally validated survey	267	19.1

Observing solely the titles of studies presented-on table 3 it is possible to perceive how unemployment is a present theme through the most distinct lenses of observation. We are briefly assessing three of those on 10 most frequently cited documents presented just for illustrative purposes on how intricated and complex the unemployment phenomenon is and how it could be tangent on different areas of research interest.

"Economic growth in a cross-section of cities" by Glaeser, Scheinkman & Shleifer (1995), most cited manuscript in the database, found that cities with tend to follow a similar economic growth pattern as the countries they are located. Meaning that would be a rare case to have a city with low unemployment levels in a country where the overall unemployment is high, and the opposite also is presumable.

Blanchflower's (2000) "Self-employment in OECD countries" study found an overall negative relationship between self-employment rates and unemployment. However, Blanchflower (2000) found no substantial evidence that increasing self-employment would as well increase the real growth rate of a given economy; similar result obtained by Glaeser, Scheinkman & Shleifer (1995) study.

"The sum of all FEARS investor sentiment and asset prices" draws attention for its higher average in total citations per year, 61.0, as presented in the table 3. Study developed by Da, Engelberg & Gao (2015) uses unemployment as a proxy in an index they called FEARS (Financial and Economic Attitudes Revealed by Search), composed by the volume of online search queries for terms such as "recession", "bankruptcy", and "unemployment", believing that these queries could be quantifiable to predict risk aversion and sentiment of investors.

More prominently by Da, Engelberg & Gao (2015) study but considering the other two mentioned as well, it is possible to observe that unemployment is in fact a multifaced phenomenon that could be analysed in a variety of manners, both its direct indirect influence. Moving forward, to unveil the most frequently cited documents we retrieved from Scopus and Web of Science databases the `bibliometrix` offers the function "`citations`", that enables the print of this information on the R-Studio environment.

Following table 4 presents the 10 most referenced documents on our working dataset. It is important to do a clarification: Although previously was presented most cited documents within the database of 2.334 documents, here the presentation is about studies that were referenced in overall literature, being the authors of these references inside the sample we are using or not.

Table 4
Top-10 documents referenced.
(Elaborated by the author).

References	Frequency of citations
Meyer, B. D. (1990). Unemployment insurance and unemployment spells. Econometrica, 58(4), 757–782. https://doi.org/10.2307/2938349	13
Shimer, R. (2005). The cyclical behavior of equilibrium unemployment and vacancies. American Economic Review, 95(1), 25-49. doi:10.1257/0002828053828572	13
Merz, M. (1995). Search in the labor market and the real business cycle. Journal of Monetary Economics, 36(2), 269-300. https://doi.org/10.1016/0304-3932(95)01216-8	11
Wanberg, C. R. (2012). The individual experience of unemployment. Annual Review of Psychology, 63, 369-336. https://doi.org/10.1146/annurev-psych-120710-100500	11
Paul, K. I.; Moser, K. (2009). Unemployment impairs mental health: Meta-analyses. Journal of Vocational Behavior, 74(3), 264-282. https://doi.org/10.1016/j.jvb.2009.01.001	11
Rosenbaum, P. R.; Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41–55, https://doi.org/10.1093/biomet/70.1.41	10
Nickell, S. (1997). Unemployment and labor market rigidities: Europe versus North America. Journal of Economic Perspectives, 11(3), 55-74. DOI: 10.1257/jep.11.3.55	10
Caliendo, M. Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. Journal of Economic Surveys, 22(1), 31-72. https://doi.org/10.1111/j.1467-6419.2007.00527.x	9
Spence, M. (1973). Job market signaling. The Quarterly Journal of Economics, 87(3), 355–374. https://doi.org/10.2307/1882010	8
Winkelmann, L., & Winkelmann, R. (1998). Why are the unemployed so unhappy? Evidence from panel data. Economica, 65(257), 1–15. http://www.jstor.org/stable/2555127	8

It appears that at least three topics are most recurrent considering the documents presented on table 4: Cyclicality of jobs opportunities (fluctuations of employment and unemployment); labour market in a broader perspective, and the potential effects of unemployment on health-statuses. Knowing the most frequently cited documents is useful to acknowledge who as well would be the referential authors cited on our working sample of documents. Following table 5 presents 10 first authors that are recurrently cited within the collection of 2.334 documents.

Table 5
Top-10 most frequented cited authors.
(Elaborated by the author).

Authors	Frequency of citations
BLANCHARD, Olivier	329
CARD, David	291
HECKMAN, James J	267
CLARK, Andrew	228
HALL, Robert E	207
CALIENDO, Marco	202
PISSARIDES, Christopher	196
MORTENSEN, Dale T	187
BLANCHFLOWER, David G	185
LEE, Hoyoung	161

Table 5 illustrates that by some distance Olivier Blanchard is the researcher more frequently cited within the merged database. Professor Blanchard, who is an economist, have indeed contribute largely on macroeconomics topics including unemployment. According to Google Scholar at least two unemployment related study are in his top-10 most cited document in the Google repository. In 1986, Blanchard & Summers discussed effects of macroeconomic shocks on unemployment rates, that were steadily growing at the time (Blanchard & Summers, 1986).

Later, in year 2000, Blanchard now with co-authorship by Justin Wolfers takes a step further on his previous study, arguing that the continuum on European unemployment rising since the 1960 years, besides of economic shocks, could be influenced by country-related heterogeneity factors (Blanchard & Wolfers, 2000). Countries heterogeneity as we mentioned on this thesis introduction and would be cleared on sequential chapters is one of our guiding themes on the countries' data that will be later analysed.

Another assessment about authorship beyond absolute numbers depicted on table 5 is the h-index factor, a metric that considers both productivity and citation impact. H-index or Hirsch index is an index h, defined as the number of papers with citation number $\geq h$. It is a useful indicator to characterize the scientific output of an academic contributor (Hirsch, 2005; Schreiber, 2008).

The package `bibliometrix` offers the function `Hindex`, which calculates an author H-index, sources H-index, as well its variants (g-index and m-index). It is possible to see authors individually searching by their referred name as well in groups. Table 6 presents the results regarding h-index of the first 10 most productive authors in our working collection of documents. We are focusing on h-index but table 6 presents as well other useful indicators for deeper analyses, such as g-index, m-index, total citations, number of produced documents and the year the author being observed started to write about the unemployment theme.

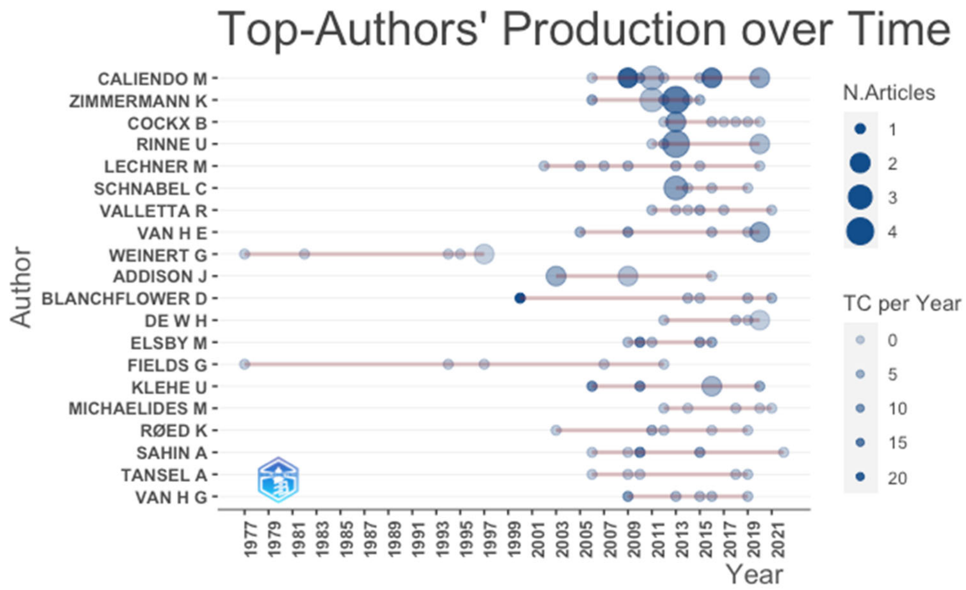
Table 6
Top-10 most productive authors by h-index.
(Elaborated by the author).

Author	h-index	g-index	m-index	Citations	Number of documents	Starting year
CALIENDO, M.	11	13	0.64	637	13	2006
ZIMMERMAN, K.	10	11	0.59	435	11	2006
RINNE, U.	7	8	0.58	184	8	2011
LECHNER, M.	6	7	0.28	287	7	2002
ADDISON, J.	5	5	0.25	173	5	2003
BLANCHFLOWER, D.	5	5	0.21	621	5	2000
COCKX, B.	5	7	0.45	148	7	2012
ELSBY, M.	5	5	0.36	375	5	2009
FIELDS, G.	5	5	0.11	41	5	1977
KLEHE, U.	5	5	0.29	443	5	2006

Table 6 reiterates some of the results before presented. Marco Caliendo for example, appears a strong contributor on the unemployment-related literature considering our working dataset. Regarding the h-index specifically, considering the total of 13 documents that Caliendo has authorship, 11 of these have been at least 11 times referenced. Caliendo’s g-index in 13 is also a significant number, since this should be equal or higher than h-index whereas the former suggests more relevance to highly cited documents than h-index (Egghe, 2006).

Further analysis about author’s productivity over the timespan we are covering could be presented using `bibliometrix` function `authorProdOverTime`, a visual illustration that considers the calculation of an authors’ production in terms of number of documents in the sample and total citations per year over the span analysed. Figure 6 depicts these results.

Figure 6
Authors’ production over time.
(Elaborated by the author).



Bigger and darker circles represent the more productive authors' considering the figure 6. Again, Caliendo stands out from others in our sample, being consistent both in number of articles produced and citations received. Furthermore, Zimmerman has two relevant documents on the sample around 2012 and 2013, being co-author in studies that observes unemployment and labour market in the European context (Rinne & Zimmermann, 2012; Giulietti et al., 2013). Blanchflower has a highly cited document in the beginning of the 2000s years, most recurrently visited than many articles written later.

Blanchflower (2000) study discusses a theme that have been replicated and seems to be a persistent related phenomenon to unemployment: Entrepreneurship. Proposing a manner to measure self-employment rates, the author intended to verify if this form of alternative employment would be sustainable and if affects or have any relationship with the final levels of unemployed people. From Blanchflower (2000) and other studies that appeared so far, it seems a trend that self-employment indeed present some relationship with unemployment theme.

2.4.2. Bibliometric Analysis Results – Bibliographic network matrices and associations.

Aria & Cuccurullo (2017) in their designed study that complements the usage of the `bibliometrix` package suggest that attributes in any bibliographic database are inherently connected through several links: Authors to journals, keywords and publications, publications on dates and many others. These linkages may not be easily perceived but could be assess through the creation of networks, that consequently can be presented in the form of matrices (Aria & Cuccurullo, 2017).

Still according to the authors, co-citation and coupling are probably the most acknowledged networks (Aria & Cuccurullo, 2017) and are the ones we are deepen our analysis from hereafter. `bibliometrix` has an inside function called `cocMatrix` that allows the creation of bipartite networks using two bibliographic attributes to assess the interconnectedness between them. Table 7 presents the network between documents and source of publication.

Table 7
Network documents and sources.
(Elaborated by the author).

Sources	***Number of documents***
LABOUR ECONOMICS	181
INTERECONOMICS	73
JOURNAL OF MONETARY ECONOMICS	68
IZA JOURNAL OF LABOR POLICY	56
INTERNATIONAL JOURNAL OF MANPOWER	54
JOURNAL OF LABOR ECONOMICS	40
JOURNAL OF MONEY CREDIT AND BANKING	39
IZA JOURNAL OF EUROPEAN LABOR STUDIES	37
JOURNAL FOR LABOR MARKET RESEARCH	35
SOUTH AFRICAN JOURNAL OF ECONOMIC AND MANAGEMENT SCIENCES / WORK, EMPLOYMENT AND SOCIETY	30 (in each)

Table 7 illustrates that by some distance most of the documents in the merged database are published in one source, the Labour Economics journal. Even adding the second and third most relevant sources presented on table 7 does not equal the number of documents on Labour Economics. Since first volume, dated from 1993, and according to their website (https://www.sciencedirect.com/journal/labour-economics/about/aims-and-scope), this journal has been calling for international researchers that had developed solid empirically and tested studies with strong economic interpretation about topics of interest from labour economists.

Besides Labour Economics, it seems clear that unemployment studies have been largely published on journal having an economic scope. At least considering the table 7 results, it is rarely present journals scoping to analyse unemployment and related themes with business and management lens, a gap that this thesis intends to respond when we have our study finalised.

We move other networks presentation similar as the one presented in table 7. Again, using `bibliometrix`, we present on the following table 8 the association between countries production and number of documents originated by each country. A complement for what we early illustrate on the figure 4.

Table 8
Network countries and documents.
(Elaborated by the author).

Country	Number of documents
UNITED STATES OF AMERICA	400
GERMANY	267
UNITED KINGDOM	240
SPAIN	117
SOUTH AFRICA	93
NETHERLANDS	78
ITALY	71
FRANCE	67
AUSTRALIA	60
POLAND	52

With some margin, table 8 shows presents that most studies were developed by authors' that have as their first affiliation United States. The 400 USA affiliated documents represent approximately 17.4% of our entire working merged database. Germany and United Kingdom, two next countries have a 22.5% summed. Considering the three nations, almost 40% of documents we are assessing is referring to advance and solid economies, corroborating what we also present in figure 4 and our insight to move away our analyses from these countries.

Continuing to observe bibliographic connections, it is possible to have an initial insight about the keywords frequently used on the documents within our database. Results presented on table 9 will show frequency of keywords used, which is essentially a first indicative about potential driving themes contiguous to unemployment. Further analysis and more elaborated networks will be presented later but this first insight already suggest some emergent themes.

Table 9
Network authors keywords x frequencies of use.
(Elaborated by the author).

Keywords	Frequency of usage
UNEMPLOYMENT	463
EMPLOYMENT	83
ENTREPRENEURSHIP	74
COVID-19	54
LABOUR MARKET	53
YOUTH UNEMPLOYMENT	52
UNEMPLOYMENT INSURANCE	49
INFLATION	47
ECONOMIC GROWTH	45
EDUCATION	38

It is no surprise that the most recurrent term would be "unemployment" given not only this were one of the key terms initially searched both on Scopus and Web of Science repositories but is the overall theme for this chapter and for the thesis at large. Counterpart term, "employment" could also be inferred. Table 9 presents more than the employment and unemployment terms, presence of "entrepreneurship" for example confirms some results previously obtained within the database as the study of Blanchflower (2000), for example. Blanchflower contributes on a study that discusses other keyword present on table 9, discussing the effect of recession period started on 2008 has on young people being unemployed (Bell & Blanchflower, 2011).

These types of networks may also occur among documents within a same database. Kessler (1963) defines that two studies may be bibliographically connected if at least one of their cited sources appears in the references (or similar) section on both articles (Kessler, 1963). Association, matrices and coupling of items in a bibliographic dataset may occur in different level from words and usage as we just presented to documents and authors and so on. The usage of `bibliometrix` enables the assessment of every potential association. We do not present all of these to not overextend beyond the purpose of this research although we recommend and make or own database available, at request, for further checking's.

2.4.3. Bibliometric Analysis Results – Bibliographic collaborations and networks.

Collaboration networks may be understood as an illustrative presentation where nodes in the network represent the authors and the links between other nodes are co-authorships (Aria & Cuccurullo, 2017). Analysis of these networks are relevant because those could be helpful to better understand the structure of scientific collaborations. It could be argued that the comprehension of collaborations between authors within a theme, in our case the unemployment, could be more assertive than the number of citations an author has by its own.

Using function `biblioNetwork` from `bibliometrix` we were able to observe the form of collaboration mentioned above. Figure 7 presents the authors' collaboration network considering our working merged database composed by Scopus and Web of Science documents.

Figure 7
Authors' collaboration network.
(Elaborated by the author).

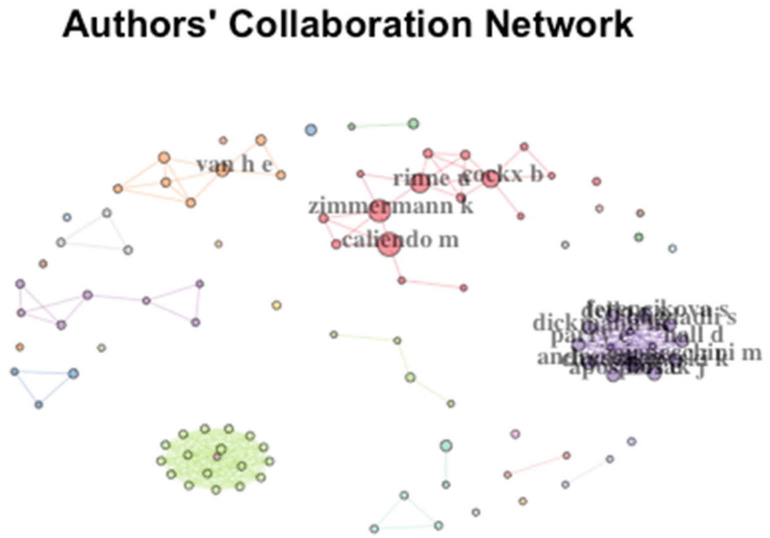


Figure 7 illustrates three larger groups of collaboration networks. Red groups seem to be the most relevant following by the purple network that has a higher concentration of authors although more disperse. The red network has the presence of authors recurrently appearing in our dataset, names of Caliendo, Zimmermann, Rinne and Cockx were also present on most cited authors on the earlier presented tables 5 and 6. Similarly as presented in figure 7, other bibliographic networks can be plotted into graphic visualization and from hereafter this will be our focus regarding results presentation.

So far, we have been analysing the merged database, the one with 2.334 documents from both Scopus and Web of Science repositories. Continuing with further analysis and considering some specificities, the idea is to run separate plots to compare documents from the two used sources of documents. These separate plots, when presented, will be useful to identify if exists any major differences between Scopus and WoS.

When doing these separated presentations we will be using the interface `biblioshiny`, also available through `bibliometrix`. To start these simultaneous assessments, following figure 8 illustrates a network for countries' scientific collaboration considering documents extracted from Scopus; after this, figure 9 presents the same type of network but considering the Web of Science documents.

Figure 8
Countries' collaboration network on Scopus database.
(Elaborated by the author).

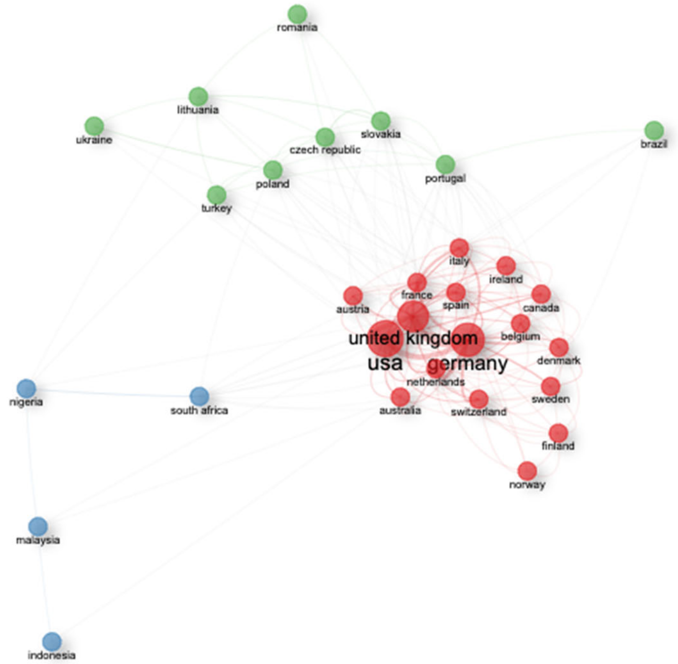
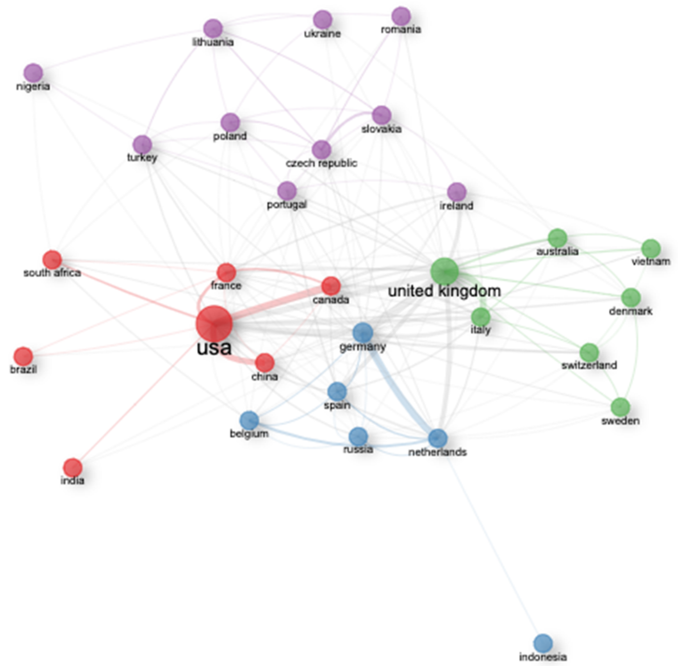


Figure 9
Countries' collaboration network on Web of Science database.
(Elaborated by the author).



We may notice on both figures 8 and 9 that United States seem to have a major role leading networks in Scopus and Web of Science documents as well. USA being the nation exerting scientific dominance about unemployment-related studies as presented in figures 8 and 9 reinforces results previously presented here at the figure 4 and table 8. Furthermore, results presented in the last presented figures reflects some distinction between the two repositories of documents we are dealing here.

Considering the presented-on figure 8, United States scientific production about unemployment seems to be more diversified with other countries, although most of them are in the European continent. In figure 9, the strongest USA collaborations appears to be with China, having as well a relevant connection with India and Canada. These distinctions about a same topic under analysis is convergent with other bibliographic studies (Chirici, 2012; Echchakoui, 2020) that argues that across some disciplines or themes Web of Science could offer a wider range of information.

Going back to authorship analyses, figure 10 shows a co-citation network. Here we use again the merged database as biblioshiny enables this specific format of observations using two repositories simultaneously.

Figure 10
Co-citation collaboration network.
(Elaborated by the author).

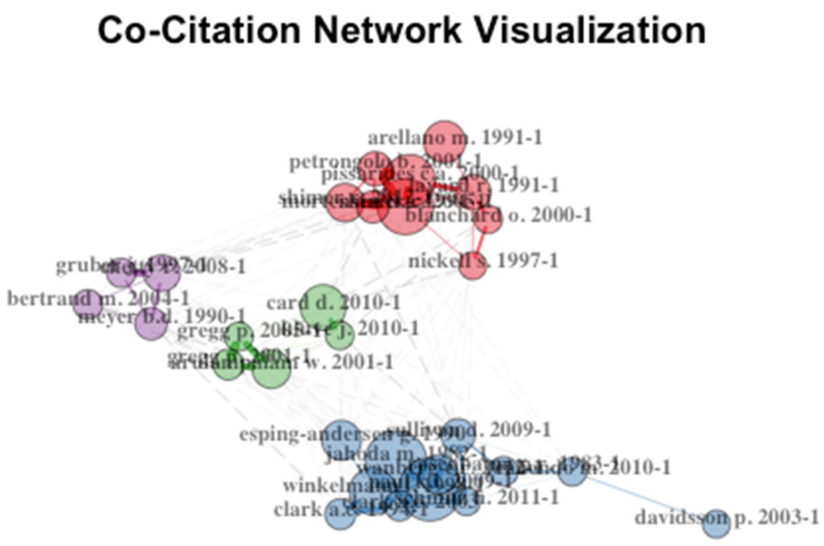
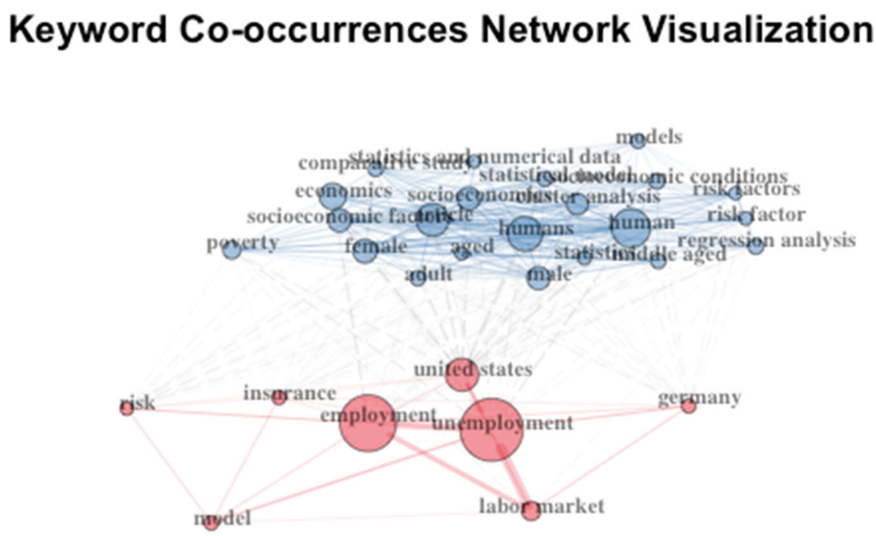


Figure 10 presents four major groups of co-citation collaboration. It seems that the co-citation effort within the groups is intense, especially in the red cluster, where are located some of most recurrent cited authors' that we have been discussing up to this point, such as Caliendo, Blanchard, and Pissarides. However, collaboration outside the groups presented does not seem particularly apparent. This could be a suggestion for researcher's efforts in more collaborative studies outside their own "bubbles".

Now as complement to what was presented in table 9, following figure 11 illustrates the keyword co-occurrence frequency. Once again, the proceeding was to run a plot visualization considering our merged database, since the fields used in both Scopus and Web of Science regarding the keyword usage are the same, enabling e to analyse their documents at the same moment.

Figure 11
Keyword co-occurrence network
(Elaborated by the author).



As we already noticed by the table 9 presented outputs, figure 11 indicates a strong usage and recurrency of use in a triad of terms: Employment, unemployment, and labour market. This is no surprise given these are probably the most frequently focus on most of the studies that are composing our working merged database. The keyword "United States" appears, and it seems to be the term linking the two groups illustrated in the figure. This could suggest that most of terms or even a major part of studies uses USA as their context or as a reference for their own research, once again, reinforcing our intention to further analyse a different context and countries' reality.

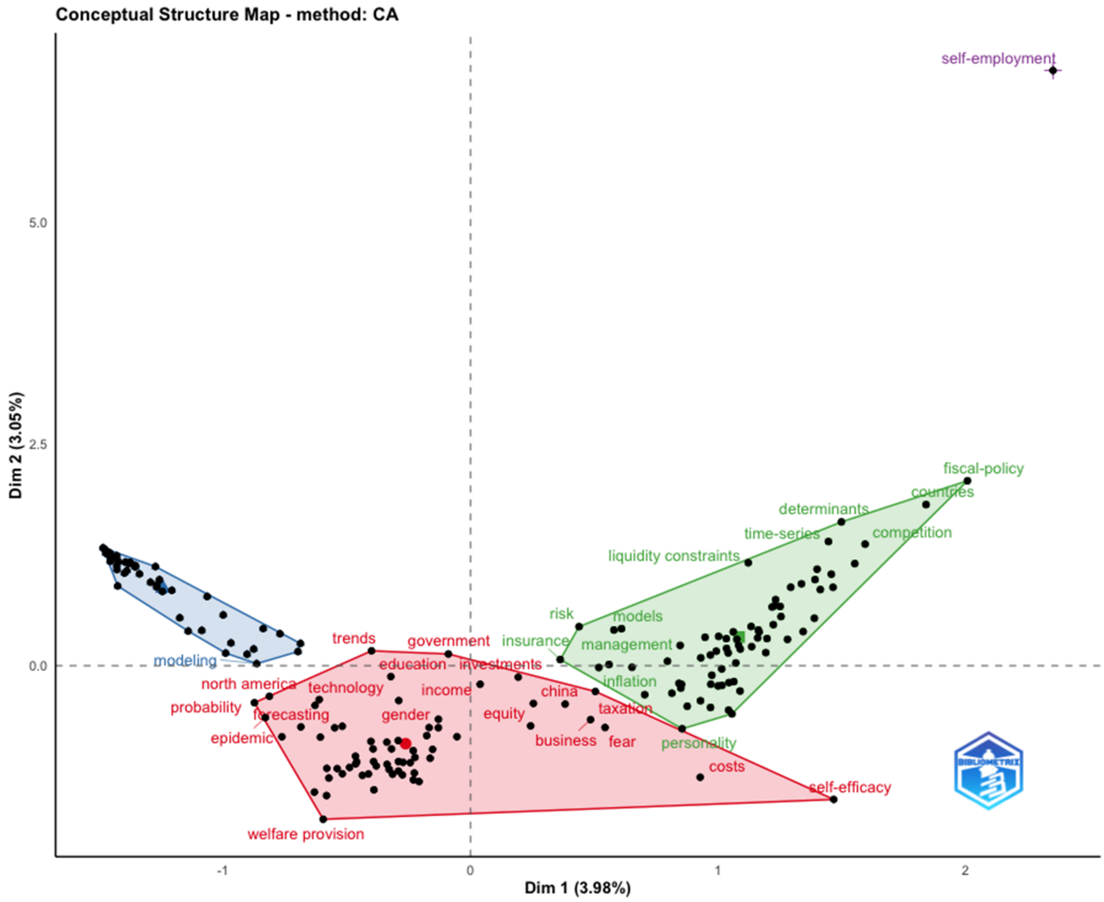
Overall, figures 7 up to 11 presented in this topic confirms visually some of the results previously indicated in previously presented results. Keyword co-occurrence analysis presented in figure 11 is particularly insightful, as shed some light and suggest some relationships between terms and what could be emergent themes of discussion most frequently used on unemployment-related studies. Next topic will be further analysing the usage of terms and words on the working sample of documents, deepening the assessment about emergent themes on unemployment research field.

2.4.4. Bibliometric Analysis Results – Co-word analysis, thematic evolution, and emerging themes.

Co-word analysis enables an identification of the conceptual structure in a collection of documents using word co-occurrences number within a sample of documents (Aria & Cuccurullo, 2017). This type of assessment is particularly useful to achieve the purpose defined for this chapter: To have an identification of main topics and themes related to unemployment currently discussed (emergent) and the ones that endures on the field. By co-word analysis, it is expected a better comprehension about discussions in the literature related to unemployment, both these enduring and the recently emerging.

The function `conceptualStructure`, available via `bibliometrix`, can perform a Correspondence Analysis (CA) or Multiple Correspondence Analysis (MCA) to draw the conceptual structure in a bibliographic database. Following figure 12 presents the default conceptual structure map for our working merged sample of documents.

Figure 12
Conceptual structure map via Component Analysis.
(Elaborated by the author).



Four groups (blue, red, green, and purple) were formed and presented-on figure 12 using the Component Analysis default proceeding. There is a group separating itself from the others: “Self-employment”, in the purple colour. This is in line with previously obtained results, it really seems that when unemployment is analysed, in some form the self-employment appears. That results with early mentioned studies such as the the contributions by Blanchflower (2000; Bell & Blanchflower, 2011). As for the blue group, the leading term is “modelling” which could be an indicative that maybe a significant part of studies within the sample had the intention to develop or test some models discussing labour market or unemployment.

Groups in green and red appears as the ones with most quantity of words inside them. Red group have words more related to individual’s characteristics, such as “Income”, “gender”, “education”, “welfare provision”, “government” and even “epidemic” appears correlating between themselves which could lead this to be understood as a socioeconomic group of terms. Green group to finalise the four that were formed, have more econometric words, such as “Fiscal policy”, “liquidity constraints”, “risk”, “insurance” and “inflation”.

Figure 12 analysis, illustrated by the four groups formed, presents how different terms may ensemble between themselves and how they converge when considering the larger unemployment scope. From the co-word analysis we may infer three potential driving themes emerging: Self-employment, economic and political indicators, and socioeconomic factors, all relatable direct or indirectly with unemployment.

To complement this initial understanding, we move to the usage again of the `biblioshiny` interface, aiming to expand and refine results about driving themes. Going forward we will one more time run separately the samples of documents extracted from Scopus and Web of Science databases. By performing the analyses considering each repository individually we believe to have a larger comprehension of recurrent terms, words used, and keywords, enabling a more assertive view on how the unemployment-related literature field have been evolving thematically.

In fact, if even considering distinct databases, Scopus and Web of Science in our study’s case, themes included on both appears to have similar results, more relevant these themes appearing will be. To assess this premise, it will be presented on the following figures 13 and 14 two illustrations created by `biblioshiny`, first figure will consider documents extracted from Scopus whereas the second one is based on the Web of Science sample of studies.

Important to remember some characteristics about our retrieved sample of documents and that guides the following figures presentation. Timespan covered and studies available for each database differs; Scopus have the first document extracted dated from 1971 whereas Web of Science has as starting point the 1976 year. Cutting point at the last document retrieve is also distinct, being 2021 for Scopus and 2022 for WoS documents. Further evaluations and analysis will be following the presentation of figures 13 and 14.

Figure 13
Thematic evolution on Scopus database.
(Elaborated by the author).

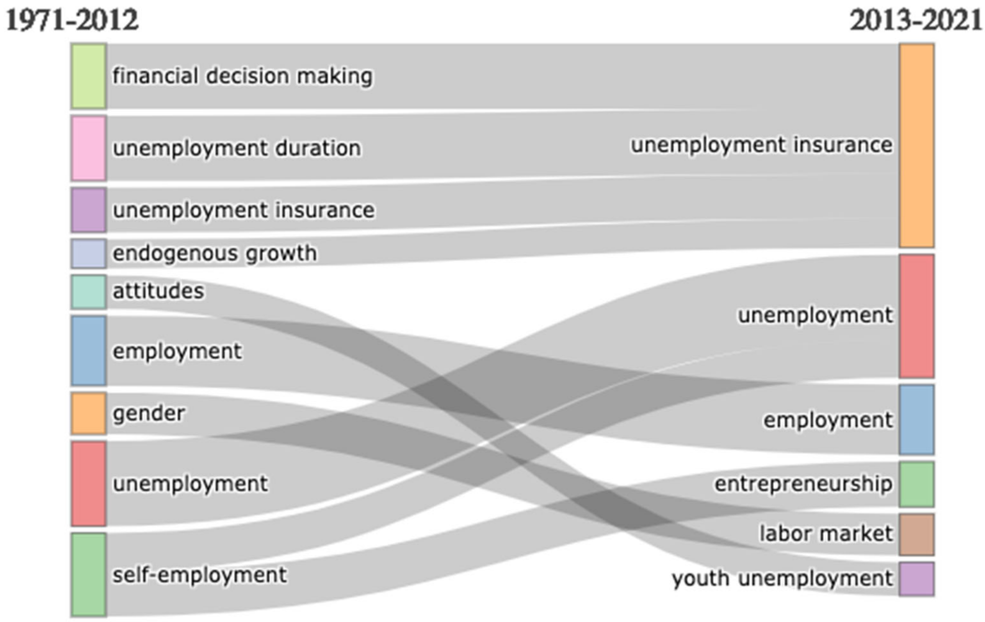
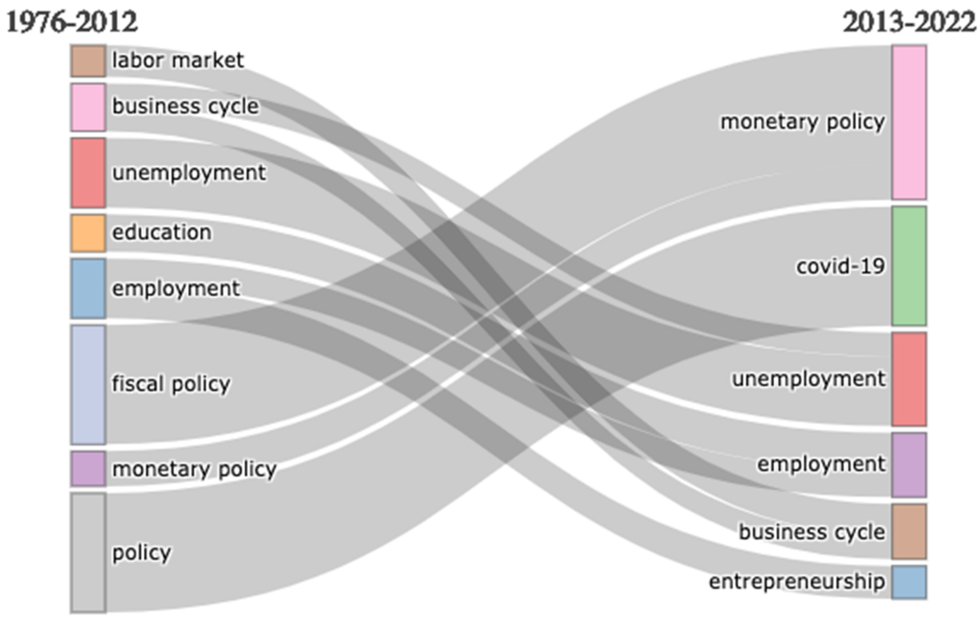


Figure 14
Thematic evolution on Web of Science database.
(Elaborated by the author).



Both figures 13 and 14 have on their left-side the earliest driving themes on unemployment-related studies while in the right-side are illustrated how these initial themes evolve or remained relevant topics. Overall, the terms presented on the figures indicates convergent topics, differences and reinforce some of the insights formerly founded and mentioned during other topics on the results section. We further analyse some of these topics from hereafter.

Although in figure 13 “unemployment insurance” being present in both sides of the image, it is most recently that topics such as “financial decision making” and “unemployment duration” were absorbed by the unemployment insurance. Literature suggests that these topics seems indeed to be connected. Baily (1978) related insurance and income arguing that the possibility of becoming unemployed directly implies an uncertainty about financial revenue for workers and the dependency those would have on insurance programs. Acemoglu & Shimer (1999) found empirically that insured workers when looking for a job tends to search for the high-wage jobs, even incurring a higher unemployment risk.

Unemployment insurance has also some relationship with health-related topics O’Campo et al. (2015) argues that psychological distress among unemployed and even the employed people, could be impacted by generosity of eventual unemployment insurance they receive, the duration of it, and wage replacement if not receiving any insurance. Taking into consideration these examples, it seems justifiable to presume unemployment insurance not only as an emergent theme but also having the potential to be a determinant that could explain unemployment rates.

Still on Figure 13, another theme presented on the right-side of emerging topics, and a matter of interest already mentioned here and constantly included on the unemployment literature is “youth unemployment”. This theme has appeared early in table 9 as one of the highest used keywords on the merged database and as a thematic focus on some of the most relevant authors present in our sample of documents (e.g., Bell & Blanchflower, 2011; Caliendo & Schimdl, 2016).

Breen (2005) discuss cross-national differences in youth unemployment even when assuming a restricted group of nations, in his case on the OECD. Author’s results show that youth unemployment tends to be high in regulated labour markets and low in liberal markets. Youth unemployment has also been solidly analysed in moments of economic crisis. Choudhry, Marelli & Signorelli (2012) study for example, finds that in a crisis context, rates of youth without a job are greater than in other type of labour force.

Marques & Hoerisch (2020) discussed massive youth unemployment on Europe during the 2008 crisis period, suggesting that higher numbers of youth unemployed are common in economic crisis. Many fronts in the literature seems to indicate that any intention to better understand unemployment phenomenon must consider the youth portion of the labour market, justifying this not only as a relevant theme on the field but as well one of the potential factors that composes overall unemployment.

Moving to figure 14, it is possible to perceive a significant presence of the term “monetary policy”, appearing as the leading term in the illustration absorbing terms previously used such as “policy” and “fiscal policy”. Lockwood, Miller & Zhang (1998) argued that the design of a monetary policies is directly dependent of the type of unemployment to be mitigated, potentially leading to a twofold phenomenon, more insurance programs and higher inflation indexes to foment these insurances.

A frequently appearing author, Olivier Blanchard, also discuss monetary policy and labour market. Blanchard & Gali (2010) proposed a utility-based model, considering fluctuations between rigidity on monetary policies and unemployment levels. Monetary policies are regularly viewed as a managerial response to deal with economic crisis or unexpected shocks that not rarely tends to increase unemployment. Stockhammer & Sturn (2012) for example, suggests that the extension on hysteresis in the aftermath of recessions depends on monetary policy reactions and the easiness on it during the recession period. Monetary policies may also be useful to responds to external shocks and dampens their impact, especially on emergent economies (Horvath & Zhong, 2019).

Important to notice that as a term, “monetary policy” is wide and could have different meanings. Therefore, to use this as potential determinant for unemployment rates a general indicator must be used as a proxy for these policies. Most frequently studies tend to use inflation as a proxy to represent a measure or eventually an outcome of monetary policies. The before referred Lockwood, Miller & Zhang (1998) study for example uses inflation as indicator, something we are aligned with and will be further detail on sequential chapters of this thesis.

Going back for figure 14, “COVID-19” appears in as the second high on emergent relevance considering the Web of Science portion of documents. COVID-19 derived from “policy” and “fiscal policy” terms, probably representing the efforts made through these practices to cope with the pandemic scenario. When we are writing our study COVID-19 was an ongoing crisis and most of its outcomes are not yet perceived or measurable with preciseness. Hence, COVID-19 studies remain fuzzy since the world is still coping with the pandemic and most of studies are local to specifics contexts such as Germany (Bauer & Weber, 2020), Portugal (Almeida & Santos, 2020), United Kingdom and India (Victor et al., 2021), and other European countries (Su et al., 2021).

Nonetheless, COVID-19 indeed seems to already have a relevant effect on labour market and its influence probably will endure in short and medium terms of time. The availability of measurable data to use COVID-19 as a direct determinant on the unemployment rates composition remains to be seen in the next chapter, but the influential power on the phenomenon is already acknowledged as presented-on figure 14 and by some studies that we found on our retrieving process of documents, particular in the Web of Science.

At last, but not less important, both figures 13 and 14 brings to the table other topic already discussed in other topics of this chapter: Entrepreneurship. Entrepreneurial studies are in both right sides on the figures, meaning that is emergent topic of discussion coming from employment and self-employment terms. Baptista & Preto (2007), considering Portuguese context, found a strong relationship; for their reality, elevated unemployment rates lead to self-employment whereas in opposition higher self-employment would potentially decrease unemployment in long-term.

This consequential effect pointed out about Portugal labour market by Baptista & Preto (2007) was also observed in other countries such as US, UK, Spain, and Ireland (Faria, Cuestas & Gil-Alana., 2009). Some studies have been discussing public policies applied aiming to foment a more active labour market programs, promoting individuals’ interest on start their own business to, consequently, reduce unemployment levels (Laffineur et al., 2017; Michaelides & Davis, 2020).

The literature we reviewed also suggest some relationship between entrepreneurship and unemployment insurance. Xu (2022), for example, shows that higher insurance usually deters unemployed individuals from forming their own businesses. Youth unemployment and entrepreneurship are also related, Awogbenle & Iwuamadi (2010) examining the African context, advocates in favour of vocational and entrepreneurial programs as a short-term intervention to tackle unemployment, specially at early stages of youth lives.

Studies with a similar argument has been made for different countries, such as Lithuania and Poland (Greblikaitė, Sroka & Grants, 2015), Indonesia (Ridha, Burhanuddin & Wahyu, 2017), Philippines (Camba, 2020), Kazakhstan (Zhartay, Khussainova & Yessengeldin, 2020), South Africa (du Toit, 2021) and others. Connections on entrepreneurship and self-employment with the terms mentioned here goes on and on and would not be extinguished here neither this is the aim for now. Nonetheless, it seems reasonable and well literature-based that some form of entrepreneurship must be considered as a potential determinant that would explain unemployment rates composition.

To summarize the presented in figures 13, 14 and what we have been arguing up to this point, it is our understanding that five major themes have emerged from our bibliometric analysis: Unemployment insurance, youth unemployment, monetary policies, COVID-19, and entrepreneurship. Presuming that our target variable to be assessed is the unemployment, our proposition on sequential chapter of this thesis is to ascertain if these five terms could be useful to explain how unemployment rates is composed.

Specificities about the terms and themes we identified here and how they could or could not be measurable as variables will be considered and further explained on the following chapter. Nonetheless, we believe to have a solid starting point about factors that relates and, therefore, could be potential explicative factors on how unemployment rates are composed.

Closing out the presentation of results from our bibliometric analysis we return to this chapter's research question (What are the main topics discussed in the unemployment-related academic literature?) and defined objective (To have an identification of main topics and themes related to unemployment currently discussed (emergent) and the ones that endures on the field). Considering the depicted on this section, we believe to have fulfilled the purpose as well responded the research question satisfactorily.

Following figures 15 and 16 presents some of the "trend topics" unveiled considering the timespan of years covered by documents retrieved from Scopus and Web of Science databases and analysed during this chapter. Again, these illustrations were created via `biblioshiny` interface and `bibliometrix` R-Studio package developed by Aria & Cuccurullo (2017). Figure 15 presents the trending topics from our Scopus documents while figure 16 is referring to studies extracted from the Web of Science.

Figure 15
Trending topics on Scopus database.
(Elaborated by the author).

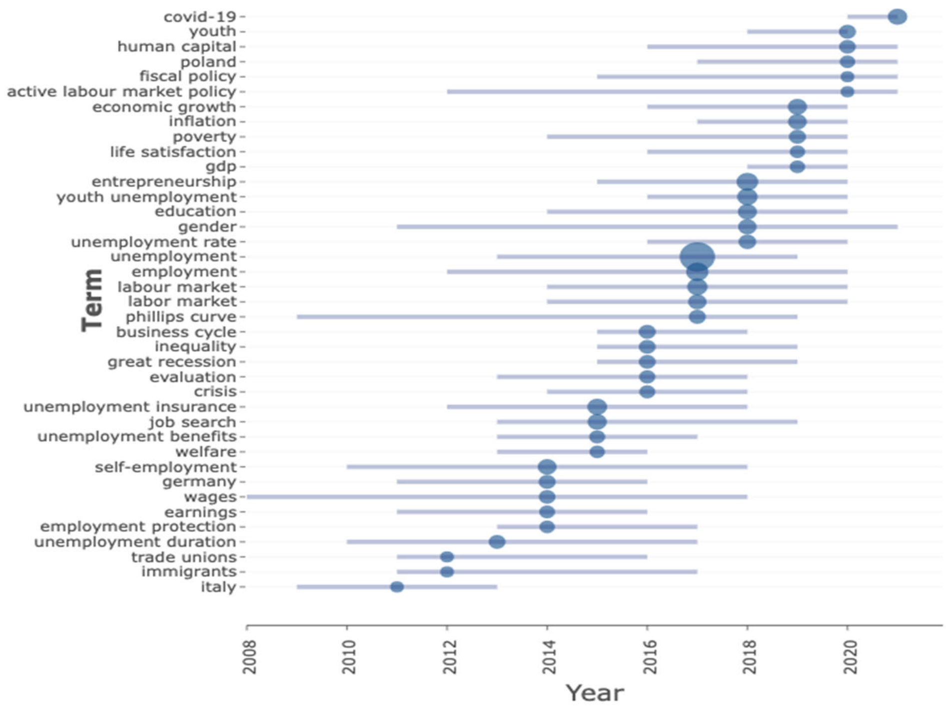
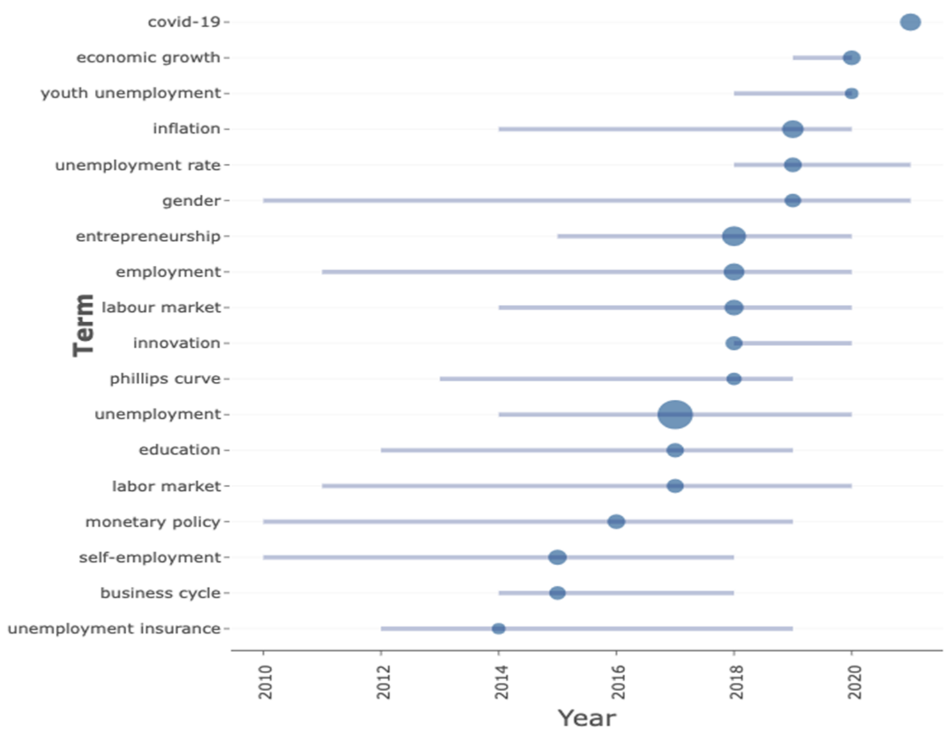


Figure 16
Trending topics on Web of Science database
(Elaborated by the author).



Figures 15 and 16 presented other terms beyond those we already discussed and that could as well be presumed as potential determinants for explaining the unemployment rates. Terms presented on both figures could be useful to extract some insight of potential first order variables or proxies to be used representing other factors. Any additions or eventual changes of the emergent terms we identified here, if necessary, will essentially depend on the availability of reliable data to be assessed.

Nonetheless, we believe to have achieved a reasonable starting point of factors emerged from our bibliometric analysis. Our revision of past literature on unemployment-related studies offered a well-established group of driving and emerging themes depicted on figures 13 up to 16. From hereafter we may proceed for more econometric analysis on chapter 3 that will be developed on the sequence. Before that, next section concludes the ongoing study on our multiple studies approach for this thesis.

2.5. Conclusions.

All things considered and presented during the bibliometric analysis allowed an overview about emerging topics sub-topics and research tracks in the unemployment-related literature considering this theme analysed on the business, business finance, management, and accounting areas. Our presented revision on this topic may be presumed as a summarization that presents a solid basis about topic that could be determinants to unemployment rates and be helpful to advance research and collaborate with other studies in the field.

We unveil that labour market, employment, and unemployment studies are enduring on academic research over past decades whereas considering recent socioeconomic scenarios, such as the post-pandemic context, the relevance of these topics has the potential to continue relevant and growing on importance. Developed bibliometric analysis enables to see in practice that indeed unemployment, even observed in isolation, is a multifaceted and complex phenomenon. Considering the sample of documents we have worked here we have early documents dated from 1971 up to 2022, covering more than 50 years of studies discussing unemployment by distinct lenses of analysis.

Diversity on topics, sub-topics, and the diverse contributions by authors from all around the world, it was possible to identify the existence of some leading themes emerging more significantly. About this and going back to the main purpose defined for this chapter (To have an identification of main topics and themes related to unemployment currently discussed (emergent) and the ones that endures on the field), we have reason to believe that the delimited main purpose was responded. Responding to the delimited purpose as well the research question presented on table 1 and on this chapter, introduction is also answered.

Emergent topics unveiled by our bibliometric analysis included unemployment insurance, youth unemployment, monetary policies, COVID-19, and entrepreneurship. Thus, we have some evidence that unemployment phenomenon has inside it other themes that are not less important and, in fact, could be a driving force on the unemployment rate composition. These identified core themes suggest some pathways for practitioners and academicians to conduct future research to better understand unemployment whereas in this thesis scope, these topics are the starting point to assess potential influential factors on unemployment rates composition.

The bibliometric analysis was useful to reveal some inherent characteristics on the unemployment-related literature that are insightful beyond the emerging themes. For example, it was possible to identify that geographically most of studies produced about the theme are being made from high developed economies, such as United States, Germany, and United Kingdom. This is a relevant output that suggest we should move away from these countries when performing our cross-country analysis, aiming to ascertain nations that are usually neglected on the types of studies we analysed here.

Most productive authors within our retrieved database of documents (according to H-index) are Marco Caliendo, Zimmerman, K. and Rinne, U. Caliendo especially has been consistently appeared as one of the leading authors on unemployment thematic. Still on our merged sample containing 2.334 documents, highest cited studies are Glaeser, Scheinkman & Shleifer (1995), discussing economic growth and its relationship with labour market; and Blanchflower (2000) that assess self-employment as an alternative for unemployment. Meyer (1990), analysing unemployment insurance, and Shimer (2005), with the continuum between employment and unemployment are highly cited as well.

Sources that frequently are publishing studies about unemployment are Labour Economics, Intereconomics and Journal of Monetary Economics, suggesting that usually these studies are developed by an economic approach, illustrating the opportunity to shift the observing to a business and managerial perspective, as we intend to offer in our study. Overall, all findings on this chapter helped to perceive complexities on unemployment phenomenon.

Co-word analysis on keywords, thematic evolution and trending terms within the field provided useful insights on this topic that could be further developed from here. This chapter is an attempt to conceive a solid presentation about unemployment-related studies on the business and management areas. Results of this study are not necessarily proposed on originalism basis but complements well the past literature on this research field.

Nonetheless, this research has its own limitations that are important to acknowledge. Although the Scopus and Web of Science repositories, the ones used here, being composed by studies published and rigorously reviewed to ensure their quality, it is possible that other relevant research have been not reached on the working sample of documents analysed here. The usage of only two databases could be dismissing other relevant documents that are not included in neither of the repositories. Other sources such as Google Scholar and EBSCO, for example, could be used to complement the information we are offering here.

Limitations arise also due to the methodological proceeding. Findings here are extracted essentially from one method only, by the application of software R-Studio features as the package `bibliometrix` and the interface `biblioshiny`. On R-environment alone there are other packages (CITAN, ScientoText, H-index Calculator and Scholar) beyond the one we used, which could implicate on different results than the ones we presented. Other software's could perform a similar analysis, such as CitNetExplorer, VOSviewer, SciMAT, and BibExcel.

Considering the limitations we just presented findings offered on this chapter are not presumed as absolute. Potentially, other bibliometric studies, applied with different tools would indicate distinct results. Nevertheless, even taking into consideration the potential biases and limitations that naturally permeates every research, we believe that this chapter is insightful to assess the unemployment-related literature and develop the sequential chapters that continues this thesis.

3. ASSESSING UNEMPLOYMENT DETERMINANTS: CROSS-CLUSTERING ANALYSIS USING WORLD BANK COUNTRY-DATA IN A PAST DECADE (2012-2021)

3.1. Introduction.

Building upon about we presented in the previous chapter, it is reasonable to understand the importance of a better comprehension about labour market complexities, being unemployment one of these complex factors in it. Labour market is regularly a reliable presentation about economic momentum and overall market efficiency (Bosna, 2022) despite the unity of analysis, if a city, a state, a country, a region or even the entire world. Hence, labour market components are such important factors that some (e.g., Bayrak & Tatli, 2018) argue that labour force, employed and unemployed levels, cannot be presumed exclusively as productive variables, but also as strategic determinants.

A better knowledge about labour market constituents could be a primordial input for decision-making process by any country in any moment, especially due the direct or indirect effects that labour has with other markets such as the capital market, the market of production and offering goods or services, for example (Draskovic et al., 2021). Labour market however despite the importance it's not an easy task to understand. At minimal, has a twofold avenue of analyses: Employment and unemployment. For our research intention and scope, we are focusing on the unemployment side of the labour market.

We decide to focus on unemployment in particular believing that a better comprehension about those excluded from labour opportunities could have for themselves as well in a wider scope a more damaging outcome, that could result in enormous social cost and costs of welfare-reduction (Bell & Blanchflower, 2011; Çelik & Lüküslü, 2018; Bayrak & Tatli, 2018). From this perception, this chapter will be dedicated for an econometric approach to obtain a better understanding of unemployment, particularly the rates of unemployed people.

About these rates in specific, World Bank refers to International Labour Organization definition, which defines unemployment rate as the share of individuals in a giving labour force that is without work, at an observed moment of time, but are available for and actively seeking for employment allocation. Although general ideal on this concept may seem simple, there are several factors that could contribute for a better or a worse interpretation of these rates when analysing it.

Shimer (1998), for example, discussed that even if all unemployed persons were eligible to apply for a job, the conditions available, or not, for their interest on job allocation are beyond their own hands. Furthermore, within this absolute share of jobless people could also exist some individuals that are not employed by discretionary choice, considering that their qualifications and the job availability are not well aligned (Montgomery et al., 1998).

Complexity on the analysis of unemployment rates arise also when it is included as unemployed those individuals waiting to return for a job after being discharged from a previous occupation (Javaid, Akbar, & Nawaz, 2018). Not to mention individuals that are not pursuing for a job allocation because their reasons (Choudhry, Marelli & Signorelli, 2012), as a better qualification of themselves up to be receiving insurances for their unemployment status, for example.

Factors mentioned so far do not extinguish all complexities about unemployment rates, are mere presentation on how intricated the understanding, therefore empirical analysis about it could be. Even more intricacies may emerge due unexpected social or economic phenomenon's that could spill effects on labour market, employment, and unemployment levels. Coronavirus pandemic, for example, affect unemployment all over the world, since governments have enforced social distancing, businesses being suspended of operate and a severe shrink on the demand by non-essential workers (Couch, Fairlie & Xu, 2020).

This COVID-19 scenario apart is important not only because the ongoing effects, considering the moment we are writing this thesis, and the significant emergence of this term on the previously applied bibliometric analysis in chapter 2. Coronavirus affected everyone in every aspect, including job relations and labour market. Nonetheless, it is not the intention for this chapter to exclusively endeavour in efforts to explain unemployment rates in this pandemic context, especially because the outcomes of this phenomenon are far to be settled and completely measured.

Considering the presented so far and our intention to better understand the unemployment portion of the labour market, research question to guide this chapter emerge as: What are the main determinants that composes the unemployment rates? Responding to this question we as well intend to compromise with the main objective to develop a supranational cross-country analysis that enables the identification of determinants composing the unemployment rates.

The cross-country portion of the defined objective will be applied using a clustering technique to establish countries relationship without forcing them into predefined groups. Country's data will be extracted from the World Bank and the premise is to use as many countries as possible, believing that unemployment rates composition would be more assertively explained.

About the assessment of potential determinants composing unemployment, a Vector Error Correction Modelling (VECM) will be conducted to ascertain the influence of each determinant (independent variables) on the dependent variable (unemployment rate), allowing the identification of which factors are more influential and how they could be treated for decision-makers have the possibility to confront unemployment rates with more assertiveness. Details about variables selected, data extraction, collection, analysis, and methodologies applied on this chapter will be further presented on following sections. Literature review, a methodological, results and discussions, and a conclusions section continues this chapter.

3.2. Literature review.

This section will focus on some factors that, derived from the related literature, could be presumed as potential determinants on the composition of unemployment rates. Structurally, this topic will be divided into subtopics referring to each determinant that will be later used.

3.2.1. Economic growth and unemployment: Okun's Law.

Arthur Okun in his seminal study proposed that 1% change in the unemployment level would be correlated with an approximate 3% of change in overall economic productive outputs, respecting opposite directions on these fluctuations, if one is elevating the other is decreasing and vice-versa (Okun, 1962; Lee, 2000). 'Okun's Law' have been consistently analysed by several studies after its proposition although not necessarily confirming the original values of one and three percentual variations but still validating the57ssociateion between unemployment and productivity, usually proxied by gross domestic product (GDP) (Prachowny, 1993; Cuaresma, 2003).

Okun's (1962) proposition have some real-life practical applicability. Not rarely public policies and governmental decisions are made presuming that unemployment rates could be reduced and influenced through management stimulus or constraint onto productivity (Marconi, Beblavý & Maselli, 2015). A counterargument exists, however. Blinder (1997) suggest that Okun's premise is built upon an atheoretical assumption, considering that the presumed 3-to-1 relationship on GDP and unemployment tend to be wrong in practice (Blinder, 1997).

Observing literature documented about empirical usage of this theory, since its original proposition, a large number of studies could be found building upon the seminal Okun's (1962) work. Some studies extending the original cyclicality between unemployment and productive capacity, whereas some studies are more aligned with Blinder (1997) criticism, contrasting, or adapting the first premises on what would be later known as a law (Huang et al., 2019).

Zanin (2014), for example, found that the estimated Okun coefficients are not always significant when observing the effects in specific subgroups but tends to be higher on youth individuals. Butkus et al. (2020), examining a 28-sample of European countries, obtained that other factors are influencing the unemployment and productivity relationship, such as age, gender, and educational levels (Butkus et al., 2020).

Guisinger et al. (2018) assess the labour market influential power on the Okun's presumed relationship, unveiling that more flexible labour markets proxied by higher levels of education, lower rate of unionization, and higher share of non-manufacturing employment, would be more productive, dampening the unemployment rates.

Overall, it seems natural that some level of cyclicality surrounds unemployment and economic growth, being this an inherent characteristic of these variables, more important would be the extension of this feature on empirical analyses by economists, managers, policymakers, banks, and many other stakeholders interested to know how the unemployment rates of various labour force groups are sensitive to economic fluctuations (Zanin, 2018).

Considering the non-exhaustive literature examples presented so far and how this topic goes deep in further discussions, the rapport of economic growth (sometimes here referred as "output") and unemployment seems undeniable. We build on this relationship well-based theory, particularly by Okun (1962), to select economic growth as a potential composing part of unemployment rates.

Economic growth, that we will be measuring using GDP as a proxy (Prachowny, 1993; Cuaresma, 2003), also emerge on the trending topics on unemployment-related literature results from the early applied bibliometric analysis in figures 15 and 16, solidifying our selection to use on potential explaining unemployment rates later, on our empirical assessment using the Vector Error Correction Modelling method.

3.2.2. Inflation and unemployment: The Phillips Curve.

Another recurrent classical theory that relates unemployment rates with macroeconomic variable, similar as the Okun's (1962) proposition, is the Phillips Curve econometric modelling. Named after William Phillips, this theory presumes a correlation of reduction in unemployment whereas rates of wage within a country's economy are increasing (Phillips, 1958). Therefore, considering that some nation has any changes in its unemployment levels, this would implicate in a predictable effect on price inflation indexes (Phillips, 1958; Bayrak & Tatli, 2018).

In a Nobel worthy contribution study and building on the earlier proposed by Phillips, Milton Friedman confirms that the relationship between inflation indexes and unemployment are aligned (Friedman, 1977). If inflation increases, unemployment tends to decrease, and vice versa. This relationship being often represented by a downward-sloping curve on a graph with these two variables, indicating that policymakers face a trade-off between these two factors.

Several studies have been adapting, extrapolating, and overall assessing the concept originally proposed by Phillips. Stock & Watson (1999), for example, focusing on inflation side of the relationship, found that when forecasting future values for inflation, Phillips (1958) premises, suggesting an accompanying variation by unemployment, tend to be more assertive factor in comparison with other variables. Phillips idea however is not immune to some criticism, Karanassou, Sala & Snower (2010), for example, contest what they call "conventional wisdom", especially with the "stagflation" happening after the 1970s.

Indeed, inflation indexes are very particular depending on the country and its economic situation under analysis. Thus, it would not be surprisingly to obtain different results even using the same original proposition by Phillips (1958) study. For example, Warsame et al. (2022) examining the Somalia's context have results confirming the premises of Okun's law but contradictory with Phillips curve, unemployment and inflation varying in the same direction whereas Fratianni, Gallegati & Giri (2022) considering United Kingdom found negative and stable relationship between money wages inflation and unemployment, converging with original Phillips premises.

For the purposes for this chapter, we acknowledge distinctions inherent for each country reality, sharing some form of medium-run relationship between inflation and unemployment, similar as the one found on Fratianni, Gallegati & Giri (2022) research. Meaning that the association among these variables above all exists, if this influence is strong or not remains to be conferred on the empirical analysis later developed.

In summary, as the Okun's law (1962) Phillips curve have a solid grounded literature sustained the selection of inflation as a potential determinant of unemployment rates composition (e.g., Bryan & Cecchetti, 1993; Boskin et al., 1998; Bayrak & Tatli, 2018; Cepal, 2020). Inclusion of inflation on our assessment is also justified by bibliometric analysis especially considering the direct relationship that inflation have with monetary policies, for better or worse, given that public administration manipulates inflation levels to cope with unemployment as Phillips (1958) himself argued.

3.2.3. Insurances, beneficial programs, and unemployment.

So far, our later presented model to assess the unemployment rates composition has two variables on studies that also test or empirically analyse this phenomenon: Economic growth (here proxied by GDP) and inflation. Both variables appeared as emergent themes on the bibliometric analysis presented on the previous chapter of this thesis and we remain going from these results to select other variables, more specifically for this subtopic: Unemployment insurance.

Unemployment insurance programs and overall policies to financial aid are a complex subject of analysis. Engen & Gruber (2001) study have results showing that this could be a twofold avenue; while could increase overall welfare, being an income to those without a job, could also deter household accumulation and savings, lagging the economic cycle. Thomas & Worrall (2007) observed the moral hazard about the receiving of insurance by unemployed people, discussing if this type of benefits should come from public or private mechanisms to mitigate eventual disfunctions on the income distribution.

Insurance benefits for unemployed people has a well-documented presence in academic literature. Nakajima (2006), for example, examinates how the extension, both in time and values, of insurances could lead to voluntary unemployment. Suggesting that is a major task for policymakers to establish an optimal level to not incur in the possibility of shirking for labour by individuals. However, this is not an easy task, Moyen, Stähler & Winkler (2019) proposed a model for a theoretical optimal unemployment coverage policy, considering a general equilibrium where a group of countries decide to introduce a common unemployment insurance system.

This matter is from recent interest by academics and policymakers, as we could notice on our bibliometric analysis proceeding, urging other researchers to follow the Moyen Stähler & Winkler (2019) proposition. Some authors (e.g., Ignaszak, Jung & Kuester, 2020; Ábrahám et al., 2023) have been dedicating effort to unveil what could be an ideal metric to have a balance of non-neglect beneficial assurances whereas not deterring the cyclicality of employment and unemployment rates.

We believe that this theme is indeed an emergent matter, particularly following the COVID-19 pandemic that escalated the number of people in need for insurance income after losing their jobs by the economic crisis. Despite the intricacy of this topic go beyond our intended scope of research we do not disregard the relevance.

In fact, unemployment insurance appeared with some consistence on the bibliometric analysis applied earlier and building on that we intend to select this as a potential variable to explain unemployment rates in our modelling. We acknowledge however that insurance measurement is not easily available; our intention is to use this as a variable indeed, but that will depend on the availability of information we could or could not retrieve from World Bank.

3.2.4. Youth unemployment and its importance on overall unemployment rates.

Youth unemployment have been a matter of interest extensively research over the years as we could perceive when performing our bibliometric analysis (e.g., Hammarström, 1994; Mroz & Savage, 2006; Awogbenle & Iwuamadi, 2010; Caliendo & Schmidl, 2016; Bayrak & Tatli, 2018; Marques and Hoerisch, 2020). It is indeed a theme within the wider unemployment subject that deserves some particular attention about it. United Nations (UN) defines, and we align with it, a youth person objectively as those on the 15 to 24 years old range being therefore the youth unemployment referring to this share of population not allocated on the labour market.

West (1987) noticed how youth unemployment was persistent and suggested some public policies to mitigate the problem, exemplifying with the implemented Participation and Equity Program by the Australian federal government and how it was better than expected good for Australians youth people. Mroz & Savage (2006) assessed long-term effects of early unemployment and how this could affect incidence and extension to later and persistent unemployed status.

If indeed a persistent long-term influence exists, in a short or medium time it would not be different. Issue of youth unemployment tends to increase specially in moments of great economic and social adversities, as the one we recently experienced with COVID-19 pandemic outbreak. Unemployment grows in general in these types of situations, but this growing occurs in a faster pace considering the youth population (Gough, Langevang & Owusu, 2013; Barford, Coutts & Sahai, 2021).

A consistent increase in youth unemployment rates has been occurring since the 2008 economic crisis, both globally and within countries, but this has accentuated significantly since February 2020 with a higher rate on the loose of jobs by young people, more particularly young women living in middle-income countries (Barford, Coutts & Sahai, 2021). This is illustrated on Peñaloza & Rincón (2022) for example, analysing informal economy, unemployment, and other features on the Ecuador post-pandemic scenario.

More developed economies have not suffered less however, unemployment for youngers is also a damaging effect from Coronavirus pandemic, as observed by Fana, Pérez & Fernández-Macías (2020), finding that countries like Spain, Italy and United Kingdom are more likely to suffer by implications of the confinement measures giving that these countries, when COVID-19 exploded, were already facing high unemployment and high temporary work levels among youth population.

It could be the case for youth unemployment be in isolation a study topic by itself. Tomić (2018) for example has a very analogous approach as the one intended on this chapter, exploring main determinants of youth unemployment in specific whereas our approach is to observe unemployment in wider scope but using the youth portion as a latent variable composing overall unemployment. Results on Tomić (2018) resonates some of the before mentioned potential determinants, there was found that GDP growth and decreasing are consistently related with levels of youth unemployment.

As already delimited, this chapter's proceeding is to observe youth unemployed as a part on the overall unemployment rates composition, our hypothetical inference is this factor would be one of the main determinants to explain the broader unemployment. When observing youth portion of unemployment with the already selected GDP, inflation, insurance programs and other variables later explained we believe that will be possible to achieve a more extensive comprehension of the labour market and unemployment phenomenon.

3.2.5. Entrepreneurship, self-employment, and unemployment.

As we proceeded in previous topics, our intention is not to overextend in explanations about entrepreneurial activities nor conceptualisation on this subject. There is a large literature dedicated for this and spend much time on this matter would escape from the scope of this chapter. However, we shed some light on a few studies that specifically discusses entrepreneurship related with labour market and unemployment. Also, this was another theme that emerged from the previous applied bibliometric analysis, extracted from there to further exploration on the sequential study we are developing here.

Conceptually, Earle & Sakova (2000) suggest that a self-employed person could be both an individual improving products or service while working or someone unable to find an opportunity on the labour market, being therefore forced to create a job for itself. Thurik et al. (2008) goes further on the relationship of self-employment and unemployment, observing a sample of OECD countries, authors found that variation in unemployment levels have a positive impact on the self-employment rates whereas the opposite has a negative impact, being the latter association (self-employment to unemployment) stronger.

Laffineur et al. (2017), focusing on the role of Active Labour Market Policies (ALMP), defends this is the most effective manner to promote entrepreneurship while coping better with elevated unemployment rates. Results suggests that countries more flexible on their labour market institutions should allocate resources for promotion of entrepreneurship activities as they already have a favourable scenario for better outcomes of their ALMPs (Laffineuer et al., 2017). It is argued that policymakers must be aware that ALMP are a good driver for necessity-based entrepreneurs although may not be that good for opportunity-driven entrepreneurial efforts (Laffineur et al., 2017).

Overall, and considering the studies we have been assessing since the bibliometric analysis, it would be an understatement to believe that entrepreneurship has a strong relationship with unemployment. There is in fact a growing relevance on the so-called unemployment-entrepreneurship nexus as a driver for job creation when governments and policymakers are facing momentum or perennial crisis on labour market (Cheratian et al., 2019).

As the purposes of our research is not on the complexities about entrepreneurship-unemployment relationship, we will acknowledge entrepreneurial activities in a broader definition and use the self-employed rates as a proxy for entrepreneurship and insert it on the potential determinants composing unemployment rates. Therefore, we finalise on the factors that emerged from earlier bibliometric analysis, having for now: GDP, inflation indexes, unemployment insurances benefices, youth unemployment rates and self-employment. Five potential determinants to one target variable are already a solid, nonetheless, we decided to add others to have a more robust equation. These other variables are summarized in the next topic before to proceed for the methodology section.

3.2.6. Other potential determinants for unemployment rates.

First, it is important to acknowledge that does not exist an ideal number of determinants to explain any form or type of a phenomenon. This is particularly correct when observing a high complexity factor as unemployment. In the end, is predominantly a researcher decision on which and how many factors is better to work with. It could be the case that the already presented five determinants (GDP, inflation, unemployment insurance, youth unemployment and self-employment) could be sufficient. Nonetheless, as we intend to have a wide-ranging scope while maintaining a feasible accessible data, we move on adding some other variables that could increase reliability and a robust assessment of unemployment rates composition.

Building on premises from previous studies related with our subject of interest (e.g., Lentz & Mortensen, 2005; Pissarides & Vallanti, 2004; Uzay, 2005), there is a suggestion about productivity as one of the main drivers influencing labour force. If productivity is high, labour demand tends to reduce while a low-level productivity could lead to more job-opportunities (Bayrak & Tatli, 2018). Relationship on productivity and labour force is well documented in the literature and considering that this data is easily available we add productivity as one of the variables that could better explain unemployment rates.

Other recurrent variable is a financial one, the level of savings by individuals and how this influences their working statuses (Bayrak & Tatli, 2018). Neoclassical economists largely discuss the influences of savings as potentially being the main determinant for economic growth (Metzler, 1951; Bayrak & Tatli, 2018; Ribaj & Mexhuani, 2021). Similar as for productivity, if the relationship exists on the employment side of the labour force, here meaning that people being able to save more income would be less inclined to accept dissatisfying job-opportunities, it is reasonable to presume a connection existing on savings and unemployment. Once again, this data is available in World Bank and therefore saving has been included in the mix of variables.

When reviewing the unemployment-related literature, beyond the topics already discussed from subtopics 3.2.1 to 3.2.5, another aspect is recurrently appearing related with our theme: Education. Nickell (1979), for example, proceed with one of the earliest estimations on the impact of schooling and the probability of entering unemployment, founding that individual's level of education tends to reduce the spell on unemployment. Riddell & Song (2011) results shows that a schooling period from 12 to 16 years impacts higher on re-employment for jobless workers.

Results on this relationship not always positive on the perspective that more education reduces unemployment. Altindag, Dursun & Filiz (2022), for example, found that higher-educated individuals, when receiving unemployment insurances, tend to use this income for a longer time being less likely to leave unemployment status. Reckoning the influence that education may have on unemployment and re-employment we decided to add this as a determinant for unemployment rate as well, depending on the consistence of information available to work it.

It would be possible to go and on about which or how many variables could offer a solid presentation of the unemployment phenomenon and its rates, and we still would not exhaust all the intricacies of unemployment composition. Adding to the already discussed variables, there is a plenty of institutional factors such as wage systems, employment protection legislation, labour taxes, labour mobility and overall active labour market policies that could be largely influent on unemployment and labour market (Nickell, Nunziata & Wolfgang, 2005; Bayrak & Tatli, 2018).

Demographic variables also can play a significant role and add even more complexity. Population density, migration flows, ethnicity, and gender, all of those are potential determinants to unemployment levels (Pissarides & McMaster, 1990; Bayrak & Tatli, 2018). As we mentioned before, the addition of variables could be overly extensive and still not achieve a definitive answer. Nonetheless, based-on the previous bibliometric analysis and some additive inputs literature-based, we believe to have a makeable number of potential determinants to be assessed and better understand unemployment rates composition using, at first, GDP, inflation, unemployment insurance, youth unemployment, self-employment, labour productivity, savings, and education.

Intending to work with these variables, the assumption is that it will be possible to accomplish a solid response to the before stated chapter's research question: What are the main determinants that composes the unemployment rates? On responding to that we are as well aiming to fulfil the chapter main objective: Develop a supranational cross-country analysis to identify main determinants composing unemployment rates. The eight variables usage will be dependent on availability of consistent information among the many as possible number of countries to proceed with a supranational analysis. About that and further methodological procedures discusses the following section.

3.3. Methodological procedures.

This section is subdivided in four topics aiming to describe the proceedings followed to achieve the main objective designed for this chapter. First, is presented how the dataset was built. Aiming to perform a cross-country examination on unemployment determinants, the starting point is to describe which and how many countries will be analysed. Second, will be detailed the equation and first theoretical econometric model to assess the determinants that may compose unemployment rates. Third section explains data cleaning proceedings to create a definitive dataset. Finally, on the fourth section, will be depicted the econometric proceedings up to the application of VECM method.

3.3.1. Building the dataset.

Contemplating all the presented in introductory section and to satisfactorily respond the defined research question while respecting the selected variables mentioned in the section before, an analysis of annual cross-country data will be applied considering a 10-year timespan going from 2012 to 2021. Some guiding questions to be answered before the effective method application: Which countries would be included in the analysis? Following what parameters? How reliable and how much of data is available? These questions and how we assess them for have a solid dataset will be further discussed.

Giving that the idea is to achieve the most informative and consistent results on unemployment phenomenon as possible, we decided to leave the availability of data to delimit the number of countries to be analysed. Also, with the intention to access information on a diversified group of countries, the data was retrieved on World Bank repository, the Databank. Databank and all its material available are free of charge while much of the information comes from statistical systems of member's countries, providing reliable and good-quality data in one aggregate repository (World Bank, 2023).

Initial retrieving of countries' level data is regarding to the employment rates, although all our scope being the counterpart of employment. This two-way avenue is particularly important giving that any analysis of unemployment phenomenon that does not include employment could be a research malpractice misleading an adequate understanding of unemployment phenomenon (Margo, 1993; Farber et. al, 2019). Therefore, we initially assess the employment side of the labour force to later deepening ourselves into our interested targeted variable, the unemployment.

Hence, first data extracted from Databank, dating from February 23, 2023, is referring about employment to population ratio[1], a spreadsheet containing a raw presentation of working-age people inserted on the labour force considering the years we defined as timespan (2012-2021). This initial data was pre-assessed and treated with the removing of elevated missing information, aggregate data, non-quantitative information, and all out of our interests' inputs to have a primary working sample of 185 countries.

Considering the now restrained sample of 185 country-level employment information, we move forward to clustering technique that we decide to use to supranational and cross-country analysis as alluded on our introductory section of this

[1] Available at: https://data.worldbank.org/indicator/SL.EMP.TOTL.SP.ZS

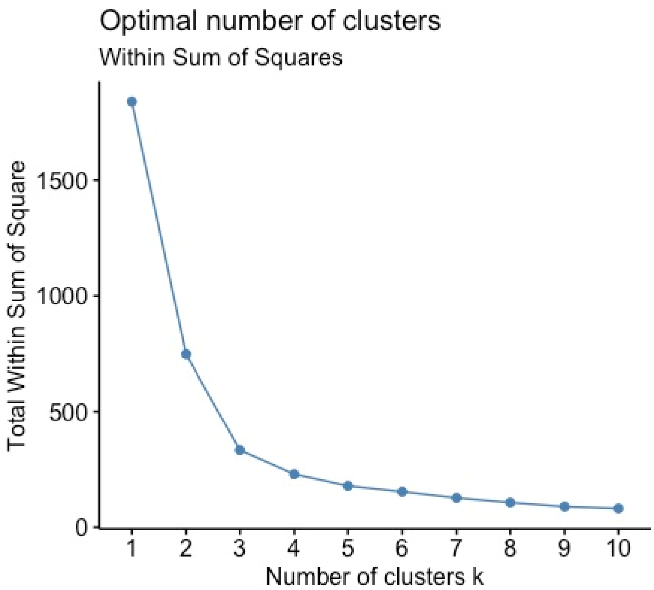
chapter. Our intention is to be clustering countries considering their employment levels from later assess these clusters created about unemployment rates composition, therefore we have a suggestion of both sides (demand and supply) of the clusters' labour market.

Clustering is a tool for grouping observations that was proposed and applied for decades. Cluster analysis is an unsupervised technique useful to create subsets, in a larger set of data, in a manner that those in the same subset, the cluster, are more alike within the group than with the ones participating in other subsets (Jain, Murty & Flynn, 1999; Tsai et al, 2019). Considering our research purpose, we are relying on cluster proceeding to aggregate countries that have most similarities, considering as parameter their employment to population ratio, despite their geographic region not being the same.

Clustering technique is a solid approach given that geographical distribution may not be the most accurate manner of aggregate countries. There is in fact more similarities between Brazil and Ecuador, both geographically in South America, than among Brazil and Russia? An analysis of their medium values of employment in the dataset gives that from 2012 to 2021 Brazil has approximately 56.3% of people employed, and Ecuador has 63.0% while Russian Federation has 59.3%, suggesting more similarity between the Brazilian and Russian labour markets. This is an illustrative comparison and there are numerous factors that could lead to nations' grouping. Nonetheless, we believe it is reasonable to use clustering to have an initial insight on countries' resemblances.

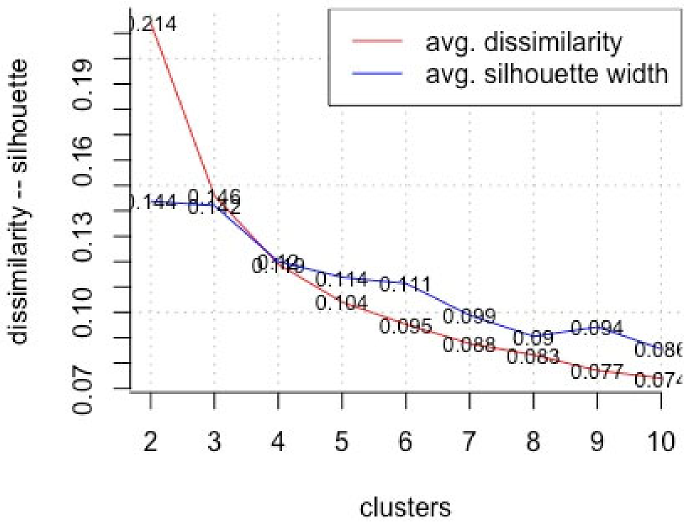
Most used algorithm to cluster is the K-means (Tsai et al, 2019), which perform a partition in data sample (n) into clusters (k) where each observation, in our case each country, belongs to the cluster with the most approximate mean with itself data (Hartigan & Wong, 1979; Frey & Dueck, 2007; Tsai et al, 2019). We used the Within Sum Squares (WSS) method on the software R-Studio to have a visual suggestion of an ideal number of clusters considering our 185 countries' sample. Figure 17 illustrates this assessment.

Figure 17
Optimal number of clusters by Within Sum of Squares.
(Elaborated by the author).



WSS visual presentation on figure 17 as we can perceive suggests as ideal number of clusters, *k*, 3. Therefore, our dataset building has an initial indication to be formed by three clusters. To complement the solely visual signal, we also run a dissimilarity analysis, again using R-Studio to check if 3 clusters is indeed reliable. Figure 18 shows the statistical and graphic results of this application.

Figure 18
Dissimilarity analysis.
(Elaborated by the author).



Once again, observing the dissimilarity values and lines presented in the figure 18, the *k* = 3 seems the most accurate to proceed. 3 is the point where values in lines red and blue are very approximate (0.146 and 0.142 respectively) without overlap between themselves as happens on the case of 4 clusters. Other possibilities as work with 5, 6, 7 and more clusters gives a larger distance between the lines which would not be ideal (Gowda & Diday, 1991).

All country-data numbers extracted from Databank regarding employment to population ratio are in percentage values and a scaling proceeding were applied to not incur in potential dysfunctional groupings of countries with higher or lowest percentual values. Therefore, clustering procedure were made with the scaled numbers imputing the maximum of three clusters.

Considering that our clustering is using employment to population ration while the targeted interested of our study is the unemployment rates composition, clusters were labelled on labour market terms, since the expression gather both employed and unemployed individuals. Three clusters subsampling of countries is depicted on the following table 10.

Table 10
Clusters and countries distribution.
(Elaborated by the author).

CLUSTER	COUNTRIES *	COUNTRIES WITHIN CLUSTER	MEDIAN VALUES OF EMPLOYMENT **
Low Labour Market	AFG, BIH, COM, DZA, EGY, ESP, GAB, GRC, GUY, HRV, IND, IRN, IRQ, ITA, JOR, LBN, LBY, MAR, MDA, MKD, MNE, MRT, NAM, NPL, PNG, PRI, SDN, SEN, SRB, STP, SWZ, SYR, TJK, TKM, TUN, TUR, YEM, ZAF.	38	-1,51
Medium Labour Market	ALB, ARG, ARM, AUS, AUT, AZE, BEL, BGD, BGR, BLR, BLZ, BRA, BRB, BRN, BTN, BWA, CAN, CHI, CHL, CIV, COG, COL, CPV, CRI, CUB, CYP, CZE, DEU, DJI, DNK, DOM, ECU, EST, FIN, FJI, FRA, GBR, GEO, GIN, GMB, GNB, GNQ, GTM, GUM, HKG, HND, HTI, HUN, IRL, ISR, JAM, JPN, KGZ, KOR, LAO, LCA, LKA, LSO, LTU, LUX, LVA, MDG, MDV, MEX, MLT, MNG, MUS, MYS, NCL, NGA, NIC, NLD, NOR, PAK, PAN, PHL, POL, PRT, PYF, QAT, ROU, RUS, RWA, SAU, SLB, SLE, SLV, SOM, SSD, SUR, SVK, SVN, SWE, TCD, TGO, TON, TTO, TZA, UKR, URY, USA, UZB, VCT, VEN, VIR, WSM, ZMB, ZWE.	108	0,05
High Labour Market	AGO, ARE, BDI, BEM, BFA, BHR, BHS, BOL, CHE, CHN, CMR, COD, ERI, ETH, GHA, IDN, ISL, KAZ, KEN, KHM, KWT, LBR, MAC, MLI, MMR, MOZ, MWI, NER, NZL, OMN, PER, PRK, PRY, SGP, THA, TLS, UGA, VNM, VUT.	39	1,22
First Sample of Countries	**185**		

Notes: * Countries are listed in alphabetical ordination respecting to their respective IDs and according to the World Bank definition. Available at: https://data.worldbank.org/country.
** Is referenced to the median values within the cluster, which is formed by each country values about employment to population ratio from 2012 to 2021 timespan. Values are presented in their scaled form and non in the percentage numbers as originally retrieved from Databank repository.

Table 10 clusters' labels are not presuming one group of countries as better or worse than the other. Names assignments are illustrative, and we use the observed values of employment to population ratio as parameter to each cluster. Low labour market (LLM) group are the lower scaled values, the countries were the proportion of employed people in the overall labour force is noticeably small, higher unemployment therefore; Medium labour market (MLM) is composed by transition countries, employment rates are around mid 50 % up to high 60 % of employed people in the labour force; and High labour market (HLM), has higher scaled values for employed individuals, unemployment is lower for these countries in comparison.

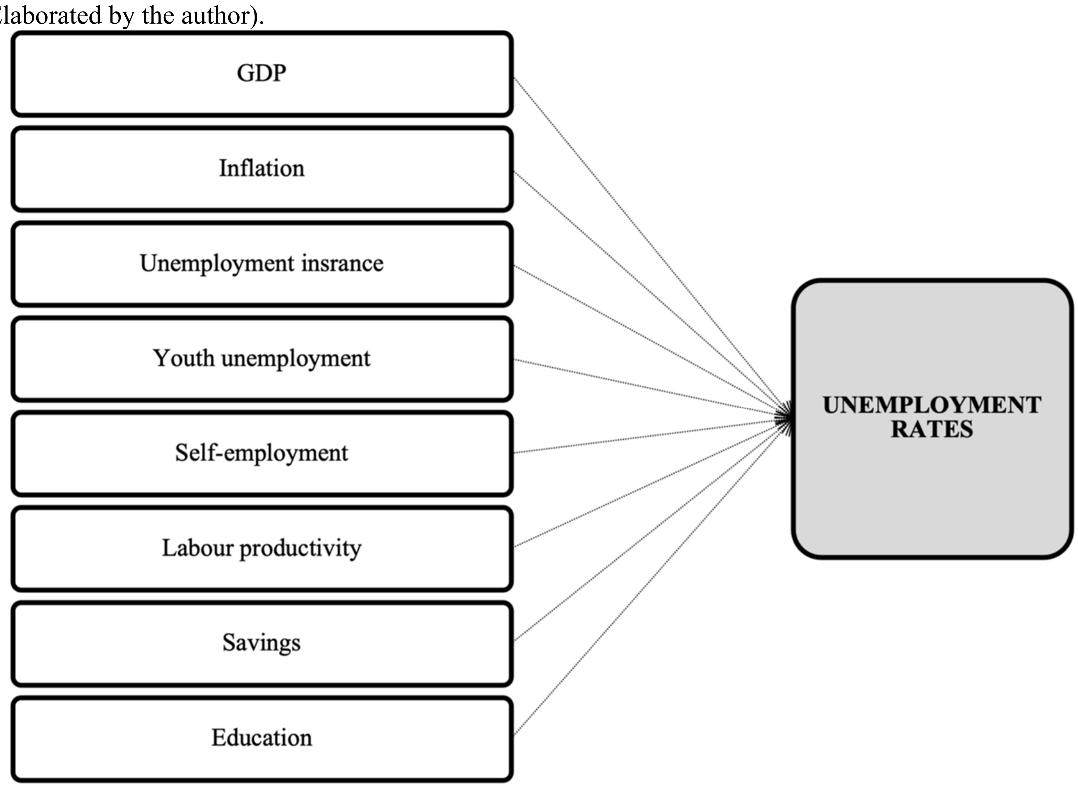
Beyond the referred visual and statistical results presented on figures 17 and 18, three clusters seem consistent, even with a larger group. Clustering considering the employment to population ratio it is reasonable as most countries are neither super high or low with employed people, suggesting that most countries would be in a mid-term (a larger cluster) whereas other countries suffering with low employment are in a pole whereas the ones more stable considering population and employed people in the other end. Thus, the building of an initial sample for cross-country examination is completed and further analysis will be possible to be performed considering now the similarities and distinction between the clusters and, from hereafter, their respective unemployment levels.

3.3.2. Methodological pre-proceedings.

Main technique to be used in this chapter is the Vector Error Correction Model (VECM). VECM have been used consistently as a well fit modelling when the premise is to estimate influential power of exploratory factors into a dependent variable of interest. Considering unemployment-related studies, the procedure has also been of use as in Brüggemann (2006), Patel & Choga (2018) and Sevencan (2023) for example, from these and from others relatable analysis we decided to replicate this method here.

Considering the usage of the VECM technique, the concepts for variables selected and depicted on the literature review section, and earlier the emergent themes presented on the chapter 2 of this thesis, it is possible to conceive a conceptual framework for our intended unemployment rates composition. The basis are the eight potential determinants we believe to explain and overall unemployment in our supranational cross-country analysis. This conceptual structure is illustrated in the following figure 19.

Figure 19
Conceptual framework.
(Elaborated by the author).



As we acknowledged before, other variables could be added in the framework we are proposing in figure 19; also, some of these eight selected factors could produce a reasonable study in isolation or even be removed due non-availability of data. Nonetheless, it is our belief that a solid background both in number variables, eight, and presumed relationships is offered. GDP, inflation indexes (INF), unemployment insurance benefits (UI), youth unemployment (YU), self-employment rates (SER), labour productivity (LP), savings (SAV) and education (EDU) are therefore the initial explanatory variables in our panel data analysis using VECM considering the timespan of 2012–2021.

Unemployment is the target variable, and the others are independent, potentially influencing unemployment rates composition. Functional form of this is the following: $UR = f(GDP, INF, UI, YU, SER, LP, SAV, EDU)$. Table 11 summarizes our initial variables.

Table 11
Variable description and source.
(Elaborated by the author).

Variable	Description	Units	Source	Variable ID on World Bank
UR	Unemployment rate: Share on the overall labour force (15+ years old) without job but available and seeking for it.	Annual percentage.	World Bank	SL.UEM.TOTL.ZS
GDP	Gross domestic product as a proxy for economic development and Okun's Law. Annual percentage growth rate of GDP at market prices based on constant local currency.	Annual percentage.		NY.GDP.MKTP.KD.ZG
INF	Inflation indexes as a proxy for the Phillips Curve. Inflation as measured by the consumer price index.	Annual percentage.		FP.CPI.TOTL.ZG
UI	Percentage of population participating in unemployment compensation, severance pay, and early retirement due to labour market reasons Estimates include both direct and indirect beneficiaries.	Annual percentage.		per_lm_alllm.cov_pop_tot
YU	Share of the labour force on ages 15 to 24 without job but available and seeking for it.	Annual percentage.		SL.UEM.1524.ZS
SER	Workers who are working on their own account or with one or a few partners or in cooperative. Have the type of jobs defined as a "self-employment jobs."	Annual percentage.		SL.EMP.SELF.ZS
LP	Proportion of the population ages 15 and older that are economically active, supplying labour for production of goods and services.	Annual percentage.		SL.TLF.CACT.ZS
SAV	Gross domestic savings calculated by the GDP less final consumption expenditure.	Annual percentage		NY.GDS.TOTL.ZS
EDU	Considers the ratio of the total labour force with advanced education in working-age and unemployed.	Annual percentage.		SL.UEM.ADVN.ZS

Table 11 summarises variables we select to further assessment plus unemployment rates and a brief description of each of the factors, their original unity of analysis and the source from where we retrieved their information. Bearing the presented-on figure 19 and table 11 our intention is to use annual values for each variable, every country in every year of the defined timespan (2012-2021).

All values retrieved from World Bank repository of data, the Databank, will be scaled to use distinct country's level of data proportionally. About this, Databank aggregates information for many sources, such as ILO, OECD, and specific countries entities. Although we are referring to World Bank as main source, information in practice comes from different origins.

Moving on to the econometric methodology approach to coordinate these variables, up to this point and considering the theoretical presumptions presented in the figure 19, our intention is to test long-run association between the set of eight variables with unemployment rates using the following model of equation:

$$\mathrm{UR}_{\mathrm{C_t}} = \beta_0 + \beta_1(\mathrm{GDP}_{\mathrm{Ct}}) + \beta_2(\mathrm{INF}_{\mathrm{Ct}}) + \beta_3(\mathrm{UI}_{\mathrm{Ct}}) + \beta_4(\mathrm{YU}_{\mathrm{Ct}}) + \beta_5(\mathrm{SER}_{\mathrm{Ct}}) + \beta_6(\mathrm{LP}_{\mathrm{Ct}}) + \beta_7(\mathrm{SAV}_{\mathrm{Ct}}) + \beta_8(\mathrm{EDU}_{\mathrm{Ct}}) + \mu_{\mathrm{Ct}} \quad (\mathrm{Eq.\,1})$$

Whereby:

- Ct is a country on the sample and a given period correspondingly.
- UR is the dependent variable unemployment rate.
- $\beta 0$ is a constant in the equation.
- μ is the error term.

GDP represents the annual gross domestic product growth rate, INF is for the inflation rates, UI considers the percentage of population participating in unemployment compensation programs, YU the share of youth unemployed in overall labour force, SER self-employed rates, LP proportion of the population economically active, supplying labour, a proxy for productivity, SAV is for gross domestic savings and EDU the ratio of people on total labour force with advanced education and in working-age.

Figure 19 conceptual representation and table 10 described variables are still on the range of theoretical inferences. How much of these intended data are in fact available considering our already delimited 185 countries sample? Is indeed the Equation 1 presented earlier feasible and with sufficient data available to proceed? We assess these questions on the following.

As depicted on table 10, all variables information is to be extracted from the World Bank repository of data, the Databank, using respective variable identification. All data was downloaded directly into the software R-Studio by using the package `WDI`[2] and later transformed into a Microsoft Excel spreadsheet for cleaning proceedings. `WDI` package allows users to search and download from over 40 datasets hosted by the World Bank.

[2] See https://vincentarelbundock.github.io/WDI/#/ for further information.

On the Excel spreadsheet and using the function ‘COUNTIF’ we identified some missing values (NAs) considering the variables presented in table 11. These results are detailed on the following table 12, where variables are arranged from the ones with less absent data to the one with most information missing. Our procedure will be to observe variables’ information first and later the variables data and country level information.

Table 12
Checking for missing data.
(Elaborated by the author).

Variable	Missing values (NAs)
UR – Unemployment rates	2
YU – Youth unemployment	2
LP – Labour productivity	2
SER – Self-employment rates	2
GDP – Gross domestic product	65
INF – Inflation	185
SAV – Savings	211
EDU – Education	851
UI – Unemployment insurance	1719

All information described at table 12 and the missing values we are indicating is referring to the moment we extract the data (22 May 2023) that will be used from hereafter. This is important to be acknowledged as World Bank updates data regularly and probable retrieving information in another moment values in the table 12 may differ. We extracted data covering the 2012 up to 2021 timespan for each of the considered variables and the initial sample of 185 countries presented in the table 10. Table 12 indicates a major problem at unemployment insurance variable, having 1.719 missing values.

Further analysing the data, we may notice that are few countries with available information regarding their unemployment insurance programs. Most of nations with this data are from countries on the Latin America & Caribbean region, such as Argentina, Brazil, Costa Rica, El Salvador, Panama, and Uruguay. However, even in these countries, none of them have data available for our complete 10-years timespan (2012-2021).

There is no clear guiding for the decision to drop or maintain a variable; nonetheless, unemployment insurance is an evident problem. Number of data missing is way larger than available information and the data that exists are incomplete and centred in a small sample of countries. Therefore, measurements to cope with missing information (Madley-Dowd et al., 2019) would almost create an artificial variable which lead us to decide to exclude UI from our model; we do this to not deviate our analysis for misleading interpretation due the usage of non-real data.

Variables with less NAs are easily to assess not going for the exclusion. For UR, YU, LP and SER variables, the missing data are on Afghanistan and Myanmar in 2021 year. Could be the case of information regarding these variables not being available when data for these countries were comprised by World Bank or when we retrieved information from there. Considering that is a 1-year missing data in a 10-year timespan, we imputed the median value considering the 9 previous years of information available.

Missing data on the remaining variables demand a more in-depth process to deal with it. The 65 NAs on GDP are mostly due to countries that have none of this information available, such as Channel Islands, Eritrea, New Caledonia, and Korea Dem. People's Rep., that have no data regarding gross domestic product considering in the entire 10-year of retrieved information. These mentioned countries are by themselves responsible for 40 of the 65 missing values on GDP variable. Further observation about these countries indicates that beyond GDP they have a good amount of missing information on inflation, savings, and education.

Having this non-availability in many of our selected variables, we decide to exclude these countries from the initial sample of 185 nations. Removing Channel Islands, Eritrea, New Caledonia, and Korea Dem. People's Rep., the missing value on GDP drops to 25 and we identified the cases of South Sudan (6 NAs on GDP and SAV and 10 on EDU) and Venezuela (7 NAs on GDP, SAV and EDU and 5 on INF), these two countries were also excluded and missing data on GDP falls to 12 and we treated this number following the same approach adopted for UR, YU, LP and SER.

To summarize our cleaning data process, countries with no information at all or that has a high number of missing values, considering the 10-year of intended data on two of the eight selected exploratory variables were also excluded. According to this filtering, Afghanistan, Argentina, Congo, Cuba, Guam, Guyana, Comoros, Lao PDR, St. Lucia, Liberia, Libya, Malawi, French Polynesia, Papua New Guinea, Puerto Rico, Somalia, Suriname, São Tome and Principe, Syrian Arab Republic, Tajikistan, Trinidad and Tobago, St. Vicent and The Grenadines, Virgin Islands (U.S.), Turkmenistan, and Yemen were also removed from the sample.

After all the first imputation of median values when exists data available to be referred, the removing of a total of 31 countries with elevated missing values, and the exclusion of one variable (unemployment insurance) on our initial model, we were able to deal with absent information on UR, GDP, INF, YU, SER, LP and SAV variables, none of these have the absent values initially depicted on table 12. However, a problem remains regarding the education (EDU) variable which even after the initial proceeding has a total of 562 missing values within the sample that now is composed by 154 countries.

If the proceeding to be performed was the exclusion of countries that have a high number of missing values in EDU, the sample of countries would decrease from 154 to 88, largely affecting the supranational cross-country we intend to apply. Considering that, our decision is to remove EDU, giving that the remotion of it would be less damaging tan the loss of more than a half of complete country-level data.

All cleaning process concluded; some updates must be made to reflect the basis on which our empirical tests will be applied. Updated formal function now is the following: $UR = f\,(GDP, INF, YU, SER, LP, SAV,)$ and the equation model for VECM application is:

$$\begin{aligned}\mathrm{UR}_{\mathrm{C}_t} = \beta_0 + \beta_1(\mathrm{GDP}_{\mathrm{Ct}}) + \beta_2(\mathrm{INF}_{\mathrm{Ct}}) + \beta_3(\mathrm{YU}_{\mathrm{Ct}}) + \beta_4(\mathrm{SER}_{\mathrm{Ct}}) + \beta_5(\mathrm{LP}_{\mathrm{Ct}}) \\ + \beta_6(\mathrm{SAV}_{\mathrm{Ct}}) + \mu_{\mathrm{Ct}}\ (\mathrm{Eq.}\,2)\end{aligned}$$

Whereby Ct is a country on the 154 nations' sample considering a certain period, UR is the dependent variable unemployment rate, $\beta0$ a constant and μ is the error term. GDP as gross domestic product growth rate, INF is inflation rates, YU share of youth unemployment, SER for self-employed rates, LP proportion of the population economically active, supplying labour for as a proxy for productivity and SAV is gross domestic savings.

Conceptual framework presented on figure 19 are not disregard but will, due data availability, remain only on conceptual level. Other research, using other countries and data source may use this framework but for this chapter we continue with the information we were able to solidly achieve.

3.3.3. Econometric methodological steps.

After data cleaning and with an updated equation model it is the moment do describe the econometric steps necessary before the application of the Vector Error Correction Modelling (VECM). Some procedures must be taken into consideration to achieve a solid basis for the method usage and, consequently, to have a reliable presentation of unemployment phenomenon. These mandatory steps before the VECM usage are described in the following paragraphs and their results will be presented on the results section to not overextend our methodological topic.

First assessment to be made is regarding unit root tests. This is important because when dealing with time series or panel analysis as the one intended to use here, if data are not complying with stationarity results and further analyses may be spurious (Choi, 2001; Shabbir, Kousar & Alam, 2021). There are several statistical tests to check for unit roots and stationarity of a panel data, however as our intention is to remain using R-Studio we will proceed with tests recurrently used in this software. Function `purtest` in the package `plm` offers seven possibilities to unit root tests being Levin, Lin & Chu (2002) and Im, Pesaran & Chin (2003) the most used and will be used here as well.

Next important step when dealing with econometric modelling is the choosing regarding the lags of the model. Most used test for lag-order determination are Akaike Information Criteria (AIC) and Schwarz Criteria (SC). As suggested on Stock & Watson (2012) the selection of a lag order when lesser than the ideal could imply on loss of data whilst an over selection number could overfit the model. Once again, the option for select an optimal lag length will be performed using R-Studio on two manners, through the packages `tsDyn` and `vars`. Ideally the optimum number of lags will be the same on both methods, suggesting consistency on the dataset being analysed.

Having tested unit roots and assessing optimal lag selection, a cointegration analysis follows. Essentially, cointegration checks the existence or not of long-run association between variables under analysis; if a cointegration is observed this suggests that the modelled variables will behave together in a similar pattern when in a long-term analysis (Shabbir, Kousar & Alam, 2021).

Three most well-known tests to check long-run association are: Pedroni (1999) and Kao (1999) tests, both following the residual-based approach; and the Johansen (1991) cointegration test, which uses the likelihood method to find out the cointegration among variables (Örsal, 2007). To assess cointegration the option selected for this study was the Johansen's (1991) maximum likelihood approach, considering that is a more complete test and the application on similar studies as the developed by us (Patel & Choga, 2018; Shabbir, Kousar & Alam, 2021). Johansen test demands the ideal lag length to be predefined and is calculated using the function `ca.jo` from the library `urca` on software R-Studio.

Having passed the unit root and cointegration assessments, it is possible to proceed with the usage of Vector Error Correction Model (VECM) enabling to recognise both long-run and the short-term relationship among our selected variables and unemployment rates.

Vector Error Correction Modelling may be used when there are at least two variables in the model which are stationary on the first difference, which is assumed to be the case here. VECM equation model for this study considering the remaining seven variables is presented on the following Equation 3:

$$\Delta UR_{C_t} = \alpha_1 + \sum_{\rho}^{i=1} \beta_{1_i} \Delta GDP_{t-1} + \sum_{\rho}^{i=1} \beta_{1_i} \Delta INF_{t-1} + \sum_{\rho}^{i=1} \beta_{1_i} \Delta YU_{t-1} + \sum_{\rho}^{i=1} \beta_{1_i} \Delta SER_{t-1} + \sum_{\rho}^{i=1} \beta_{1_i} \Delta LP_{t-1} + \sum_{\rho}^{i=1} \beta_{1_i} \Delta SAV_{t-1} + \lambda_1 ECT_{t-1} + \mu_{it} (Eq.3)$$

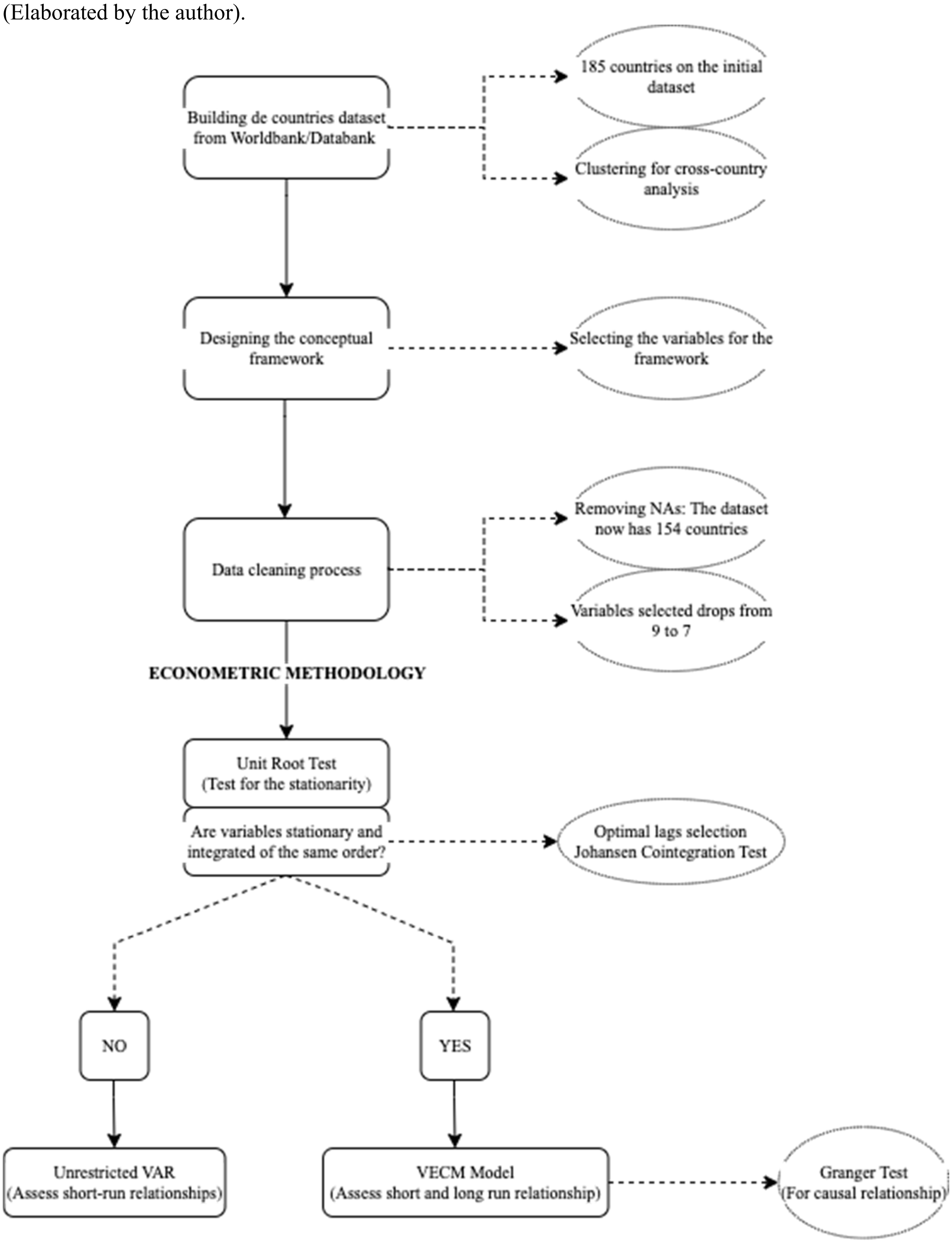
Included on the equation is overall unemployment rates, as the variable of interest on ΔUR_{C_t} and the already known and discussed variables regarding GDP, inflation, youth unemployment, self-employment rates, labour productivity and savings. Furthermore on the equation are μ_{it} for eventual random errors on the equation, ECT_{t-1} represents the integrating vectors, λ is an adjustment coefficient for disequilibrium on the year-before being analysed, α_1 is a speed coefficient adjustment regarding the long-run correction terms while on the β are the individual variables' coefficient in the error correction term.

An extra step that we will use considering this thesis later intention to forecast or anticipate some tendencies, is to verify whether an economic variable would help the estimation of another. Determination of this causal direction can be analysed following the approach given by Granger (1969). Variable X is a "granger cause" of Y if variable Yt stays as expected using previous terms regarding the variable Xt (Granger, 1969; Shabbir, Kousar & Alam, 2021). Results for the Granger test as well the other tests we are referring at this subtopic will be presented on the results section of this chapter.

3.3.4. Methodology flow.

To close this section after the final dataset building, cleaning process of this data and the econometric methodology steps to be followed, we may finally move on to the Vector Error Correction Model application. Before the application per se, figure 20 summarizes the proceedings we present on this section and most of what will be presented results presentation in the next section.

Figure 20
Methodological flow.
(Elaborated by the author).



3.4. Results and discussions.

3.4.1. Descriptive statistics.

Table 13
Descriptive statistics values.
(Elaborated by the author).

	Mean	Median	Maximum	Minimum	Standard Deviation	Skewness	Kurtosis	Jarque-Bera (X-squared)	Probability (p-value)	Observations
UR	-0.000765	-0.331828	4.112.095	-1.295.026	1.000.126	143.348.818	4.818.830	739.69	0.00000	154
GDP	0.000464	0.077882	8.043.095	-11.733.595	0.9993018	-156.398.713	23.467.550	27509	0.00000	154
INF	-0.000750	-0.119152	27.499.946	-0.435224	0.9961742	2.003.422.731	480.846.678	14754671	0.00000	154
YU	-0.000804	-0.251036	4.951.260	-1.337.232	1.000.173	138.164.303	5.541.810	904.53	0.00000	154
SER	0.0006326	-0.2212683	20.977.197	-14.771.550	0.9999833	0.44171753	1.936.508	122.65	0.00000	154
LP	0.000141	0.016596	2.768.432	-3.177.386	0.9996904	-0.04368222	3.580.861	22.14	0.00001	154
SAV	0.001483	0.019280	3.427.099	-4.909.006	1.000.393	0.03909135	4.298.474	108.58	0.00000	154

Note: Some variables originally refer to absolute percentage values whereas others are on the differences considering the year (*t*) and the year before (*t-1*). Aiming to establish a common unit of analysis between all variables, we transformed all indexes to their staggered format.

On this section we are presenting the results that emerged from our application of the Vector Error Correction Model (VECM) technique. Preceding steps to be complied before the method usage, as we described at the last section of this chapter, will be presented first and after that the presentation of VECM specific findings analysis. Table 13 that initiates this section illustrates descriptive values for a brief summarization of our data. These descriptive statistics are useful to identify general characteristics on data both individually, for each variable, and for the dataset in a wider perspective.

We will be assessing a dataset of 154 observations – countries – considering a timespan ranging from 2012 to 2021. Each of these countries having data respective to six exploratory variables plus their unemployment rates, as presented on table 13. For our complete sample of countries, the unemployment rates have a mean value by -0.000765 with a relatively high standard deviation of 1.000.126 and maximum and minimum values of 4.112.095 and -1.295.026, respectively.

Elevated deviation from the mean as well high distances from maximum and minimum values reflects well how distinctive are the countries composing the dataset on their unemployment levels, complementing well the clustering results previously presented in table 10. Statistics of Jarque–Bera (1980) and corresponding p-values are used for testing the normality assumption of the data (Cromwell, Labys & Terraza, 1994) and considering the presented for all variables under analysis we have evidence to conclude that the data we are further assessing does not follow a normal distribution, enabling the viability of the model proposed-on equation 3.

We do not need to overextend about the results from table 13 as they are already depicted in the table, but some initial insights are possible to infer from this descriptive presentation. Overall, we have a consistent dataset regarding the variables although some of them seem to detach for the others. Inflation, for example, have the highest maximum values, suggesting a significant discrepancy between countries regarding some of them being suffering more with inflationary pressures than others. Specific level of influence of inflation on unemployment rates as well for the other variables will be checked throughout this section.

3.4.2. Panel unit root results.

Table 14
Panel unit root test results.
(Elaborated by the author).

Variable	**Levin, Lin & Chu test**		**Im, Pesaran & Chin test**		**Alternative hypothesis**	**Decision**
	Z-statistic	***p-value***	***Wtbar statistic***	***p-value***		
UR	-13.546	0.00000	-3.1098	0.00093	Stationarity	Reject the null hypothesis
GDP	-23.02	0.00000	-23.953	0.00000		
INF	-17.1	0.00000	-21.243	0.00000		
YU	-11.805	0.00000	-4.8951	0.00000		
SER	-11.222	0.00000	-6.7285	0.00000		
LP	-8.5177	0.00000	-2.1005	0.01784		
SAV	-20.07	0.00000	-16.383	0.00000		

Note: Null hypothesis tests the existence of a unit root on the variable's data or, in other words, if an observed is non-stationary. Assuming a significance value on 5%, all variables described on the table are under this threshold, therefore, the null hypothesis may be rejected.

Table 14 results confirms the premise of stationarity when observing all selected variables, including unemployment rates. Which, therefore, enables the VECM application. It was assessed statistics for each variable through all 154 countries in the sample using both Levin, Lin & Chu (2002) (LLC) and Im, Pesaran & Chin (2003) (IPS) tests. Complying with the stationarity is an essential step to avoid spurious results and considering the presented in table 14 we may assume, with a 95% of confidence (even upping it to around 99%), none of the variable is non-stationary.

Having a validation for the stationarity of the analysed variables is an important premise before any further test as the VECM. Both LLC (2002) and IPS (2003) tests applied, and their results presented on table 14, checks the rejection or non-rejection of the existence of at least a unit root in the selected variables, suggesting that our data and is well suited for the application of the Vector Error Correction modelling technique, if they also met the cointegration analysis that will be presented later.

3.4.3. Optimal lag selection.

As already defined on the methodology section, lag selection will be performed using R-Studio software and its packages `tsDyn` and `vars`. Optimal lag selection for this study will involve the delimited six determinants as well the dependent variable, unemployment rate, on the entire sample of 154 countries. Table 15 presents our results.

Table 15
Lag length selection criteria.
(Elaborated by the author).

Lag	Log (L)	AIC	SC	HQ	FPE
1	-9.870592 *	-10.57951	-10.38431 *	-10.50686 *	0.00002543188
2	-10.14617	-10.59501*	-10.22902	-10.45880	0.00002504074 *
3	-10.24986	-10.57007	-10.03329	-10.37030	0.00002567327
4	-10.3395	-10.53060	-9.823015	-10.26726	0.00002670746
5	-10.38295	-10.48992	-9.611538	-10.16301	0.00002781717

Note: Table comprises results emerged on the usage of both R-packages selected to obtain the optimal lag number. Log (L) is extracted log-likelihood, AIC is Akaike Information Criteria, SC for Schwarz Criteria, HQ for Hannah-Quinn criteria and FPE for final prediction error.
* Is for the most statistically recommended lag selection.

Five applied and presented tests to identify optimal lags, as presented on table 15, has three suggesting 1 lag as ideal and two others suggesting for 2 lags. Specifically, about the most adopted criteria, Akaike Information Criteria (AIC) recommends a lag selection equal to 2, both in `tsDyn` and `vars` R-packages usage while per Schwarz Criteria (SC) the ideal lag should be equal to 1 also in both of procedures.

We presume that one, as optimal lag length, could be too little. Furthermore, considering that the VECM modelling when performed on the R-Studio packages demands at least 2 lags to proceed, for our study the decision is in favour of two lags as ideal. Overall, the criterion adopted on most studies (e.g., Liew, 2004; Shabbir, Kousar and Alam, 2021) comparable with this one suggests two as the optimal length, and, therefore, further proceedings will be applied as the optimal length being two lags.

3.4.4. Panel cointegration results – Johansen test.

Cointegration tests are mandatory to assert long-run relationship among a set of variables. Considering that variables are cointegrated, it is possible to relate and combine them using a linear approach, knowing that even if or when eventual shocks in a short-run analysis affect these variables individually, they will converge in time (in the long-run). If variables are stationary and cointegrated, the VECM could be applied. Results of the Johansen test using Eigenvalues for our working variables is presented on table 16.

Table 16
Johansen test for integration via max-Eigen method.
(Elaborated by the author).

Hypothesized number of CE (r)	Test statistics	10pct	5pct	1pct
r ≤ 6	75.46	7.52	9.24	12.97 *
r ≤ 5	83.59	13.75	15.67	20.20 *
r ≤ 4	94.66	19.77	22.00	26.81 *
r ≤ 3	109.28	25.56	28.14	33.24 *
r ≤ 2	133.26	31.66	34.40	39.79 *
r ≤ 1	381.51	37.45	40.30	46.82 *
Null hypothesis or r = 0	595.75	43.25	46.45	51.91 *

Note: Null hypothesis: There is no cointegration. CE is cointegrated equations. Comparison of cointegration must be made through test statistics values and the critical values percentages.
* Cointegration exists in this equation at this level of percentual confidence.

Johansen's test the null hypothesis (H0) that there are no cointegration among the variables and interpreting the outputs presented in table 16 we have found that the test of hypothesis from r = 0 (the null hypothesis), there is no cointegration between the variables, from a full cointegration between all variables being analysed (when r ≤ 6), presents significance. Results on table 16 enables the rejection of null hypothesis of no cointegration within the variables.

In fact, cointegration exist up to r ≤ 6, meaning that is reasonable to assume with at least 90% until 99% of confidence that the seven variables (UR, GDP, INF, YU, SER, LP and SAV) analysed here are cointegrate among themselves in a long-run analysis. This result matters to propose a reliable model that resists that even in the occurrence of unexpected shocks, such as inflationary spikes, increasing on gross domestic product or other changes that may influence unemployment rates and labour market overall.

Working with variables in a timeseries that passes unit root tests and are cointegrated let the interpretation about the selected factors to understand unemployment rates are indeed correlated with the phenomenon. Meaning that, the established relationship between our observed variables does not exist due to some chance and further analysis about it could be continued by the VECM application.

3.4.5. Long-run relationship of variables: The VECM application.

Vector Error Correction Modelling s an established technique that assess long-run equilibrium relationship between no stationary variables as well their short-run adjustment behaviour. Short-run part of this modelling indicates the speed of adjustment to restore the equilibrium of the model, suggesting how fast or slow variables can return to their natural equilibrium. Therefore, ideally coefficients of the VECM must be statistically significant and have a negative sign.

VECM overall equation model for this chapter considering the 7 variables we discussed and selected previously follows the before presented and now reinforced equation 3:

$$\Delta UR_{C_t} = \alpha_1 + \sum_{\rho}^{i=1} \beta_{1_i} \Delta GDP_{t-1} + \sum_{\rho}^{i=1} \beta_{1_i} \Delta INF_{t-1} + \sum_{\rho}^{i=1} \beta_{1_i} \Delta YU_{t-1} + \sum_{\rho}^{i=1} \beta_{1_i} \Delta SER_{t-1} + \sum_{\rho}^{i=1} \beta_{1_i} \Delta LP_{t-1} + \sum_{\rho}^{i=1} \beta_{1_i} \Delta SAV_{t-1} + \lambda_1 ECT_{t-1} + \mu_{it} \text{ (Eq. 3)}$$

The presented in the above equation is our general model, considering the entire working sample of 154 countries. As was applied a clustering process for our intended supranational cross-country analyses, the idea is to replicate the same equation within the three clusters early defined. Hence, in practical ways four equations and their indexes will be observed: Equation 3.1, for overall sample; equation 3.2 for the cluster named low labour market; equation 3.3 for medium labour market and equation 3.4 to high labour market (see table 10 for the cluster assignments).

VECM we are presenting the results on the sequence was performed using the software R-Studio and the R-function `cajorls`, which is a built-in feature of the library `urca`, available in different versions of the software R. Parameters in `cajorls` returns the Ordinary Least Squares (OLS) regressions of a previously restricted VECM model, in our case, the one proposed on equation 3 regarding unemployment rates. Following table 17 presents the results extracted from `cajorls` considering the complete dataset of 154 countries.

Table 17
Long-run relationship of the VECM general equation.
(Elaborated by the author).

Variables	Coefficients	Standard error	T-values
GDP (-1)	-0.031595	0.082091	-0.385
INF (-1)	+0.118283	0.071254	1.660
YU (-1)	+0.0598575	0.0397405	1.506
SER (-1)	-0.0161193	0.0346806	-0.465
LP (-1)	-0.0073723	0.0366554	-0.201
SAV (-1)	+0.017461	0.044790	0.390
CointEq1 (UR)	-0.084891	0.038132	-2.226
Constant		0.007377280	
R-Squared Adjusted		0.0605	
F-statistic		8.618	

Table 17 reports long-run coefficients, standard errors, and t-values for the VECM that considers the entire sample of 154 countries. Error correction terms (ECTs) coefficients can be interpreted as the rate of convergence in a current period (*t*) toward the equilibrium in a following period (*t* + *n*). Values described in CointEq1 on table 17 indicates the relative speed of adjustment in which the dependent variable (UR) returns to equilibrium due to an eventual exogenous shock.

For this first and for the future analysis by cluster an ideal scenario would have the error correction term must as negative and significant, which indicates the movement of dependent variable toward equilibrium. A positive sign would suggest a movement of dependent variable away from equilibrium (Patel & Choga, 2018; Shabbir, Kousar & Alam, 2021). Furthermore, ECTs values must be between 0 and 1, where 0 indicates no adjustment, while 1 represents a 100% adjustment.

Going back for the depicted-on table 17, the results for unemployment rates shows a value of -0.08, suggesting around 8% velocity of adjustment for this variable toward unemployment equilibrium in sequential years. Other values obtained for UR on table 17 presents a low standard error by 0.03 and a t-value of -2.26, indicating the coefficient for UR as statistically significant, as it is meaningfully different from zero.

Following equation 3.1 indicates results of long-run relationship between our selected variables and unemployment rates considering the presented-on table 17.

$$\begin{aligned} \mathrm{UR}_{\mathrm{Ct}} = {} & 0.007377280 + (-0.031595)\mathrm{GDP}_{\mathrm{t-1}} + (0.118283)\mathrm{INF}_{\mathrm{t-1}} \\ & + (0.0598575)\mathrm{YU}_{\mathrm{t-1}} + (-0.0161193)\mathrm{SER}_{\mathrm{t-1}} + (-0.0073723)\mathrm{LP}_{\mathrm{t-1}} \\ & + (0.017461)\mathrm{SAV}_{\mathrm{t-1}} + \mu_{\mathrm{it}} \text{ (Eq. 3.1)} \end{aligned}$$

Findings presented on table 17 and on equation 3.1 suggests that inflation, youth unemployment and savings have a positive association in the long run with unemployment rates whereas gross domestic product, self-employment rates and labour productivity have a negative association with our dependent variable.

Coefficient value for inflation shows that if we increase one unit of inflation, unemployment rates would also increase by 0.118283 units. Meaning is in an economic scenario where inflation rates are rising, implicating uncertainties and distortions in country's markets, this tends to reduce hiring or laying off workers to adjust their costs to inflationary pressures, resulting in a rise on unemployment rates.

Recent studies bring to discussion the importance of the association between inflation and unemployment, particularly in the long run as being analysed at this point. Tenzin (2019), for example, alerts to the idea that if inflation is not well monitored or controlled, it could propagate uncertainties leading to lowering in investments, a lower economic growth, thereby implying on a higher unemployment rate.

Youth unemployment coefficient indicates that if an increment of one unit on YU occurs, this would mean a rise on rates of unemployed by 0.0598575 units. One possible interpretation is that the portion of youth without a job affect the total labour market in the same direction. Association of youth and overall unemployment is well documented on the literature and naturally they should fluctuate accordingly, despite the signal (Mroz & Savage, 2006). Preventive methods to tackle unemployment at early stages of labour life may reduce the extent in which individuals spend their working lives unemployed (Gregg, 2001; Görlich, Stepanock & Al-Hussami, 2013; Kawaguchi and Murao, 2014).

Another factor found having positive association with unemployment rates is the variable savings. Coefficient number for SAV presents that if an increase by one unit of savings happens, unemployment rates would also rise by 0.017461 units. Suggestion is that when individuals can increment their own savings, this potentially could lead to a high unemployment. One potential explanation for this is that incomes are being saved this may implicate on less consumption, which may decrease demand for goods and services, adversely affecting economic activity and creating less job opportunities.

There are some studies aligned with this result, such as Lentz & Tranaes (2005) for example, analysed wealth effects and labour market, suggesting that levels of savings and other financial resources being accumulated impact the job-search of an individual. Those with reduced financial pressure (more savings) can be more selective in their job-search, therefore increasing unemployment rates in some level. Levenko (2020) found that higher levels of perceived uncertainty, particularly on the labour market (more unemployment) could lead to increased household savings.

Assessed the positive associations, we go back to results depicted on table 17 and mathematically presented on equation 3.1 referring to the three negative interactions for variables gross domestic product, self-employment rates and labour productivity. A one-unit increase in gross domestic product (GDP), this would be associated with a decrease of 0.031595 units in the unemployment rate. This relationship is in line with the typical expectations of classical economic theories, as a growing economy often (although not necessarily) leads to an overall increase on job opportunities and decreasing unemployment rates (Okun, 1962; Prachowny, 1993; Cuaresma, 2003; Simionescu, 2020).

Moving to the next negative association, the coefficient of -0.0161193 on self-employment rates, likewise in GDP, suggests an inverse relationship between SER and UR. A one-unit increase in the lagged self-employment rates corresponds to a decrease of 0.0161193 units in the unemployment rate. A negative association implies that if levels of self-employed individuals are high, levels of unemployment will be reduced. This result is aligned with the most usual theoretical approaches, including some of already referenced on this study (Thurik et al., 2008).

Although the increasing on self-employment levels could eventually implicate on reduced unemployment as the data 154 countries data shows, not necessarily this implies in a good effect. Millán, Congregado & Román (2012), for example, suggests that individuals transitioning from unemployment to self-employment tends to have a low longevity of their self-employment ventures, due to factors such as limited resources, lack of entrepreneurial experience, and potential financial constraints. Not to dismiss the results we found, but only to reinforce on how the complexities on self-employment and unemployment association could be.

Last variable to be assessed, the labour productivity (LP) coefficient value being of -0.0073723 shows that if we increase one unit of overall labour productivity, this will correspond to a decrease of 0.0073723 units in the unemployment rate, implicating an inverse relationship between these variables. Premise is that when labour productivity increases, this indicates that workers already occupying a job position are producing more output per unit of labour input. Hence, higher labour productivity often leads to increased efficiency which can contribute to lowering unemployment rates.

This association as obtained here confirms some studies mentioned earlier (e.g., Lentz & Mortensen, 2004; Pissarides & Vallanti, 2004; Uzay, 2005; Bayrak & Tatli, 2018), suggesting higher productivity as a driver for reduced rates of unemployment. More recent studies although aligned with the results obtained here alerts that the association of productivity and unemployment has its own limitations and not always are easily envisioned Crowder & Smallwood (2019), for example, suggests that in practice a positive productivity shock on unemployment could be overstated.

After analysed all dependent variables proposed to explain unemployment rates, it is possible to confirm that all of them have a significant, although some negative and other positive, effects on the dependent variable in the long run. Table 17 also indicates other important statistics about VECM, such as the adjusted R-squared of 0.0605, indicating that 6.05% variation in unemployment is due the projected explanatory variables, which is a relatively a low percentage but considering the complexities of unemployment phenomenon is still significant. F-statistic by 8.618 value indicates that the overall the model we propose is reliable.

So far, all the results and discussions presented are considering table 17 and equation 3.1 which are referring to the overall sample of 154 countries. On the following paragraphs will be assessed the results considering the three clusters (see table 10 for the clusters composition). Proposition is to observe if, and how, the clusters are diverging or complying with the overall sample of countries analysed.

Following equation model 3.2 is considering the cluster we label as “Low Labour Market” (LLM):

$$\Delta UR_{LLMC_t} = \alpha_1 + \sum_{\rho}^{i=1} \beta_{1_i}\Delta UR_{t-1} + \sum_{\rho}^{i=1} \beta_{1_i}\Delta GDP_{t-1} + \sum_{\rho}^{i=1} \beta_{1_i}\Delta INF_{t-1} + \sum_{\rho}^{i=1} \beta_{1_i}\Delta YU_{t-1} + \sum_{\rho}^{i=1} \beta_{1_i}\Delta SER_{t-1} + \sum_{\rho}^{i=1} \beta_{1_i}\Delta LP_{t-1} + \sum_{\rho}^{i=1} \beta_{1_i}\Delta SAV_{t-1} + \lambda_1 ECT_{t-1} + \mu_{it} \text{ (Eq. 3.2)}$$

Statistical proceedings used on equation 3.2 is the same applied on equation 3.1, therefore the analysis remains similar as well. Important to reinforce that the initial total of countries, and therefore their attribution into clusters, were updated during the data cleaning proceedings described on the methodology section. Hence, if we observe the presentation on table 10 some countries were removed of their clusters due to absence of data or other problems discussed and presented on section 2 of this chapter.

First, and specifically to the updated number of countries that are included in the Low Labour Market, this cluster had 38 countries in the initial raw data retrieval and after the cleaning proceedings this number fall to 26. This is the larger dropping of countries’ number comparing with the other two clusters. Table 18 presents once again results extracted from `cajorls` function on R-Studio and a summary of values regarding the long-run association of the model presented on equation 3.2 for the Low Labour Market cluster.

Table 18
Long-run relationship of the VECM equation – Low Labour Market cluster: Version 1.
(Elaborated by the author).

Variables	Coefficients	Standard error	T-values
GDP (-1)	-0.03922	0.01581	-2.481
INF (-1)	+0.19801	0.02121	9.336
YU (-1)	+0.009194	0.009317	0.987
SER (-1)	+0.010759	0.005244	2.052
LP (-1)	+0.0029327	0.0046157	0.635
SAV (-1)	-0.024884	0.008851	-2.811
CointEq1 (UR)	0.012070	0.009624	1.254
Constant		0.81971126	
R-Squared Adjusted		-0.0195	
F-statistic		0.4518	

Results on table 18 shows a more complex scenario comparing with the overall model coefficients depicted in table 17 and equation 3.1. First, we have a negative and low R-squared adjusted by -0.0195 indicating that considering the countries on Low Labour Market cluster, the proposed exploratory variables are not completely suited to explain their unemployment rates. F-statistic of 0.4518 also suggests that this model, having all variables in it is not a suitable modelling. Although may be a simplistic reasoning, remains relevant significant to mention that within this cluster is a considerably small sample of countries.

Cluster size however is an indication, but we will not assume this as a prohibitive to modelling. We deepened our analysis for the exploratory variables and assuming as a significant result p-value ≤ 0.05, it was possible to notice that for LLM cluster model, not all variables are statistically significant. Our results indicate: GDP and INF have a p-value of 0.00000; YU had a p-value of 0.9354, similar happening in LP, with 0.943; SER has a p-value of 0.02318 and SAV a p-value of 0.149.

Hence, p-value analysis suggests that for this model three variables have statistical significance while other three does not. If removing the non-significant variables (YU, LP and SAV) we have a second and updated values for VECM coefficients, standard errors and t-values that are depicted on following table 19.

Table 19
Long-run relationship of the VECM equation – Low Labour Market cluster: Version 2.
(Elaborated by the author).

Variables	Coefficients	Standard error	T-values
GDP (-1)	-0.065957	0.061102	-1.079
INF (-1)	-0.01739	0.08264	-0.210
SER (-1)	+0.010759	0.02023	0.570
CointEq1 (UR)	-0.01154	0.03593	-4.139
Constant		-1.232363	
R-Squared Adjusted		0.05083	
F-statistic		2.974	

We have overall improved values in table 19 if comparing with table 18. Values on CointEq1 coefficient indicates the relative speed of adjustment in which the dependent variable (UR) returns to equilibrium, therefore, error correction term (ECT) have a negative sign presenting the movement of dependent variable toward equilibrium, here it is specifically by 1% (0.01154). A relatively low standard error of 0.03593 and a t-value of -4.139 indicates a significant coefficient value, as it is meaningfully and different from zero.

On the coefficients of exploratory variables, we found self-employment rates having a positive and significant association in the long run with unemployment rates, whereas gross domestic product and inflation have a negative association behaviour with the dependent variable. In comparison on what was obtained for the general model (depicted on table 16), indexes are similar, although some of the signs of association differ. GDP remains with a negative association; INF had a positive association for the complete sample while here presents a negative one; SER also have signs reversed, from a negative association on table 16 to a positive one considering only LLM countries.

Therefore, the equation for Low Labour Market (LLM) cluster are presented in the following updated equation 3.2:

$$\mathrm{UR}_{\mathrm{LLMCt}} = (-1.232363) + (-0.065957)\mathrm{GDP}_{\mathrm{t-1}} + (-0.01379)\mathrm{INF}_{\mathrm{t-1}} + (0.010759)\mathrm{SER}_{\mathrm{t-1}} + \mu_{\mathrm{it}} \text{ (Eq. 3.2)}$$

Hence, giving that the negative association between GDP and unemployment on LLM cluster is the similar with the one found on the general model, the interpretation remains essentially the same as inferenced before, but with updated values. A one-unit increase in the lagged gross domestic product (GDP), would be associated with a decrease (due to the negative association) of 0.065957 units in the unemployment rate. These two variables fluctuate in opposite directions if one increases the other tends to decrease.

Inflation and self-employment variables on Low Labour Market countries are in contrast when comparing with the general model. Meaning that for LLM nations an increase of one unit in inflation numbers is associated with a decrease of 0.01379 in unemployment rates while a rise of one unit in self-employment associates with a reduction on unemployment of 0.010759.

A potential explanation for the INF relationship with UR could be that an eventual growing on inflation indexes may be associated with an increased prices of goods and services which could lead companies to reduce their expenses in other factors such as their hiring processes, consequently, increasing unemployment levels. There are some studies in favour of the maintaining of unemployment rates above what could be perceived as natural or an ideal rate, to eventually have a stricter policy for restoring price stabilities and better deal with scenarios of higher inflation (Friedman; 1977; Orphanides, 2002).

About the self-employment rates coefficient interpretation for Low Labour Market countries, values of a positive association between SER and UR suggest a possible counterintuitive result, as a rise in self-employment is commonly perceived as a positive thing that leads to reduced unemployment. A possible explanation for the obtained variation of these variables in LLM countries is that when in periods of economic slumps, companies diminish their hiring processes and individuals fall into gig-economy opportunities, which does not necessarily absorb formally the unemployed. This illustrates the not rarely ambiguous relationship between these factors as mentioned on Thurik et al. (2008).

Following table 20 presents the countries that are composing the LLM cluster for further assessments.

Table 20
Low Labour Market countries.
(Elaborated by the author).

CLUSTER	COUNTRIES
Low Labour Market	Algeria, Bosnia and Herzegovina, Croatia, Egypt, Eswatini, Gabon, Greece, India, Iran, Iraq, Italy, Jordan, Lebanon, Mauritania, Moldova, Montenegro, Morocco, Namibia, Nepal, North Macedonia, Senegal, Serbia, South Africa, Spain, Sudan, Tunisia.

Note: This is the final sample of countries after the cleaning process applied to remove the nations with absent data or misinformation. An update for the described on previously presented table 10.

Apart from India, one apparently common trend on these nations is that they are relatively small countries (both in land area and population). Although this is a diverse group with low and medium incomes, does not seem that any of the presumed most developed economies are part of this cluster. Furthermore, it seems that countries that faced some economic crisis in a recent time are here, as the case of Greece. Not to mention some of the relative poorer economies in African continent, as Sudan and Gabon, for example.

To have a more direct numerical analysis about countries on this and on the other clusters that will be later assessed, it was extracted from World Bank again, who comprises World Development Indicators, data regarding absolute percentage values for unemployment rates, filtering this information considering the working timespan of 10-years (2012-2021)[3]. We found a medium value of 6.00% for unemployment rates in a world-level. Worldwide numbers in this timespan have a lowest in 2019 5.54% and peaks in 2020, the Coronavirus-outbreak year, of nearly 6.90%.

This 6.00% worldwide unemployment rate will be referred to assess how countries in each cluster are far or approximate of this referential number. Low Labour Market countries described on table 20, apart from Senegal and Moldova, all have unemployment rates above the 6.00% threshold. Some countries, as Eswatini, North Macedonia and South Africa, have alarming numbers above 20%. We have some justification on the labelled low labour market cluster, as the employment-unemployment cycle for countries presented at table 20 seems to be slow.

Moving forward to the Medium Labour Market (MLM), this cluster also has a sample reduction after the cleaning process. Initially with 108 countries now the group has 93. A 15 country-drop that maintain this group with most nations considering the 154 total of countries we have been observing. The following table 21 shows the retrieved results regarding the MLM cluster after the usage of `cajorls` on R-Studio and a

[3] Data available at: https://data.worldbank.org/indicator/SL.UEM.TOTL.NE.ZS.

summary of values about the long-run association of variables considering the equation 3.3 presented next:

$$\Delta UR_{MLMC_t} = \alpha_1 + \sum_{\rho}^{i=1} \beta_{1_i} \Delta UR_{t-1} + \sum_{\rho}^{i=1} \beta_{1_i} \Delta GDP_{t-1} + \sum_{\rho}^{i=1} \beta_{1_i} \Delta INF_{t-1} + \sum_{\rho}^{i=1} \beta_{1_i} \Delta YU_{t-1} + \sum_{\rho}^{i=1} \beta_{1_i} \Delta SER_{t-1} + \sum_{\rho}^{i=1} \beta_{1_i} \Delta LP_{t-1} + \sum_{\rho}^{i=1} \beta_{1_i} \Delta SAV_{t-1} + \lambda_1 ECT_{t-1} + \mu_{it} \text{ (Eq. 3.3)}$$

Table 21
Long-run relationship of the VECM equation – Medium Labour Market cluster.
(Elaborated by the author).

Variables	Coefficients	Standard error	T-values
GDP (-1)	+0.17226	0.10977	1.569
INF (-1)	+0.087568	0.084662	1.034
YU (-1)	-0.007749	0.048711	-0.159
SER (-1)	-0.059206	0.049814	-1.189
LP (-1)	-0.0046036	0.0414628	-0.111
SAV (-1)	+0.007605	0.061786	0.123
CointEq1 (UR)	-0.155873	0.043629	-3.573
Constant		0.1151882	
R-Squared Adjusted		0.06918	
F-statistic		6.305	

Table 21 presents a scenario more similar with the modelling that considered all 154 countries. Adjusted R-squared here of 0.06918 is slightly above the value of overall (see table 17) modelling, suggesting that these six variables have a potential of explanation by almost 7% of the unemployment rates on the set of countries in the Medium Labour Market cluster. F-statistic though is lower here, 6.305 compared with 8.618 for the complete sample. Nonetheless, the model remains statistically reliable with all defined exploratory variables selected.

Values on CointEq1 presented in table 21 suggests a solid results and a movement of UR toward equilibrium on around 15% (0.155873) considering the velocity of adjustment. A relatively low standard error of 0.043629 and a t-value of -3.573 indicates significant values the unemployment rate variable. Maintained all variables for the VECM equation for Medium Labour Market, results shows that three variables (GDP, INF and SAV) have a positive association in the long run with unemployment rates while other three (YU, SER and LP) have a negative relationship.

As it is noticeable that MLM cluster has more similarities with the overall model than with the Low Labour Market cluster most of the comparison will be made with the general model and coefficients presented on table 17. There we had as well three positive and three negative associations, but not necessarily the same ones presented in the table 21. GDP have differing signs, negative on the general and positive for MLM and youth unemployment signs are also reversed, negative for the complete sample and positive for MLM; inflation, self-employment rates, labour productivity and savings have the same coefficient signs at tables 17 and 21.

All coefficients given and described on table 21, with no variable remotion, we may proceed for composing the equation for Medium Labour Market (MLM) cluster that is presented in the following updated equation 3.3.

$$UR_{MLMCt} = 0.1151882 + (0.17226)GDP_{t-1} + (0.087568)INF_{t-1} + (-0.007749)YU_{t-1} + (-0.059206)SER_{t-1} + (-0.0046036)LP_{t-1} + (0.007605)SAV_{t-1} + \mu_{it} \text{ (Eq. 3.3)}$$

Considering the countries on the Medium Labour Market have a positive association between GDP and unemployment, meaning that one unit of increase of these countries' gross domestic product is associated with a 0.17226 increase on the unemployment rates. Result may appear counterintuitive giving that is in opposition with the seminal premises on Okun's law (Okun, 1962; Lee, 2000) and most common accepted economic theories that when GDP rises, countries are more economically stable, their unemployment tends to decrease.

This form of association is not without unprecedent on the literature however, Ball & Mankiw (2002) for example, already argued that the association of economic growth (usually measured by GDP) does not necessarily implies in more jo occupancies. Bhaduri & Marglin (1990) also discussed how economic growth influence on the unemployment rate are not completely measurable. Therefore, results here have some basis to accept that does not necessarily a more developed economy, higher GDP, leads to a lowering on unemployment levels (Lavoie & Stockhammer, 2013).

Moving to the other variable with conflicting result between MLM cluster and the overall sample, youth unemployment for the Medium Labour Market group have a negative association with unemployment by -0.007749, suggesting that if this variable have a one unit increase this would implicate on a decreasing happening as well on the overall unemployment rates in the MLM countries by 0.007749. We have presented several studies (e.g., Mroz & Savage, 2006; Gough, Langevang & Owusu, 2013; Barford, Coutts & Sahai, 2021) that argues these variables fluctuate in the same direction as obtained for MLM cluster and seems more intuitive this occurring than the opposite.

Other four exploratory variables in MLM cluster have the same presumed association found on the complete dataset of countries. Self-employment rates and labour productivity coefficients are negatively related with unemployment rates by the values of -0.059206 and -0.0046036, respectively, whereas inflation and savings presented a positive coefficient association by +0.087568 and +0.007605. Interpretation about it is if SER and LP increases the unemployment levels would decrease while when INF and SAV are rising the unemployment rates accompanies this rise as well.

Comparing with the general model, the one with all 154 countries, the association found on self-employment rates, labour productivity, inflation, and savings are equal despite some numerical variation. More importantly, the presumed level of influence by these variables into the unemployment rates composition follows a similar interpretation as we already presented when analysing the overall VECM model. This implicates that most of the inferred and literature basis we present there could be replicated for the Medium Labour Market analysis.

After the assessment of the six exploratory variables, it is the moment to present which countries are in the labelled Medium Labour Market cluster. Report of the countries is important to later assess the absolute values of unemployment rates for this group of nations, similar as proceeded with the Low Labour Market set of countries. Following table 22 presents the 93 countries included on MLM group.

Table 22
Medium Labour Market countries.
(Elaborated by the author).

CLUSTER	COUNTRIES
Medium Labour Market	Albania, Armenia, Australia, Austria, Azerbaijan, Bangladesh, Barbados, Belarus, Belgium, Belize, Bhutan, Botswana, Brazil, Brunei Darussalam, Bulgaria, Cabo Verde, Canada, Chad, Chile, Colombia, Costa Rica, Cote d'Ivoire, Cyprus, Czechia, Denmark, Djibouti, Dominican Republic, Ecuador, El Salvador, Equatorial Guinea, Estonia, Fiji, Finland, France, Gambia, Georgia, Germany, Guatemala, Guinea, Guinea-Bissau, Haiti, Honduras, Hong Kong SAR, Hungary, Ireland, Israel, Jamaica, Japan, Korea, Rep., Kyrgyz Republic, Latvia, Lesotho, Lithuania, Luxembourg, Madagascar, Malaysia, Maldives, Malta, Mauritius, Mexico, Mongolia, Netherlands, Nicaragua, Nigeria, Norway, Pakistan, Panama, Philippines, Poland, Portugal, Qatar, Romania, Russian Federation, Rwanda, Samoa, Saudi Arabia, Sierra Leone, Slovak Republic, Slovenia, Solomon Islands, Sri Lanka, Sweden, Tanzania, Togo, Tonga, Turkiye, Ukraine, United Kingdom, United States, Uruguay, Uzbekistan, Zambia, Zimbabwe.

Note: This is the final sample of MLM countries after the cleaning process applied to remove the nations with absent data or misinformation. An update for the described on previously presented table 10.

Medium Labour Market group has in it around of 60% of the complete sample of 154 counties being analysed. This is the most diversified set of countries including large countries of the world as Russia, United States and Brazil, for example, and some small lands, such as Fiji and Solomon Islands, for example. Economic levels of the countries on MLM cluster are also diversified, including giants as China and USA and underdeveloped economies from the African continent.

As we done for the Low Labour Market cluster and assuming as parameter the world-level of unemployment on 6.00%, most of the countries that composes Medium Labour Market have values nearby this number, some are slightly under 6%, going to 3 or 4 percent while others are a little above, about 7 and 8 percent. Although most of the 93 countries on this cluster are about this range of percentual indexes, some outliers are identified within the cluster, cases of Djibouti for example, having a median value of unemployment rates above 26% and on the opposite side Qatar, that has a median value of 0.16%.

Reasoning for Djibouti and Qatar cases may be because of land area and population for these countries. For example, Qatar has almost no unemployment but is a relatively small and less populated territory than many others in the sample, meaning that this 0.16% on proportion and considering these caveats may be as significant as countries having around 6% of unemployment rates. Further assessment on these cases and why unemployment rates for these countries are so low or so high is something that escapes from this chapter’s scope but could be a future research avenue path to be pursued.

Continuing restrained to our defined endeavour on the comprehension of labour markets analysis on the three created clusters of countries, it is the moment to assess the third group, the named High Labour Market (HLM) group. HLM initially had 39 countries and this number, as happened for the other two clusters, have been reduced, in HLM case from 39 to 35. Comparatively, this is the cluster that suffered less with information lost during the data cleaning process applied on the methodological proceedings.

Following table 23 presents coefficients and other results for the High Labour Market cluster after application of `cajorls` on R-Studio, which gives a summarization of values regarding the long-run association for the VECM equation 3.4. This mathematic expression is depicted after table 23.

Table 23
Long-run relationship of the VECM equation – High Labour Market cluster.
(Elaborated by the author).

Variables	Coefficients	Standard error	T-values
GDP (-1)	-0.22915	0.29331	-0.781
INF (-1)	+0.065772	0.045202	1.455
YU (-1)	-0.0015493	0.0753346	-0.021
SER (-1)	+0.10694	0.12841	0.833
LP (-1)	+0.089885	0.062933	1.428
SAV (-1)	-0.10774	0.15934	-0.676
CointEq1 (UR)	-0.15182	0.07816	-1.943
Constant		2.298844	
R-Squared Adjusted		0.04354	
F-statistic		2.219	

$$\Delta \mathrm{UR}_{\mathrm{HLMCt}} = \alpha_1 + \sum_{\rho}^{i=1} \beta_{1_i} \Delta \mathrm{UR}_{t-1} + \sum_{\rho}^{i=1} \beta_{1_i} \Delta \mathrm{GDP}_{t-1} + \sum_{\rho}^{i=1} \beta_{1_i} \Delta \mathrm{INF}_{t-1} + \sum_{\rho}^{i=1} \beta_{1_i} \Delta \mathrm{YU}_{t-1} + \sum_{\rho}^{i=1} \beta_{1_i} \Delta \mathrm{SER}_{t-1} + \sum_{\rho}^{i=1} \beta_{1_i} \Delta \mathrm{LP}_{t-1} + \sum_{\rho}^{i=1} \beta_{1_i} \Delta \mathrm{SAV}_{t-1} + \lambda_1 \mathrm{ECT}_{t-1} + \mu_{it} \text{ (Eq. 3.4)}$$

For High Labour Market, similar as happened in the general model and on MLM cluster, no variable must be removed. HLM cluster seems to have a scenario alike with the Medium Labour Market and the modelling that considered the complete sample of 154 countries; however, values on table 23 still indicates some particularities for this cluster. Adjusted R-squared is the lowest on comparing with the other three modelling, the 0.04354 value suggests that the selected variables could explain around 4% of the dependent variable. F-statistic despite being relatively low in comparison still imply on a statistically reliable model.

Values of CointEq1 presented on table 23 suggest an ideally presumed value for the variable unemployment rate return to equilibrium on a long run analysis. We have found a relatively low adjustment necessary, by -0.15182, similar with the one obtained for MLM cluster. Considering that no variables were removed, results on table 23 are the basis for compose the High Labour Market VECM equation. Coefficients presented on the table shows that from the six variables we have three (INF, SER and LP) with a positive association, and other three (GDP, YU and SAV) with a negative association with unemployment.

Once again, the signs of coefficients association are mixed. Comparing with the Medium Labour Market cluster, only two associations are equal there and here on the High Labour Market group, the variable inflation being positive and youth unemployment being negative (see table 21 for MLM coefficients). While in comparison with the model for the complete sample of 154 countries, inflation is positive while GDP is negative on both (see table 17 general VECM coefficients) and the other four variables are with reversed signs. Furthermore, two coefficients are equal comparing HLM and Low Labour Market, the GDP negative and the SER positive associations (LLM coefficients are presented on table 19).

Overall, as it was possible to presume giving that different countries compose the three groups, all clusters have their own association sign that applies better for these cluster in isolation. However, some pattern seems to emerge such as the GDP usually being negatively associated with unemployment (on the general model, on LLM and on HLM clusters) while inflation (on general model, on MLM and on HLM) is commonly positively associated alongside unemployment rates. These empirical associations are complying with seminal premises of Okun's Law (Okun, 1962) and Phillips Curve (Phillips, 1958) that we briefly discussed on this chapter's literature review.

Restraining for the High Labour Market analysis and the described-on table 23, following equation 3.4 present the HLM Vector Error Correction Model mathematic expression:

$$\begin{aligned} \mathrm{UR}_{\mathrm{HLMCt}} = {} & 2.298844 + (-0.22915)\mathrm{GDP}_{t-1} + (0.065772)\mathrm{INF}_{t-1} \\ & + (-0.0015493)\mathrm{YU}_{t-1} + (0.10694)\mathrm{SER}_{t-1} + (0.089885)\mathrm{LP}_{t-1} \\ & + (-0.10774)\mathrm{SAV}_{t-1} + \mu_{it} \text{ (Eq. 3.4)} \end{aligned}$$

GDP have a negative coefficient of 0.22915 related with unemployment rates. A unit increase in GDP would be related to a decrease by 0.22915 in the long-term. Broadly speaking, past economic growth is associated with a reduction in structural or long-term unemployment. A negative relationship suggests that if economic activities (proxied by GDP) increases, this may generate favourable conditions for job-creation which, therefore, implicates on a reducing in overall unemployment rates. This association seems persistent from the Okun's (1962) proposition to nowadays (e.g., Sahnoun & Abdennadher, 2019; Miyamoto & Suphaphiphat, 2021).

Inflation positive coefficient presumes that a one-unit rise on inflation rates influence on the growth of unemployment levels by 0.065772 units, meaning that both variables fluctuate on a same direction if one is increasing the other increases as well. This is aligned with the general model and with the Medium Labour Market cluster, suggesting that in most cases inflation can negatively affect labour market, leading to higher unemployment rates in the long-term. This association complies which was initially suggested on the theory of Phillips Curve (Phillips, 1958).

Youth unemployment for High Labour Market cluster has a negative association with unemployment by 0.0015493. This negative relationship is conflicting with general VECM application, where the association is positive, but is aligned with the Medium Labour Market. Observing the three clusters and the complete sample of 154 countries, we have reason to believe that overall unemployment is in fact affected by young people that does not have a job, despite the sign of the association. Therefore, interested parties, especially the policymakers, must dedicate some attention to their youth population to cope early and effectively with this problem at the initial stages of individuals' labour-life (Bell & Blanchflower, 2010).

Moving on for self-employment rates coefficient, on the High Labour Market cluster it is obtained a positive association of this variable with unemployment rates by 0.10694, implicating that both variables change into the same direction. A positive association as the one in HLM here also happened on the final set of variables for the LLM only, other assessments on MLM and general model suggested a negative association.

The sign of association between SER and unemployment may vary due to economic context, the labour market policies, sectoral specific investments, and productivity levels that have as well an important role in overall unemployment dynamics. Relationship of self-employment and unemployment is probably one of the most complex we are analysing (Blanchflower & Oswald, 1998; Thurik et al., 2008).

On labour productivity variable, in the High Labour Market cluster we found a positive association with unemployment by 0.089885. A positive association appearing is a first on our results presentation, considering that on the general model and on Medium Labour Market cluster the relationship was found as negative. Productivity and unemployment relationship is usually volatile and highly nonlinear, which justifies the positive association found here to not be an abnormality. Labour productivity is largely influenced by policy interventions that could shift structural parameters according to their interests (Acemoglu, 1995; Murtin & Robin, 2018) even this implicates a growing unemployment.

Finally, last variable to be checked on the High Labour Market cluster, savings have a coefficient of -0.10774 suggesting a negative association of this variable with unemployment rates for countries on the HLM cluster. If savings rates rise a one unit, unemployment on countries participating on HLM cluster will decrease by 0.10774 units on the long run. Savings negative association found on the High Labour Market group is conflicting with the sign found on the general VECM application and on the Medium Labour Market cluster.

Some inferences may justify the reasoning on increase on savings reducing unemployment. For example, a higher saving rate can lead to greater capital accumulation potentially improving economic growth, which may in turn create more job opportunities and eventually reduce long-term unemployment rates. We have more theoretical background in favour of an increase on savings be connected to an elevation in unemployment (e.g., Mody, Ohnsorge & Sandri, 2012), however the volatility between these variables is inherent and, in the end, country- level context will be the main driver for this and any other association.

At last, after all six variables coefficients presented and discussed, we present in the following table 24 the 35 countries included on the High Labour Market cluster. Once again, as done on Low and Medium Labour Market groups, it will be compared the cluster and the world-level rate of unemployment and how the countries within HLM cluster are positioned in comparison with the median value of unemployment in the 2012-2021 timespan.

Table 24
High Labour Market countries.
(Elaborated by the author).

CLUSTER	COUNTRIES
High Labour Market	Angola, Bahamas, Bahrain, Benin, Bolivia, Burkina Faso, Burundi, Cambodia, Cameroon, China, Congo, Ethiopia, Ghana, Iceland, Indonesia, Kazakhstan, Kenya, Kuwait, Macao SAR, Mali, Mozambique, Myanmar, New Zealand, Niger, Oman, Paraguay, Peru, Singapore, Switzerland, Thailand, Timor-Leste, Uganda, United Arab Emirates, Vanuatu, Vietnam.

Note: This is the final sample of HLM countries after the cleaning process applied to remove the nations with absent data or misinformation. An update for the described on previously presented table 10.

Returning for the worldwide median value for unemployment being of 6%, High Labour Market countries are positioned very close this referential number. Considering the 35 countries presented in table 24, we found two outliers in the group (Angola and Bahamas with median values of 12.35 and 9.5 percentages, respectively) and all the other nations are under the 6% limit, going in a high of around 5% (Paraguay) and a lowest of 0.3% (Cambodia).

Important to acknowledge that a country being in the High Labour Market does not necessarily implies on superior economies or richest countries, these features does not are mandatory to have a somewhat heated labour market. This was not our presumption on the labels of clusters as well the dynamics of labour market have way more complexities than this. Labour markets are in their essence mutable and any improved understand about them is essential to successful implementations of policies that aims to cope with unemployment (Schmid & Gazier, 2002).

On doing all the above analyses and discussions about the three clusters our intention was to better see how countries align between themselves and how the most common regional divisions and groupings of nations, as the geographic or economic cohorts, are not necessarily effective when assessing employment, unemployment, and labour market complexities. Therefore, the clustering proceedings and the supranational cross-country analyses presented so far are to be useful in obtaining a more concrete portrait of the dynamics and potential discrepancies that exists on labour markets that emerged on our working sample of 154 countries.

Vector Error Correction Model was applied on the previous topic on four different levels. One application checks the sample of 154 countries, which is the complete dataset of nations with reliability and availability of information to proceed with the analysis intended for this chapter, considering the cleaning proceedings used on the methodological section. Other three VECM application were made considering the three clusters that emerged from the 154 country-data, the Low Labour, Medium Labour, and High Labour Market clusters were assessed and discussed before moving to the causality tests on the next section.

3.4.6. Short-run Granger causality test.

Granger causality test is a statistical examination for check whether one time series is indeed useful in forecasting another related series. Causality checks the association between cause and effect in which Cause is an independent variable whereas Effect is a dependent variable (Granger, 1969). Furthermore, exists two forms of relationship between variables assumed as Cause and Effect, the unidirectional causality, and the bidirectional causality, in which Cause exercise influence on the Effect variable, and equally, Effect variable influences the Cause variable (Shabbir, Kousar & Alam, 2021).

Considering the above and the already depicted premises described on the methodological section of this chapter, we have used the `grangertest` function that is available on the `lmtest` package on R-Studio software. On using this we may retrieve an F test statistic and a corresponding p-value, if this p-value is less than a certain significance level (i.e., $\alpha = .05$), then it is possible to reject the null hypothesis of the Granger's test.

Granger causality test is applied here to determine the direction of relationships among the selected variables on the sample of 154 countries being analysed so far. It will be checked Granger-causes of every one of the six exploratory variables we have analysed here (gross domestic product, inflation, youth unemployment, self-employment rates, labour productivity and savings) into the dependent variable (unemployment rates) as well the reverse influence of UR on the exploratory factors.

Following table 25 summarizes the results. Lag length used will be 1, as this is the default on the function `grangertest`.

Table 25
Results of the Granger causality test.
(Elaborated by the author).

Null hypothesis	F-statistics	P-values	Decision
GDP does not Granger cause UR	7.7824	0.005341 **	Bidirectional causality
UR does not Granger cause GDP	4.8949	0.02708 *	
INF does not Granger cause UR	0.8289	0.3627	No causality
UR does not Granger cause INF	1.025	0.3115	
YU does not Granger cause UR	0.5667	0.4517	No causality
UR does not Granger cause YU	2.0553	0.1519	
SER does not Granger cause UR	1.488	0.2227	No causality
UR does not Granger cause SER	0.7154	0.3978	
LP does not Granger cause UR	0.5875	0.4435	No causality
UR does not Granger cause LP	0.0082	0.9278	
SAV does not Granger cause UR	0.0859	0.7694	No causality
UR does not Granger cause SAV	0.5477	0.4594	

Notes: Decision rule: Rejection of null hypothesis (H0) if the p-value is less than 0.05.
*Signifies the rejection of a null hypothesis at the 5% level of significance; **Signifies the rejection of a null hypothesis at the 10% level of significance.

Table 25 shows that there is one bidirectional causality and five no causalities running between the proposed variables on the short run. These results suggests that most of the variables when observed in a short-term analysis does not significantly affect the unemployment rates, likewise, unemployment fluctuations do not affect inflation (INF), youth unemployment (YU), self-employment rates (SER), labour productivity (LP), and savings (SAV).

Meanwhile, the interaction between GDP and unemployment rates on a cause-effect relationship is bidirectional. Statistically the influence of GDP on UR is significant even in a 10% level of significance, implicating a strong power of the gross domestic product on unemployment levels. This association corroborates some of the other results presented on this study that discusses economic development, here proxied by GDP, and on how this influence labour market.

We believe that this short-run analysis is important to complement everything presented at subtopic 3.4.5, where the Vector Error Correction Modelling assess the relationship and signs of influence observing the long run analysis; however, as we may perceive on the results presented at table 25 what happens in short term does not necessarily endures on a larger horizon as the opposite is also correct. Remains to be checked if the proposed models are stable and statistically reliable, something that will be assessed in the following subtopic.

3.4.7. Diagnostic tests.

To check the robustness of the Vector Error Correction Modelling proposed and tested on this chapter, some statistical diagnostics will be applied to confirm model's consistency. These tests are described and discussed on the following paragraphs.

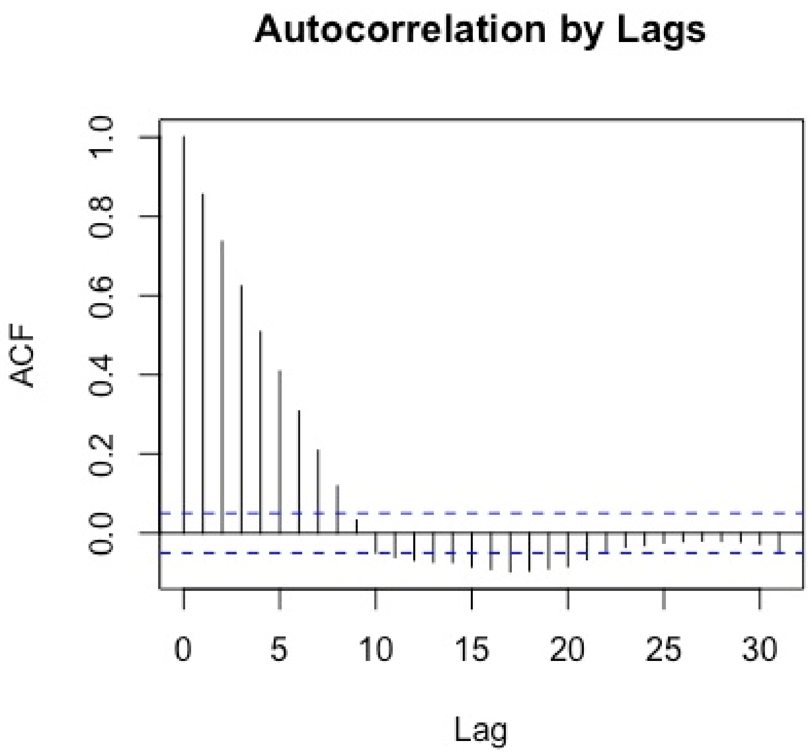
3.4.7.1. Autocorrelation.

Autocorrelation measures the degree of similarity between a time series and a lagged version of this same time series over successive time intervals. Also referenced as serial correlation or lagged correlation, this type of analysis is useful for the measurement of the relationship between a variable's current index and its historical values. Autocorrelation test for this chapter's purpose were applied using R-Studio in two manners, one statistical and one visual representation.

First, regarding the visual presentation we are using the autocorrelation plot (ACF). About this, when interpreting an ACF plot the x-axis corresponds to the different lags of the residuals (i.e., lag-1, lag-2, etc.) while the y-axis represents the correlation of each lag. The dashed blue line represents the significance level. First bar on the plot shows the correlation of a residual with itself and therefore is always 1 (lag 0). In the absence of autocorrelation, the subsequent vertical bars would drop quickly to almost zero, to between or closing to the dashed blue lines. (Venables & Ripley, 2002).

Considering the above, following figure 21 was created using `acf` function, a feature on the `stats` R-Studio package useful to plot and estimates autocorrelation. Figure 21 shows that the ACF plot of residuals of our applied VECM are not autocorrelated. After the lag-0 correlation, the following correlations (represented by bars on the plot) begins to drop into 0.0 and the majority of bar is between the limits of the significance level around the dashed blue lines. Therefore, the residual of this model complies with the premise of no autocorrelation.

Figure 21
Plotted autocorrelation.
(Elaborated by the author).



Complementing the visual presentation, we applied statistical measurement of autocorrelation among residuals by performing the Durbin-Watson tests. These tests check for first-order autocorrelation (lag-1) under the null hypothesis (H0) that first-order autocorrelation does not exist and alternative hypothesis assuming the existence of this form of autocorrelation (Durbin & Watson, 1971). Our results are presented-on table 26.

Table 26
Autocorrelation via Durbin-Watson tests.
(Elaborated by the author).

	dwtest		
D-W test statistic	1.4134		
p-value	0.000		
	durbinWatsonTest		
Lags	*Autocorrelation*	*D-W test statistic*	*p-value*
1	0.85444346	0.2908493	0.000
2	0.73645613	0.5248771	0.000
3	0.62332305	0.7511297	0.000
4	0.50891637	0.9793644	0.000
5	0.40884663	1.1789948	0.000
6	0.30769915	1.3809228	0.000
7	0.20857691	1.5789089	0.000
8	0.11885305	1.7580135	0.000
9	0.03278178	1.9298663	0.258

Durbin-Watson test statistic usually present values between 0 and 4 where values from 0-2 means positive autocorrelation, 2 means no autocorrelation and from 2 up to 4 assumes a negative autocorrelation. Test was applied by us on two-ways: Using the function `dwtest`, available at the package `lmtest` on R-Studio, and function `durbinWatsonTest`, from package `car`, to test autocorrelation imputing more than 1 lag and first-order only analysis. Results for both applications are presented on the above table 26.

Considering the `dwtest` output, we have a rejection of the null hypothesis of first-order autocorrelation, giving the 1.4134 D-W test statistic obtained at a p-value of 0.000. D-W test statistic is approximate of the most usually accepted 1.5 threshold. As far as `durbinWatsonTest`, we may notice that lags are increased the D-W statistic also increases. Optimal value obtained is by 1.76 when lag is 8 and the p-value remains significant on 0.000.

Considering the lag equal to 8, autocorrelation is dropping from 85% chance of occurrence to 11%, which suggest a reasonable acceptance. Lag ideal being equal to 8 corroborates what could be observed on figure 21, when the autocorrelation by lags is significant, on the dashed blue lines, around 8 as well. Hence, considering the presentation on figure 21 and the results presented on table 26 it is possible to accept an ideal lag length up to 8, the moment when serial correlation issues are to be settled bearing in mind the characteristics of our data under analyses.

3.4.7.2. Heteroscedasticity test.

Statistically speaking the homoscedasticity assumes that the variance of errors must be constant across all the modelled observations, which is the most recommended result. If residuals become more spread-out at higher values, this is a strong suggestion that heteroscedasticity is present on the linear regression. If homoscedasticity is not happening (therefore heteroscedasticity occurs), this could lead to biased coefficient estimates, incorrect standard errors, and unreliable interpretations (Shabbir, Kousar & Alam, 2021).

As we applied for autocorrelation, once again it will be used R-Studio in two analyses of the homoscedasticity premise, one statistical and one visual representation. First check for heteroscedasticity is by using a residual vs. fitted-plot analysis. This illustrates the distribution of the residuals of a regression model and its fitted values, trough the usage of a basic R-function for plotting graphics, the `ggplot2`.

As mentioned before, the most recommended pattern to be visualized is a somewhat homogenous distribution of observations, suggesting therefore, a homoscedastic behaviour. Eventual dispersions may be present on the plot, but ideally most observation residuals and fitted values should be aligned around a same line of distribution.

Following figure 22 presents a graphic illustration considered the modelled variables indexes and observations (countries numbers for each variable) under analysis in this study on a residual vs. fitted-plot representation. On this analysis we are considering the complete final sample of 154 countries.

Figure 22
Plotted heteroscedasticity test.
(Elaborated by the author).

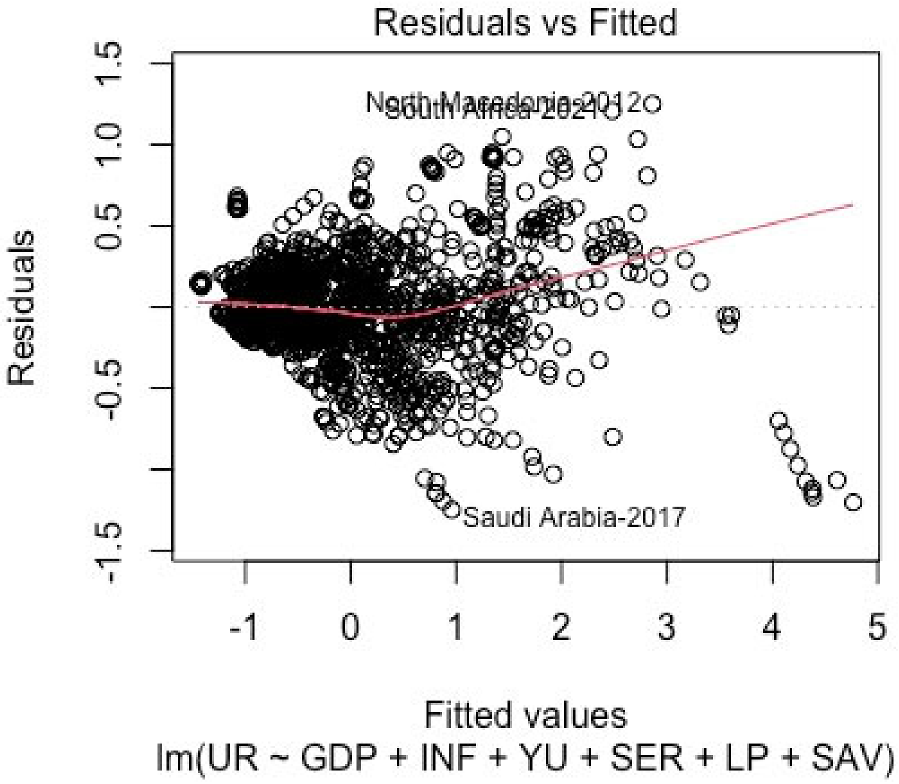


Figure 22 on showing the residual vs. fitted-plot relationship seems to suggest that indeed the variance in the residuals increases as the fitted values are increasing. This is an indication that heteroscedasticity is likely a problem in the regression model and the standard errors from the model summary may be unreliable. The modelling, with all variables, potentially violates the homoscedasticity postulation, at least considering the visual analysis. However, to complement and attest what is inferred by figure 22, we applied the statistical Breusch-Pagan test.

This test is used to determine whether heteroscedasticity is present in a regression model following null hypothesis (H0) that homoscedasticity is present (residuals are distributed with equal variance). If the p-value of the test is less than some significance level (i.e., $\alpha = .05$) then it is possible to reject the null hypothesis and conclude that heteroscedasticity is present in the model (Breusch and Pagan, 1979).

Table 27
Heteroscedasticity test via Breusch-Pagan test.
(Elaborated by the author).

bptest	
B-P test statistic	379.93
df	6
p-value	< 0.000

According to the results depicted on table 27 regarding the Breusch-Pagan test, we have found a p-value < 0.05, therefore, we may reject the null hypothesis of the test, suggesting that heteroscedasticity could be occurring on this study modelled linear regression. This would not be an ideal result, we could be following some spurious or unreliable coefficients and, therefore, interpretations about our VECM modelling. Considering results on figure 22 and table 27, we decide to proceed with a third approach, calculating for robust standard errors.

This verification was made using the `coeftest` function, also available at the R-package `lmtest` and the `vcovHC` function, which is available from the `sandwich` package. Our results have unveiled that considering the six exploratory variables of our model, it exists statistically significant heteroscedasticity problems on INF, YU, SER and SAV while the variable LP is marginally homoscedastic (0.05) and GDP complies well with the homoscedastic distribution, with values above 0.05. These results give to us some margin to consider that our modelling is not completely spurious.

Furthermore, our main objective for this chapter (Develop a supranational cross-country analysis to identify main determinants composing unemployment rates) is more concerned on influential level of exploratory than the discussion of estimations or to conduct specific hypothesis tests. Also, we have some grounding on the literature arguing in favour that heteroscedasticity is not necessarily a stoppage for any analysis (Morgan, 1976; Long & Ervin, 2000; Rubio-Aparicio et al., 2020; Đalić and Terzić, 2021), not to mention that our subject and unit of analysis is a social phenomenon, a field that naturally demands some alleviation on strictly statistical metrics that other fields cannot comply.

3.4.7.3. Multicollinearity test.

Multicollinearity exists whenever an independent variable is highly associated with at least one of its companion independent variables when in a multiple regression equation. Multicollinearity may be an issue since it could undermine statistical significance of an independent variable. To address eventual multicollinearity problems, initial steps include the proposition or evaluation of a model through different diagnostic techniques for detecting and dealing with multicollinearity in regression models (Farrar & Glauber, 1967; Mansfield and Helms, 1982).

Again, it will be ensured two checking about multicollinearity, a visual representation using a correlation matrix by using the basic R-function `corrplot`, on the package that have the same name, and later the plot-visualization will be complemented with VIF statistical measurement. Following figure 23 presents the correlation matrix for visual analysis.

Some parameters to interpret figure 23: Multicollinearity is natural when observing a variable in isolation, it is found a value of 1.00 when the analysed variable is paired with itself. This is relatable with what is assumed as a perfect collinearity (or multicollinearity), when there is an exact 1:1 correspondence between two independent variables in a model, this correspondence could be presumed as perfect either in a correlation of +1.0 or -1.0. Therefore, as a rule of thumb, correlation between two (exploratory) variables being below -0.9 or above +0.9 would be an alert to a potential multicollinearity problem.

Figure 23
Plotted multicollinearity test.
(Elaborated by the author).

Considering figure 23 it appears that this could be a problem only on unemployment rates (UR) and youth unemployment (YU), rest of variables on a first analysis passes the test. To further analyse these variables and the others on the model, next step is to apply the Variance Inflation Factor (VIF) test. This test measurement will be using the function `vif`, available at the R-package `car`. Table 28 presents these tests results.

Table 28
Multicollinearity test: VIF results.
(Elaborated by the author).

Variables	VIF
GDP	1.072240
INF	1.026140
YU	1.513946
SER	1.322847
LP	1.466506
SAV	1.250011

Note: VIF ideally should have a value under 5 and any value above 10 would indicate severe correlation.

Results on table 28 presents that in a practical analysis multicollinearity is not present on the model proposed for this study. Variance Inflation Factor (VIF) values for all exploratory variables are under the most well accepted and with relative endurance on the literature (Menard, 1995; Kim, 2019) rule of thumb by 5 and marginally around 1.5 (on youth unemployment), which is the stricter measurement limit on multicollinearities test. Although we acknowledge that rules of thumb may be misleading if considered as unquestionably (O'Brien, 2007).

To summarize our diagnostic tests: Our VECM model is well-fitted about autocorrelation of the variables; we have found a p-value < 0.00 up to an ideal lag-length of 8, suggesting that the variables being analysed are not highly correlated. Also, multicollinearity test also indicates the suitability of the model considering that all VIF values for the six selected exploratory variables are under 5 as presented in table 28. Heteroscedasticity however is present on the sample, which is not an ideal indication but considering the relatively large sample of countries, 154, and therefore the natural variability of these nations, is a somewhat expected outcome considering the supranational cross-country analysis we performed in this chapter.

Hence, it is reasonable to presume that the modelling proposed, applied, and analysed on this chapter is fit. It is presented a solid model that may be useful to achieve a better oriented and assertive understand of the unemployment, which and how the six predictors to unemployment rates could explain this phenomenon. It is our belief that the findings and discussion presented on this section complies well with the proposition to apply a supranational cross-country analysis of unemployment rates in the 2012 to 2021 timespan.

This macroanalysis, considering that we have observed data from at most 154 countries, is indeed useful to assess the present of unemployment rates. However, we have as well shed some light on the necessity to restrain even more than the clustering proceeding here applied to have factual propositions and a manageable foresight about what could happen in the future of unemployment and to have in hand more assertive avenues to cope with this perennial phenomenon. About this is dedicated the next chapter, but before that, we conclude this one.

3.5. Conclusions.

To conclude this chapter, we must return to the defined main objective of it: Develop a supranational cross-country analysis to identify main determinants composing unemployment rates. We believe that indeed, using secondary time-series and panel data available on World Bank repository of information, Databank, it was conceived a solid representation about the determinants that influences the composition of unemployment rates in a dataset of 154 countries, enabling the response as well for the research question that guided this chapter (What are the main determinants that composes the unemployment rates?).

We have used a Vector Error Correction Model (VECM) technique on data covering 2012Q1 (first quarter of the year) to 2021Q4 (last quarter of the year), allowing to to assess influential power of six selected (exploratory) variables on unemployment rates (dependent variable) composition. These selected variables were primarily extracted from the previously applied bibliometric analysis, developed on chapter 2 of this thesis. Some variables besides the bibliometric inputs were added and removed considering the availability of data and consistent country-information to be analysed.

Results for the general VECM application, that analysed the complete sample composed by 154 countries, shows that gross domestic product (GDP), self-employment rates (SER) and labour productivity (LP) have a negative association, which implies that when these are increasing this causes a reduction in unemployment rates (UR). On the other hand, inflation (INF), youth unemployment (YU) and savings-rates (SAV) shows a positive association with UR, meaning that if one of these exploratory variables increase, unemployment will increase as well.

Still considering the complete sample of countries we have analysed, we have unveiled that the leading influential variable on unemployment, considering the VECM model we have used, is inflation. INF have the highest coefficient (0.118283) among the six exploratory determinants checked. Youth unemployment and gross domestic products follows INF having 0.0598575 and 0.031595 coefficients, respectively. Diagnostic tests were applied directed to our complete sample of countries and we have found that the variables we select passes the autocorrelation and multicollinearity tests while fails to reject the null hypothesis of presumed homoscedasticity on the dataset.

However, heteroscedasticity happening on a sample composed by 154 countries with different income levels, population number, economic momentum and distinctions in many other factors that permeates labour market are not necessarily an unexpected result and does not invalidate our results inferring presented throughout this chapter. Findings on this study are both in line and contraposing two of the most seminal econometric propositions. Regarding the GDP relationship with unemployment, results are in line with Okun's law, which presupposes an inversely proportional association between changes in the growth rate, we proxy this as GDP, and fluctuations of unemployment rates (Okun, 1962; Hjazeen, Seraj & Ozdeser, 2021).

Relationship we have found about inflation and unemployment is conflicting the original proposition of the Phillips curve, that presumes these two variables would have an inverse relationship, meaning that higher inflation rates would be associated with lowering on levels of unemployment and vice versa (Phillips, 1958; Alisa, 2015). Our results, on the 154 countries sample, showed a positive association among INF and UR, suggesting that these two fluctuates on a same direction, therefore if one is increasing the other would also be growing. It is important to notice that Phillips curve concept was proposed on a very particular context, oriented by macroeconomic policies in the 20th century, but original assumptions was called into question by the stagflation phenomenon on the 1970s and after (Bruno & Sachs, 1985; Ormerod, Rosewell & Phelps, 2013).

On conclusion, it is possible to say that this present research allowed some solid empirical support on regard unemployment rates composition while alerting to the development of better oriented active labour market policies. These policies, according to our results especially oriented to good levels of economic growth, therefore GDP, as this would lead to lower levels on unemployment and a stable labour market. Furthermore, as much as possible, an effective curbing to inflation rates, giving that this is the leading variable in our sample and if inflationary pressures are increasing more people would have to cope with a loss of their jobs.

Adding to the overall analysis that is focused on the sample of 154 countries, deepened assessments were applied on subgroups within this dataset, aiming to check if eventual disagreements on the coefficient associations would occur when the country-data is being observed in groups composed by countries that have more similarities between themselves on regard their labour market. A clustering technique was used and then three cluster of countries where created, the ones named Low, Medium and High Labour market clusters.

Low Labour Market cluster, which has 26 countries, we had three of the six variables analysed as statistically significant, p-values ≤ 0.00. GDP and inflation show a negative association (-0.065957 and -0.01739, respectively), while self-employment rates have a positive one (+0.010759). Showing that countries within this cluster are highly dependent of their economic growth, proxied by GDP, to cope with unemployment rates, hence, complying with Okun's law and with the Phillips curve presumed inverse relationship between INF and UR.

Medium Labour Market, the largest cluster with 93 countries, have all the six proposed variables statistically significant for the VECM application. Three variables were positively associated with unemployment: GDP, INF and SAV whereas YU, SER and LP showed a negative relationship with unemployment rates composition. Countries within this cluster have gross domestic product as the most influential factor on their unemployment composition (+0.17226).

High Labour Market has 35 countries and as happens for the overall modelling and on the Medium Labour cluster, all variables selected are statistically significant, having their p-values ≤ 0.00. On this cluster as well, the higher coefficient is on GDP, a negative association by 0.22915. Youth unemployment and savings also have a negative association while the other three variables (inflation, self-employment rates and labour productivity) show a positive influence on unemployment rates.

Throughout our application of the Vector Error Correction Modelling as the checking for the statistical significance of variables as well their coefficient association it was possible to go deep on the supranational cross-country assessment. This analysis was in-depth considering the cross-clustering observations to assess how countries that have some level of similarities between themselves may have distinction when comparing results of the overall sample and modelling with the 154 nations' sample.

Each interested agent involved with countries included on one of the three suggested clusters have then a more empirical oriented path to follow and discuss how and the level of urgency some factor should be addressed when the intention is to better cope with unemployment and promote a more stable labour market. The belief and the aim when designing this study is to assert what are the main factors that exert influence on unemployment levels while, consequently, offer some insights on what should be confronted first and foremost to better deal with such a complex phenomenon.

In summarization, all findings presented and discussed on this chapter lead to a better comprehension about some factors and their influence on unemployment rates. On the scope of this thesis the study is successful on building upon elements from the early developed bibliometric analysis and from there evolve to a more empirical comprehensions about unemployment that first was centred in the literature of this theme essentially.

However, some limitations inherent of this chapter conception must be declared. Relying exclusively on one source of data may be harmful to our results. Although World Bank and Databank are a robust source of information, some of these could be directly extracted from country entities, and potentially the data from this origin would be more precise and up to date. Nonetheless, the task to collect this level of information in a significant number of countries it would not be simple. Also, different variables from the six we assessed here could be verified and proxied in different manners as well, expanding the scope of exploratory variables that relates with unemployment.

The Vector Error Correction Model, the technique applied here, although feasible and with consistent results on its output could be performed in other manners. Not to mention the tool used, the R-Studio software and packages, that by itself have other approaches of the same method, other statistical tools could be used and suggest different results than the ones we have obtained here. It is our recommendation for future studies that may be developed from this one the usage of other techniques to check and compare results with the propose by us. For example, a similar analysis as the one presented here could be performed by simplistic regression equation and other autoregressive models.

Also, the proposed on this chapter suggested for us the mandatory necessity to develop Some proceeding with a restrained focus, being this scope a country in specific or even in a region, to enable more assertive results for these countries reality. It was not the intention here, as the aim for the chapter is broader, but could be a direction to be followed on other studies that intend to be more directional to a particular context. Nonetheless, limitations and biases considered, our belief is that this chapter offered a reasonable presentation on how unemployment rates are composed in a general perspective. Remains to be seen how these rates could evolve on the future, something to be assessed on the following chapter.

4. COMBINING FORECASTS TO FORESIGHT THE FUTURE: A SCENARIO-BASED PROPOSITION FOR THE UNEMPLOYMENT ON BRICS COUNTRIES

4.1. Introduction.

As happens with many other macroeconomic indicators as gross domestic product, economic production of goods, monetary policies and inflation, labour statistics are a particular matter of interest in different areas. It is important for investors, if interested in anticipating economic trends and how extract profit on it, private companies aiming to maintain their employees or to attract a better workforce to their groups and to governments intending to better allocate or reallocate their labour force.

Unemployment rates in specific is an important topic of attention particularly to policymakers as well other relevant economic stakeholders that have some level of interest on labour market context, giving that this may be one of the most accurate approximation that illustrates the relationship between a country business cycle, their monetary policies directed to employment and unemployment and the active population involvement within this market (Blanchard & Leigh, 2013; Chakraborty et al., 2020).

Therefore, to have in hand reliable data about unemployment may be pivotal to decision and policymaking, enabling the recognition of eventual economic and social disfunctions to, consequently, be able to design better plans to cope with these problems. Unemployment rates as some other macroeconomic indexes are usually lagged informed, officially published after some revision which implies in a delayed data. Hence, up to date information and, when or if it is possible, real-time data may offer more effectiveness to deal with labour market problems (Fondeur & Karamé, 2013; Simionescu, 2020).

Real-time as well historical information on a subject is particularly important if the intention is on anticipate tendencies into a foreseeable future. This form of studies, that have a forecasting approach of macroeconomic variables are not necessarily recent and started to spread in the middle of the 1990s (e.g., Agnew, 1985; Aggarwal, Mohanty & Song, 1995; Marcellino, 2004; Chakraborty et al., 2020) and since them many time-series modelling have been applied to shed light and eventually to predict macroeconomic variables behaviour on the future being among these variables, the unemployment.

Initiatives to forecast, foresight and on foresee future scenarios are useful to individuals, organizations, and governments on each of their own efforts on being able to achieve desired outcomes and be better prepared with assertiveness for future-oriented decisions (Hogarth & Makridakis, 1981; Yoshida, Wrigth & Spers, 2013). Bearing this premise in sight, this chapter will be developed upon some of the basis and results previously presented on this thesis to propose here a final piece on the contributions aimed by this study.

As happened on the chapters developed before, the focus on this one will once again be on unemployment, which is the key macroeconomic factors of interest for this research as well on many practical levels, from individuals' interest to public management. The aim here is to assert unemployment rates evolution on a foreseeable future, designing potential future scenarios for labour market and unemployment in specific. Before goes in depth about objectives on the chapter we believe it is important to differentiate premises regarding forecast and foresight, because although both have a similar end on to have an insight or a look into a possible future, some distinction exists on conceptual level that led to this shared purpose.

Foresight has been used to refer about some level of promptness to confront long-term issues, especially on the behalf of governmental interests (Miles, 2010) whereas forecast is most commonly understanding as a methodological approach and not a final intent per se. Therefore, foresight is the concept hereafter adopted because the intention is to combining forecast and scenarios techniques to not only have an insight into the potential future in isolation but using these methods to establish and propose information that could offer better readiness to cope with unemployment dysfunctions that may come in short-medium time (Cuhls, 2003).

A deepening on the forecast and foresight ideas will be presented later but it is important to state that on this study will be used future studies techniques, more prominently the forecast, and scenario building as well, to achieve a foresight proposition, that will include but not be limited to the forecast methodology and its results. On regard the endeavour on foresight unemployment, this presents as relevant because it may offer useful inputs to help policymakers, economists, and other stakeholders to have a better idea of what the future labour market may present itself.

Therefore, it would be of both managerial and social interest to produce solid estimation for unemployment rates and when having availability of information to help, anticipate these future unemployment rates with reliability and up to date information. In doing so, it will be possible to design more adequate policymaking while warning and foreseeing signals of economic labour crises that may occur in the future.

Unemployment indeed is a complex and multifaceted as not many other social and economic phenomenon could be. On individuals' perspective, may bring harmful financial effects for the ones unemployed and for those on their household, having even potential psychological consequences (Eisenberg & Lazarsfeld, 1938). In a macroeconomic angle, a high unemployment rate may implicate on negative effects on overall economy, diminishing taxes income, increasing governmental spending to secure those without a job and so on (Gogas, Papadimitriou & Sofianos, 2022).

On this presented scenario, this study is proposed as an effort to achieve some level of anticipation on how unemployment rates can evolve, considering past and present rates to have an insight on a foreseeable future. As mentioned on Blustein et al. (2020) governments and organizations oriented by a forward-looking approach more likely than not will infer beneficial results if conjecturing sooner about on how to deal with the unknown and potentially enduring consequences of the crisis started by COVID-19.

Considering all the above introduced and the relevance of labour market and unemployment within this context, the expected objective on this essay is to forecast future unemployment rates and from these the proposition of potential future scenarios for the foresight of unemployment rates in the BRICS countries group. On creating these scenarios, the expectation is to have an answer to the following research question: What are the foreseeable tendencies for unemployment in the BRICS countries?

Methodological approach to answer this question and met the proposed objective is to ensemble statistical forecast proceedings (Exponential Smoothing Technique – ETS, Seasonal and Trend decomposition using Loess – STL, and Autoregressive Integrated Moving Average – ARIMA, and Artificial Neural Networks - ANN) to foresee future values of unemployment rates for each of the five BRICS countries (Brazil, Russia, India, China, and South Africa) in a 10-year into the future timespan. Statistical results in hand, next step is the proposition of scenario forecasting, suggesting how may be the potential future of labour market in these countries.

On mixing statistical and judgmental approaches to forecast, the belief is to offer a solid and replicable process with reliable outputs on how unemployment rates may evolve on the BRICS countries and, therefore, produce significant insights to the interested parties on this matter to be better prepared to cope with unemployment in the future and mitigate socio-economic effects that may be harmful.

Innovations that emerge from this approach are: (a) this study uses data on a supranational level, considering the BRICS group of countries, while most of the related studies are made with data from isolate nations and, not rarely, mostly European (b) the intended foresight proposition are based on a solid statistical proceeding that is the ground to be complemented from judgmental forecast, mixing both methods with a premise of complementarity among them. To the best of our knowledge, up to this moment, no study attempt to foresight in this approach, particularly focusing on the BRICS context.

Relevance of predictive studies of unemployment is growing on interest since the 1970s and the stagflation problem (Brunner, Cukierman, & Meltzer, 1980; Bruno & Sachs, 1985; Gogas, Papadimitriou & Sofianos, 2022). However, literature on this topic is largely focused on the forecast on United States unemployment rates (D'Amuri & Marcucci, 2010; Xu et al., 2013) and other developed economies as Germany, Belgium, and United Kingdom (Askitas & Zimmermann, 2009; Bughin, 2011; McLaren & Shanbhogue, 2011). Thus, the possibility to contribute with the theme bringing to discussion a still unexplored context as the BRICS presents as an opportunity by itself.

The decision to focus on BRICS countries is due to their economic relevance as a group and that despite being from different regions, from different levels of economic development, they still share some similarities between themselves. Brazil, Russia, India, China, and South Africa are expected to play a significant role on the world's economy, having the potential to eventually evolve from the "emergent market" status to high economic potencies (Betul, 2015) if they can manage well their resources, which includes a latent labour market to be better understand and assigned.

Rest of this chapter is organized as it follows: Second section presents a literature review of forecast and foresight and their usage on unemployment-related studies, more specifically on some applications on the BRICS countries individually. Section three presents the methodological proceedings to be used through this study: Exponential Smoothing Technique – ETS, Seasonal and Trend decomposition using Loess – STL, Autoregressive Integrated Moving Average – ARIMA, Artificial Neural Networks – ANN, and Scenario forecasting. Section four presents the empirical results of these methods on isolation and combined, their statistical results and the building of future scenarios projected; and to finalize, section five concludes the chapter.

4.2. Literature review.

On this section the emphasis will be directed into topics that offers theoretical background on some premises for this chapter development. First topic will assert future studies, forecast and foresight premises. Second topic will present studies that develop a similar approach as this one, and the third and final topic is focused on unemployment within BRICS countries as they are the final context of analyses on our study.

4.2.1. Future studies, forecast and foresight.

Future studies are a relatively new strain both on research efforts and on practical applications. Nonetheless, Slaughter (1998) argues that although the future may seem an abstract concept, it inevitably exists and will come, sooner or later. Yet, future cannot be predicted. Hence, any effort to better understand, explore it, mapping it and if it is possible to create some desirable outcome, any study for the future should be encouraged and oriented (Slaughter, 1998).

On the promotion of future studies research field, Wendell Bell who is one of the pioneers on this field, argues in favour of the idea of prospective thinking. Thinking ahead, not aiming to predict but to prospect, enables the discovering, examination, evaluation and, therefore, a better-informed situation on future outcomes (Bell, 1997). The purpose for those Bell called "futurists" should be on have the knowledge about what can or could happen (the possible), what is more likely to occur (the probable) and what ought to be (the preferable) (Bell, 1997).

Again, according to Bell (2001), any prospective scenario, from the worst to the better as possible, must be substantiated on a solid and appropriate previous knowledge, be this a theoretical or empirical. That premise aligns with the declared purposes of this study, as we intend from empirical results infer about theoretical futures – scenarios – that may come to fruition regarding unemployment rates. However, prospective thinking as a general idea although not always named this way, has a long history, going since educated guess about some topic up to the always growing number of data available to be assessed statistically or through machine learning techniques. One way or another, people still want to know what is coming ahead despite the undeniability that the future is yet not full predictable.

Future studies as a research track on specific, since the initial studies up to more recent approaches remains as a considerable fragmented field (Kuosa, 2011; Fergnani, 2018). Academics and practitioners contribute on the area coming from different backgrounds, ranging from climate change to cybersecurity, the future state of economies and many other subjects (Fergnani, 2018). Given this diversified source of contributors, aiming to establish some rigour to the future studies field, forecast and foresight techniques and their respective tools emerged and started to spread.

Later, this section will present some studies that used these methods but before that, the idea is to close the wider scope of future studies, using as main reference the three major phasis on the field, following the Kuosa (2011) study on this theme. Kuosa (2011) identified three on future studies research field, from 1940s years up until the first decade of the 21st century. First phase was largely centred on planning programs through quantitative methods and a positivism orientation. Second phase was from 1960s to 1970s, marked by the spread of international research that started to be developed beyond the USA frontiers. Third phase goes from 1980s up to 2011, when Kuosa's study was published, up to nowadays. Key features of this phase include a more consolidate topic despite the fragments still exists.

On all these phasis and in any future study probably the main difficulty is to bring something that could be assumed as "futurology" or conjecture to practical and effective propositions. Maybe the first attempt to give a more oriented process to generate reliable outputs from future studies came in the form of a derived concept from the future studies, the idea of a foresight approach. Richard Slaughter about foresight core premise said that is to "*Expanding awareness and understanding through futures scanning and clarification of emerging situations*" (Slaughter, 1990, p.82).

Over the years up to nowadays these premises defended by Slaughter (1990) only grow on urgency. Studies developed after Slaughter's work have been advancing the conceptualisation of foresight, as a terminology, as well the effective applicability of this concept. Major, Asch & Cordey-Hayes (2001) for example, defend foresight as a core competence, considering that to foresight is to be opened and better prepared to the future using wisely tools at disposal for develop visions of future avenues and then, hopefully, be able to choose the best path to proceed (Slaughter, 1995; Major, Asch & Cordey-Hayes, 2001).

Foresight is in many cases associated with other fields; examples are particularly present on the strategy literature. Major, Asch & Cordey-Hayes (2001) claims in favour of assume foresight as a strategic capability that, if well applied, could enhance an organization's ability to navigate uncertainty and proactively shape its own future. Results obtained by the authors suggest that the concept of foresight and core competencies of pathfinding, a common feature on strategy literature, are describing in their essence essentially a similar concept (Major, Asch & Cordey-Hayes, 2001).

This managerial approach and the benefits that foresight could offer to companies has only evolved both in management practice and on academic literature. Even a track for this foresight/organization's relationship has emerged, the corporate foresight (CF). CF gained traction as a fruitful avenue that could enable managers to have a long-range understanding about their business and to act more proactively despite being surrounded by the uncertainty about the future (Rohrbeck, Battistella & Huizingh, 2015).

Corporate foresight origins go back to the 1950s years, driven by two main theoretical roots: The French 'prospective' school, leaded by Gaston Berger, and the 'foresight' school guided primarily by the USA initiatives and the RAND Corporation efforts, when one of the most prominent and still acknowledged technique on corporate foresight was initially developed, the Delphi technique (Rohrbeck, Battistella & Huinzigh, 2015).

On the 1960s and 1970s corporate foresight started do spread on repertoire beyond the Delphi technique, more particularly the usage of scenario's building and forecasting. Scenario's technique attracted interest more especially due the successful application of it by Royal Dutch/Shell company, that using a scenario programme were able to increase their awareness regarding the oil crisis that would come and affect their business interests (Jefferson, 2012).

Following the success of Royal Dutch/Shell programme, other companies tried by themselves achieve a manner to anticipate what the future could present for them, increasing the professionalization of methods and processes by corporations that started to have a focus on foresight. CF then begins to evolve for continuous scanning and interpretative approaches for potential future scenarios, which later lead to techniques more rigorous, such as technology roadmaps (Battistella & De Toni, 2015).

Corporate foresight may indeed be presumed to be of great relevance to strategy and management stakeholders considering the innate potential of integration of dynamic capabilities and future-oriented outcomes desired by organizations. These potential positive outcomes may emerge from a better designed firm's learning process, creativity, innovation, and performance, leading to competitive advantage that could go beyond what traditional strategy and management scholarship have reached in the past (Fergnani, 2022).

Marinković et al. (2022) doing a systematic literature review on the theme found some interesting insights about theoretical approaches on corporate foresight. Both qualitative and quantitative studies were conducted being the first type the most used. Most of qualitative research was based on case studies, supported by interviews and Delphi technique (Marinković et al., 2022). Quantitative studies were predominantly guided throughout the treatment of surveys and basic statistical methods, more descriptive results, and some usage of regressions (Marinković et al., 2022).

Despite being qualitative or quantitative approach pursued, it should foresight be completely apart from forecast or maybe it presents in this duality an opportunity to improve foresight field through a more robust forecast proceeding that may offer more assertively results than simple descriptive statistics. The premise assumed for this study is that forecast remains important as was in the initial future studies and in fact can be improved, mixing it with more judgmental approaches as is the intention to proceed here.

Following this, we intend to answer the call made on Marinković et al. (2022), that claims for empirical research to be concerned with combining different tools instead the usage of a method in isolation. The belief is that quantitative methods, as the forecast proceedings, can profit from more abstractive approaches, as the scenario building and forecasting, the procedure that we intend to perform here. Nonetheless, before that, we return to some forecast specificities as was done for the foresight until now.

Early concepts for forecasting are designed to propose estimations for short, medium, or long-term futures, when considering a specific topic of interest. To forecast then, these research questions must be well-known and defined in advance to assess those with more focus and directly (Cuhls, 2003). Still by Cuhls (2003), forecasting application may range from possible futures envisioned by normative statements, potential scenarios, positive or negative, and single linear trend extrapolations, that must be based on actual data.

Forecast usage have a particular range on studies related to economics and public policies analysis, which is the track that this study intends to contribute as well. For this context, Klein (1984) argues that forecast proceedings, if solid and well-used, would be an important tool to identify how much stimulus or restraint should be implemented by economic measures and policies, and, therefore, how to better orient these practical on effective manners of fiscal, monetary or trade policies.

De Gooijer & Hyndman (2006) study presents an extensive review of forecast methods that had been applied on future studies efforts. Authors illustrates how this track of research has a consistent growth in both theoretical propositions and empirical usage. Technological advances, greater computational power available, improved, tested, and robust statistical models all align may implicate on mature approaches to forecast, while still unknown improvements on methods are also emerging (De Gooijer & Hyndman, 2006).

Petropoulos et al. (2022) suggest the literature on the theme is solidified, evolving from an initial educated guess to the “wisdom of crowds” that contributes with future studies continuously. Theoretical premises of forecasting is particularly rigid on the idea that current and past data must be used to make any predictions about the future (Petropoulos et al., 2022). When patterns and historical values are well identified, implement them into the process of predicting future values may generate more accurate outcomes although the exact prediction of future values are still not achievable.

On this ongoing study we share this premise presented on Petropoulos et al., (2022), assuming as orientation from past and current data and steps to be followed to reach an objective. Therefore, we rely on data available of unemployment rates on Brazil, Russia, China, India, and South Africa, to from these data apply forecast techniques that gives as output a solid point to describe and propose scenarios for the future of unemployment in these countries. Before further discussion on the methodologies to be applied, next topic covers the overall usage of future studies on the unemployment theme.

4.2.2. Future studies on unemployment theme.

It is not an understatement to say that work activities, as in perform a job or a craft in exchange of some form of remuneration, as well the work relationships have been changing rapidly and continuously over the years. Wave of change emerge from the Industrial Revolutions and the implement of new methods of production, from factory-based production passing through steam-powered machines to electricity-related machines up the expansion of information and computational technologies (Castells, 1996; Barbosa et al., 2022).

Nowadays relationships on this field may be more complex as the crescendo of new technologies as artificial intelligence methods, robust robotic advances, biotechnology, and nanotechnology are pushing the limits of work in a broader perspective. All these factors and changes over the years, from the first tool developed to help on a craft previously made solely by an individual from a robot effectively substituting this same person, it is possible to associate with socioeconomic effects and transformations that already has and still will new scenarios for the future of work (Barbosa et al., 2022).

Job relations on the future and, therefore, labour market, employment, and unemployment, tends to remain a challenging topic of interest due to its influential range of effects, from a country economic development to individuals' social well-being. Directly on the scope of this study, attention is toward unemployment, as this is one of several perennial issues in society. Within this context, future studies that intent to forecast, or foresight, have been appearing as a challenging task for economists, forecasters, and policymakers, especially during economic downturns (Barnichon et al., 2012, Boga, 2020, Jo, Kim & Lee, 2023).

Bearing this in mind, being able to appropriately anticipate changes on unemployment levels is important for fiscal and monetary policies regulation, both on the proposition and the implement of the most adequate policies (Gogas, Papadimitriou & Sofianos, 2022). About future studies on unemployment theme, these have been developed since the 1970s, interested on the relationship about high inflation, high unemployment, and slow economic growth (the stagflation), going even before to the seminal study and proposition of the traditional Phillips Curve concept (Phillips, 1958; Bruno & Sachs, 1985; Gogas, Papadimitriou & Sofianos, 2021).

Efforts on anticipate the behaviour of pivotal macroeconomic variables has only growing, from the 1970s up to now. About unemployment rates in particular these studies started to spread in the middle 1990s, when many time series models have been analysed extensively (Chakraborty et al., 2020). Estimation of unemployment levels on the future has been a topic of interest most regularly on developed countries, offering for their realities some reliable projections (Lasso & Snijders, 2016).

Many statistical and econometric techniques have been employed on future studies for unemployment. Mostly of them using forecasting methods that presents some benefits and shortcomings as well, from high sensitivity to modelling specification, to extreme requirements with respect to the data (Cook & Smalter Hall, 2017). Attempts to foresight unemployment, in the past 20 years of literature rely heavily on Auto Regressive Integrated Moving Average (ARIMA) and Generalized Autoregressive Conditional Heteroskedasticity (GARCH) forecast models (Gogas, Papadimitriou & Sofianos, 2022).

Dobre & Alexandru (2008) used ARIMA method to forecast short term values of unemployment rates for Romania, Floros (2005) used GARCH to assess unemployment in the United Kingdom, founding a close relationship between the forecasting theory premises and labour market conditions. Kurita (2010) used a variation of ARIMA, the ARFIMA method, to successfully forecast unemployment rates for Japan.

ARIMA methodology appears to be the most recurrent tool used on future studies dedicated to observing unemployment rates. However, as far as our review achieved, most of these applications are centred on developed countries and economies. Beyond the above-mentioned examples, Funke (1992) used ARIMA for the German unemployment rates and Vicente, López-Mendez & Pérez (2015) uses another ARIMA variation, the ARIMAX, to forecast Spanish unemployment.

Despite ARIMA models consistently being used and providing insightful results for developed economies within the context of future studies about unemployment, other methods are not disregarded, and some have been growing on influence. Ongoing advances in modern statistics and machine learning have offered researchers with some robust nonlinear forecasting tools beyond classical forecast methods. Artificial neural networks (ANN), decision trees, random forests, deep learning processes, support vector machines (SVM) and others (Katris 2020; Chakraborty et al., 2020) are some examples of the expansive options available to perform foresights.

For our study we have the premise that no method is completely superior to other. As any method offers their contributions our intention is to analyse unemployment data from the past and present to forecast from these through a combination of methods to later produce solid-based judgmental proceedings as we offer some scenario forecasting for the unemployment rates on BRICS countries.

Caperna et al. (2022) developed a study with a similar approach with the one intended to be performed here. Authors proposed a data-driven procedure using machine learning techniques aligned with other methods, in their case, internet-based queries, more specifically the Google search data, to better understand unemployment phenomenon. Using a cross-country setting during the pandemic scenario authors found that those web searches are well suited to understand labour market unexpected shocks (Caperna et al., 2022).

Gupta, Pierdzioch & Salisu, 2022 performed a foresight intending to anticipate changes in the United Kingdom unemployment rate using data from 1859 to 2020, more specifically information about unemployment and oil prices, founding a significant predictive power of the second to forecast UK joblessness levels. Other macroeconomic variables as well are forecast both on classical methods and machine learning techniques.

Being through the usage of the most classical forecasting techniques, as ARIMA, or by the recently adopted machine learning processes, as ANN, most labour-related and unemployment foresight studies are focused on European context (Gogas, Papadimitriou & Sofianos, 2022; Gupta, Pierdzioch & Salisu, 2022; Celbiş, 2023), which gives for our study some novelty on the proposition to proceed with a less scrutinized BRICS countries scenario. This move for outside the mainstream of research is possible because there is data available nowadays for almost every country in the world. Smit (2018) argues in favour of this matter, considering that considering the availability of big data information, researchers should be encouraged to explore other contexts.

Nonetheless, studies from other economies that are not the ones we intend to assess but were constructed with similar purposes and proceedings as the one we aim to produce here are referential for this one. The examples are many, as the Constant & Zimmermann (2009) on Germany, McLaren & Shanbhogue, 2011 on UK, the ones on United States (Choi & Varian, 2012; D'Amuri & Marcucci, 2017; Maas, 2020), Canada (Dilmaghani, 2019), France (Fondeur & Karamé, 2013), Spain (Vicente, López-Mendez & Pérez (2015), Turkey (Chadwick & Sengül, 2015), Italy (Naccarato et al., 2018) and so on.

However, there are on the literature some studies directed to forecast and foresight unemployment on the BRICS countries. Chinese labour market is for example analysed on Su (2014), Brazil have been observed on Lasso & Snijders (2016) and other examples will be further presented on the next topic of this literature review. Yet, a pattern on to proceed with analyses on countries individually is consistent and, as far as our knowledge found, no study has been dedicated on the BRICS as a group which is one of the attempts to be made here.

As source of data information regarding Brazil, Russia, India, China, and South Africa unemployment rates, it will be used the official reports aggregated by World Bank in their repository, the Databank. Most of these data are compiled by International Labour Organization by their own and official country's entities. As Databank comprises these sources and make available all in a same interface, there is the place we decided to extract information. Further detailed explanations on the data extraction process and the methodologies to be used will be presented later.

Before that and to close the topics covered on this literature review, next we present an overview specifically about BRICS countries and their specificities regarding to unemployment. Intention is to shed some light on some characteristics of these countries in the nowadays economic and political scenario to justify the option to dedicate those nations as a relevant subject of analysis. Special attention on the following is on labour market and unemployment in Brazil, Russia, India, China, and South Africa.

4.2.3. BRICS and unemployment.

BRICS had an unconventional start as a group of countries. The year was 2001 and an economist from the Goldman Sachs proposed the term BRIC, to put a name on some of the, at the time, world's fastest growing emerging economies. Namely, Brazil, Russia, India, and China (Betul, 2015; Suchodolski & Demeulemeester, 2018). That was on a Jim O'Neill report, envisioning these countries as a latent potential to become, in group, a bigger force in the global economy (O'Neill, 2001). This indeed comes to some fruition given that, nowadays, China is solidified as the world's second largest economy, being even the first one in some metrics (World Bank, 2023).

O'Neill ideas about BRIC as group started to spread and attract attention both in media outlets and literature academic research. Popularity about the potential of the countries would have if working together increase in interest that eventually leadership of each country embraced the idea, being pioneered by the Russian side on the interests of President Vladimir Putin. Putin and Russia suggested in 2006 that indeed the four countries should meet and coordinate themselves at a supranational level (de'Robertis, 2016; Suchodolski & Demeulemeester, 2018).

First official BRIC summit took place in 2009 on Yekaterinburg (Russia), where an official statement was released outlining the foundations of a shared perspective between Brazil, Russia, India, and China, aiming to collectively become a new driving force on economic advances and global governance, more specially on what they present as an effort for a global South cooperation (Suchodolski & Demeulemeester, 2018; Duggan & Azalia, 2020).

Working in line with the premise of the group, South Africa officially joined the BRIC in the 2011 summit, which happened in April of that year in Sanya (China). BRIC then becomes BRICS, in a coordinate effort by the five countries to promote new areas of cooperation within the members as well non-BRICS countries while representing a voice to be heard on international agendas such as the development of a clean energy agenda, for example (Smith, 2011; Betul, 2015; Suchodolski & Demeulemeester, 2018; Duggan & Azalia, 2020).

While writing this study, is being formalized a new expansion on the group that would add six more countries as members beginning 2024. Iran, Saudi Arabia, Egypt, Argentina, the United Arab Emirates and Ethiopia are expected to join the current five BRICS members. Expansion was announced at the 15th summit of BRICS that happens in Johannesburg (South Africa). This admission of new members happens as an increased effort by BRICS on reshape global world order aiming to provide a counterweight to the United States and European supremacy (Borger, 2023).

Effective results of BRICS expansion, for better or worse, remains to be known as it is still on early stages. However, the acknowledgment of a growth within the group may be twofold. At one side could be an indicative of expanded relevance by these countries in a global perspective whereas may as well suggest that the original collective (Brazil, Russia, India, China, and South Africa) do not have enough power by themselves, even working together, appealing to new incomes as potentially stronger representation.

On the scope of this research every analysis will be focused on the five original members on the BRICS. Before to proceed to their macroeconomic assessments, particularly about unemployment within these nations, it is important to recognize that even being participants of a collective group, these countries have significant differences on their political systems, economic structures, historical backgrounds, and many other features (Suchodolski & Demeulemeester, 2018).

Despite these natural dissimilarities, Suchodolski & Demeulemeester (2018) argues that BRICS have in common a perception of underrepresentation in important matters despite their increasing economic strength. Hence, BRICS may be directed to a union because of their dissatisfaction with current global governance framework, focused on European potencies and the United States. BRICS presents itself as an alternative force for themselves as well for other emergent economies (Suchodolski & Demeulemeester, 2018; Borger, 2023).

As a collective and historical perception, Brazil, Russia, India, China, and South Africa have a well-suited definition as emergent economies, although nowadays some (particularly Russia and China) could be assumed as leading economies. Nonetheless, BRICS importance on worldwide economics should not be underestimated (Baloch & Wang, 2019). For example, GDP on BRICS increased from 416.4 billion in 1990 to 18,188.4 billion in 2018, which represents 22% of global GDP in that year (Wang & Zhang, 2020).

Collectively, Brazil, Russia, India, China, and South Africa economic growth has been overall consistent from early 2000s years when these nations present progress that surpasses many other more developed countries (Radulescu, Panait & Voica, 2014; Yao, Li & Li, 2023). By the end of 2022 BRICS countries had around 31.5% of worldwide GDP and this numbers were projected to go up to 50% by 2030, specially led by Chinese growth, that has passed United States since 2015 if comparing two economies in purchasing parity terms (PPP) (Devonshire-Ellis, 2023).

Not all is necessarily good, however. If BRICS economic advances over the past years continues to be persistent or higher, it is more than likely that these nations still would have to confront some considerable effects on their paths to better development. Political and social instability, widespread corruption, and a considerable dysfunctionality of institutions are just a few of potential problems that may also grow due to an enduring social inequality that exists on BRICS countries (Baloch & Wang, 2019).

On this not equally distributed economic growth and advances, potentially one of the problems that may not only emerge but in fact endures in higher levels, is the labour market friction and unemployment statuses. About this, on doing this literature review we do not find too many studies that assess the unemployment phenomenon within the BRICS group. Proceeding a basic search on Scopus repository using keywords "*unemployment*" and "*BRICS*" filtering for available on journals, in final stage of publication, only 21 documents were found (query conducted on August 28, 2023).

Among these 21 studies collected, most of them uses unemployment as a proxy for other subjects of analyses, especially on the stream of financial studies (e.g., Umar & Sun, 2016; Grima & Caruana, 2017; Syed & Tripathi, 2019). Some of them however are somewhat aligned with our study. Lalthapersad-Pillay (2014) observed labour market on BRICS using as focus the gender influences on employment and unemployment situation. Results shows that most of BRICS countries had a skewness on unemployment levels regarding gender with a remarkable presence of women working in services and an only a small portion in industrial sector (Lalthapersad-Pillay, 2014).

There are two research on this Scopus sample that relates with some topics discussed previously on this thesis. Hashimi et al. (2021) proposes an evaluation of the relationship between unemployment rates and economic growth, assessing the Okun's law (1962) in BRICS from 1991 to 2008. Tahir & Burki (2023) analysed evidence of entrepreneurship and economic growth, suggesting that high youth unemployment rates in these countries could be mitigated if more support for entrepreneurial activities started to be promoted.

About future studies, at the best of our knowledge no previous research assess unemployment within BRICS collectively, as a group. However, when searching for Brazil, Russia, India, China, and South Africa individually, there appears on the literature a few reports dedicated for the analysis of labour market conditions in these nations. Observing the countries separately, the usage of forecast techniques is particularly prominent, which substantiates the study design we intend to perform on our study.

In Brazilian context, Lila, Meira & Oliveira (2022) proposed a methodological approach named robust reconciliation, aiming to improve classical methods applied on forecast studies. Authors found encouraging results in favour of their proposed modelling while providing some insights for policymakers and administrators on their decision-making processes, academics interested in sharpening the quality of their own models and business managers to cope better with unemployment levels in their area (Lila, Meira & Oliveira, 2022).

In Russia, most of the research seems to relate to migration flows within the country and how labour market changes following those movements. Chereshnev & Vasilyeva (2013) and Vasilyeva & Tarasyev (2014) proposes a modelling for predict migration flows using as potential predictors the wage differentials, distances between populations of the regions in Russia and unemployment levels. Vasilyeva (2017) develops upon these models with a forecast, anticipating that by 2030 could exist a structural imbalance of supply and demand in Russian labour market, leading to lower GDP and labour productivity.

About India, Pattanaik & Nayak (2013) discusses the enduring labour problems in the country examining trends of employment to forecast the employment–output over the years. Findings by the authors suggests some policies proposition that could be more effective to assess employment and unemployment of Indian labour market. Using the ARIMA method, it is anticipated that in Indian future may occur a substantial dependence on agriculture for employment, which would damage on other sectors leading to overall negative productivity on outputs and less job-opportunities (Pattanaik & Nayak, 2013).

More recently, studies in Indian scenario are dedicated to understanding potential effects of COVID-19 pandemic on unemployment rates levels, as India was one of the world's most affected countries. Agrahari et al. (2021) for example, combined Indian data from various sources for aggregate monthly information about unemployment. It was used by the authors a combination of methods, named Heuristic-based model using least squares approximation for their unemployment rate forecasting. Kumar et al. (2023) used four distinct modelling having a better performance of the Holt's and Winter's one and suggest that Indian authorities should handle fast and on anticipation to diminish the damaging effects following COVID-19.

Moving to studies about China, Jia & Meng (1997) alerts for governmental attention to maintain the, at their study's time, overall unemployment rates below 5%, forewarning the problem faced by the country regarding invisible unemployment, especially in agricultural sector. Wong, Chan & Chiang (2005) argued that future levels of unemployment rates, underemployment and real employment should be forecast to anticipate important information for labour-related planning and policy making directed to the best for Chinese interests. Their results offer important signals to direct training and employment policies propositions onto about labour market environment that could mitigate these already anticipate problems.

To close on the covering of five BRICS countries, Nkoane & Seeletse (2021), build a time series model to forecast unemployment on South Africa using robust estimators that enables the detection of outliers from their established time series data in a manner that these outliers does not need to be removed. Authors using a timespan from 1972 to 2014 applied an ARIMA modelling to forecast future of unemployment rates based on past data, similar as we intend to perform in this study.

Nkoane & Seeletse (2021) results for South Africa presented solid forecasting values for the future of unemployment that enables the suggesting of proceedings that may help South African government on efforts to strengthen employment by having the main shar of ownership on as many as possible companies in different sectors. To this, a review on stockholder policies in the country, considering part-ownership for ensuring that State have a say for implement the main governmental interests (Nkoane & Seeletse, 2021).

It is not the intention to discuss all studies developed on Brazil, Russia, India, China, and South Africa related to future studies and unemployment although we believe a solid portrait of the research developed on these countries considering the shared topic of interest that this study has with them. We offer some novelty to the literature at first on the combination of forecast proceedings to offer a better result considering more than one method isolated. Later, based on this ensembled forecast, proceed to a scenario forecasting considering the previous results obtained.

This methodological approach and the premise to not focus solely on one of the BRICS countries, expands our intended analyses and, therefore, the findings achieved. All of this brings to the table what we expect as contribution for this research field. Beyond this, on managerial and policymaking interests, we expect the emergence of useful insights, particularly for BRICS stakeholders, on how to be prepared to face future status of such a perennial macroeconomic variable of attention as unemployment. On to the next section, we will present the details about methodological proceedings and how we aim to use them to meet the objectives designed on this study proposition.

4.3. Methodological procedures and dataset.

This section is dedicated on further explanation about the intended methodology approach adopted to fulfil the main objective proposed for this chapter. Our intent is to proceed with a hybrid methodological approach and the hybridism happens as we rely on statistical proceedings as the foundational basis to produce initial forecasts and based on these results, we proceed to a judgemental scenario forecasting. Combining these two approaches, we intend to offer a foresight suggestion for 10-year in the future of unemployment rates within BRICS countries.

Subtopics on this section will briefly detail each of the methods. Namely, the methods there will be presented and later applied are: Seasonal and Trend decomposition using Loess – STL, Exponential Smoothing Technique – ETS, Autoregressive Integrated Moving Average – ARIMA, and Artificial Neural Networks – ANN. These are the statistical and quantitative part and to complement them, we have a dedicated topic to scenario building, and particularly the scenario forecasting procedure.

After we go through each of the methodologies that made our hybrid modelling proposition, other subtopic is dedicated to the collection and building of the working dataset that comprises data from Brazil, Russia, China, and South Africa. Once the methodologies specifics are presented and the data to be used is delimited, we conclude this section before proceeding to the presentation of results of topic.

4.3.1. Seasonal and Trend decomposition using Loess – STL.

STL main distinguishable feature from other methods resides on the possibility to perform a decomposition on a given time series of data. The paths for this decomposing process are presented on the STL acronym by itself, which is a reduction for "Seasonal and Trend decomposition using Loess". Method was originally proposed by Cleveland et al. (1990) and on this research we use the function `STL` available at the R-Studio package `fable` following the Hyndman & Athanasopoulos (2021) as referential.

Central idea on the Cleveland et al. (1990) STL proposition is to decompose a time series into three principal components: Seasonal, Trend, and Remainder (or residual). The seasonal component refers to the repeating patterns within the time series under analysis, being these patterns identified as daily, weekly, or yearly recurrent cycles. Trend identifies the direction of this same time series, observing if values are increasing, decreasing, or staying constant over time. Remainder component represents the residual component, including noise or irregularities in the data, on the time series after removing the seasonal and trend components (Cleveland et al., 1990).

For illustration purposes, we are presuming a theoretical time series, that we are labelling as UR_t for unemployment where t represents the timespan of unemployment rates obtained. STL method aims to decompose unemployment time series into S_t for seasonal, T_t for trend, and R_t for residuals. Therefore, values for unemployment can be expressed as: $UR_t = S_t + T_t + R_t$.

Loess part of the method is referring to a non-parametric smoothing technique that is applied in the same time series (assume UR_t once again) that fits a polynomial regression model to specific subsets of data, giving more weight to more recent points on the time series and less weight to the earliest. This locally weighted scatterplot smoothing (where Loess acronym came for) proceeding is applied on the effective decomposition for both trend and seasonal components (Cleveland, 1979).

According to Hyndman & Athanasopoulos (2021), two main attributes to be defined using STL are the trend-cycle and the seasonal windows. These can be defined discretionally by the researcher or identified automatically. `STL` function on the `fable` R-Studio package provides an automated decomposition using a seasonal window default on `season(window=13)` while the trend window is selected depending on the seasonal period obtained on the behaviour of the time series being analysed.

Both seasonal and trend parameters are helpful to control how rapidly the trend-cycle and seasonal components can change. Smaller values enable more rapid changes as the opposite is true for higher values. The automated process, in most cases, tends to a good balance between overfitting the seasonality while allowing it to slowly change over time. However, as is the case for any automated procedure, depending on the time series characteristics, some adjustments will have to be applied (Hyndman & Athanasopoulos, 2021).

As we intend to work with yearly unemployment rates data for the five BRICS countries individually and as a collective as well, the adjustments will be made when and if necessary, according to the data availability and its natural characteristics. Detailing about data specifics will be presented later in this same section and the countries outcomes aspects will be better understood on the section dedicated to results and discussion.

4.3.2. Exponential Smoothing Technique – ETS.

The exponential smoothing technique was originally proposed on the late 1950s early 1960s years and beyond (Brown, 1959; Winters, 1960; Holt, 2004), spreading on relevance by being used on many applications but with relevant influence on forecasting procedures (Hyndman & Athanasopoulos, 2021). Exponential smoothing methods usually involves the usage of weighted averages from past observed values where weights decaying exponentially as these observations get older, similarly as the other known proceeding the STL, also discussed on Hyndman & Athanasopoulos (2021) and applied on the previous topic.

More recent an observed value for unemployment rate is, for illustration, higher should be the associated weight for this observation. There is not much reasoning to believe that future unemployment rate in the year 2025, or further than this, would be more influenced by 2004 than 2024. This assumption may generate reliable forecasts rapidly in a wide range of time series and on this study, we use the function `forecast` also available at R-Studio package `fable` again following Hyndman & Athanasopoulos (2021).

There are three basic models on the ETS usage: First, Simple Exponential Smoothing (SES), that is appropriate for data without a clear trend or seasonality observed; The Holt's Method, or Double Exponential Smoothing, suitable for data that has an apparent trend but no seasonality; and third, the Triple Exponential Smoothing or

Holt-Winters Method, that is adequate for data that has both trend and seasonality among the observations.

As we are going to use the Hyndman & Athanasopoulos (2021) R-Studio package `fable`, one of the advantages of ETS statistical framework is that the smoothing parameter may be determined by the information criteria's (AIC, AICc and BIC) for ideal model selection automatized by the available function `ETS`. The function returns estimated values for α, β and γ, as the default smoothing parameters, and as higher these are obtained, more weight should be on recent observations which, consequently, would implicate on forecast more responsive to changes in medium-short term.

To summarize, the methodological proceeding of this method that we are using on this research follows the Hyndman & Athanasopoulos (2021) referential: First, an ideal model and its smoothing parameters must be identified, we perform this in an automatized manner via `ETS` function; having the best model defined, considering a times series of (unemployment rates in this case) forecast values may be suggested through the usage of `forecast` function. Both tools are built on the `fable` package.

Exponential smoothing technique in comparison with other advanced methods is relatively simple to understand and implement, as we illustrated that most can be performed automatically, whereas it does not fail on providing reasonably accurate forecasts, especially for short to medium-term predictions (Holt, 2004; Hyndman & Athanosopoulos, 2021). Therefore, ETS by itself could produce solid results and as we intent to combine this and other methods, the expectation is on have further better outputs.

4.3.3. Autoregressive Integrated Moving Average – ARIMA.

Considering that exponential smoothing models, as STL and ETS, are sustained on trend and seasonality in the data observed, ARIMA modelling shifts the attention to describe autocorrelation in a same data information. There is however some shared perception on the methods presented so far that may as well be a common premise to most of the forecasting attempts. We take this apart before the ARIMA specificities because we rely on these premises to move forward for further analyses.

Any predictive effort must be delimited for a specific time-horizon; historical values must be taken into consideration, with their respective weights associated; overly complex and elaborated models are not necessarily the best one, parsimonious is mandatory (Box & Jenkins, 1970; Newbold (1983). We follow these premises as some form of rule of thumb, not for ARIMA application but to STL, ETS and the later presented Artificial Neural Network (ANN) as well.

Onto the ARIMA modelling per se, hereafter we present parameters and the basic mathematical expression for the usage of this technique. In a very summarized definition, ARIMA is a linear time series model useful on the identification of linear tendencies when analysing stationary data within an observed time series. General model is denoted by $\mathrm{ARIMA(p, d, q)}$, where p and q represents the order of the autoregressive model (AR) and the moving average (MA), respectively, while d is the level of differencing which is used on the conversion of a nonstationary series into its stationarity form (Chakraborty et al., 2020).

Mathematically, the ARIMA model may be expressed as it follows:

$$y_t = \theta_0 + \sum_{i=1}^{p} \phi_i y_{t-1} + \varepsilon_t - \sum_{j=1}^{q} \phi_j \, \varepsilon_{t-j}$$

y_t stands for the factual value of the variable time series being analysed, in this study's case the unemployment rate on BRICS countries at a time t. ε is the random error at this period assumed as t and ϕ_I and ϕ are the coefficients of the model.

`ARIMA` also available in `fable` R-Studio package uses a variation of the Hyndman-Khandakar (2008) algorithm, combining unit root tests, minimisation of the AICc and maximum likelihood estimation, MLE, to obtain the most fitted ARIMA model. Beyond the automatized best fitted model that may be selected by `ARIMA`, parameters (p, d, q) must be assessed on autocorrelation function (ACF) and partial autocorrelation function (PACF) plots, and, to conclude, the model must be diagnostic checked for white noise behaviour of the residuals before to forecast (Hyndman & Athanosopoulos, 2021).

Although ARIMA modelling may be applied with the addition of seasonal terms, we intent to use the simple non-seasonal models, as we aim to work with yearly available unemployment rates on the five BRICS countries to foresight possible values for these rates on the future. Methodologically, we share a similar modelling procedure following a three-stage iterative procedure as proposed by Box, Jenkins & Reinsel (2008), plot visualization of autocorrelation, parameters definition and diagnostic checks for the residuals. All steps of these process are best illustrated with practical data and will be further detailed on the results and discussions section.

4.3.4. Artificial Neural Networks – ANN.

Derived from biological systems, the human brain particularly, Artificial Neural Networks can learn from past signals and from these generate based results. Imagine a network of "neurons" organised in layers where inputs form a bottom layer, and the forecasts (or outputs) form the top layer (Hyndman & Athanosopoulos, 2021). ANNs nowadays may be applied on a wide variety of intents in many different fields of business, industry, and science (Widrow, Rumelhart & Lehr, 1994; Zhang, Patuwo & Hu, 1998).

An overview of basic ANN in special on the interest of forecasting purposes could be summarized on three basic levels: The input layer, which is the one that receives original data as basic input; hidden layers, the intermediate level that process the designated information; and the output layer, where are the final output produced, the forecast in this case. Again, following the Zhang, Patuwo & Hu (1998), considering the intention assumed for this research we aim to forecast these rates on a 10-year future horizon from the available data, the inputs are presumed to be past values for unemployment whereas the output is their potential future values.

ANN in these parameters follows the function: $UR_{t+1} = f(\, UR_t, UR_{t-1}..., UR_{t-p})$ where UR_t is the factual unemployment rate in a time t. While dealing with time series data in a unique variable, as is the case on the application here, lagged values of this variable can be assumed as inputs for the neural network which can be defined as a Neural Network Autoregression (NNAR) model (Hyndman & Athanosopoulos, 2021). On the operational process of this method, we follow again Hyndman & Athanosopoulos (2021), considering only feed-forward types of networks that have one hidden layer and the notation $\mathrm{NNAR(p, k)}$, where p are the lagged inputs and k the number of nodes in the hidden layer level.

`NNETAR` function from the `fable` R-Studio package fits for the best $\mathrm{NNAR(p, k)}$ model automatically if p and k are not predefined they are selected automatically according to the optimal number of lags based on the AIC criteria. For one period ahead of forecast it is used available historical inputs, for two steps ahead, the previous one-step forecast is an input added to historical data and so on up to all intended period is covered (Hyndman & Athanosopoulos, 2021).

4.3.5. Scenario forecasting.

Having in hand the results from the four above presented applications we move to the fifth and final methodological proceeding, the scenario-based forecasting. We understand scenario forecasting here as an intersection between two separate and most solidified methods, the quantitative forecasting ones (in this research being used the STL, ETS, ARIMA and ANN) and the judgmental qualitative methods, such as scenario planning. Aim of this approach is to generate scenarios that are pre-substantiated by the statistical results potentially indicating a more robust effort as we intent to go further the numbers on isolation and neither rely on completely qualitative proceedings (Hyndman & Athanosopoulos, 2021).

Applying this scenario forecasting we presume that the before-achieved statistical results gives some reliable outputs to build scenarios more adequately informed and subsided by real data, which consequently, would be helpful for decision makers on mitigating judgmental biases while offering statistical basis to the foresight outcomes of unemployment rates in the future, specifically on the BRICS countries context, that could lead to a better understanding of projected results and therefore better informed decisions.

On taking by account what could be predictable, through statistical forecast methods, and judging the potential uncertainties, via scenario building, the idea is to improve analysis and eventually decision-making processes having a solid understanding of unemployment data information by in advance learn more about the systematic past values of this variable to later organize the potential future of these in feasible scenario propositions (van der Heijden, 2000).

As future remains uncertain and cannot be predicted accurately, instead of relying on a single forecast, scenario planning propositions gives the opportunity to envision multiple alternatives for what lies ahead (Coates, 2000). Spreading on relevance of scenario planning as an organizational and institutional tool for future thinking goes back to the 1950s, by Herman Kahn, and 1970s by the Royal Dutch Shell case, one of the pioneering companies on the incorporation of scenario planning into its strategic management process (Coates, 2000; Jefferson, 2012).

Scenarios applied in business and government planning according to Coates (2000) are into two categories. First category is predominantly developed to stimulate future thinking while the second one is more a practical tool, that aims to explain or explore consequences of some policy decisions, either on theoretical or factual applications (Coates, 2000). As we are dealing with actual data regarding unemployment rates on BRICS countries and the forecasts to be made are data-driven by the available information, we believe that our scenario forecasting resonates more with the second.

Still on the specific purposes of our research, scenario forecasting can offer supplemental information from the original scenario planning and the forecast methods alone as we propose scenarios based considered a predefined outcome of possibilities derived from STL, ETS, ARIMA and ANN methods, being particularly useful as comprises more factual possibilities that may happen in the future basing on past information data.

Therefore, our scenarios proposal is built on the aim of examinate unemployment rates in the BRICS group in a foresight for 10-year in the future of unemployment rates within Brazil, Russia, India, China, and South Africa, using an empirically sound basis of four methodological forecast methods (STL, ETS, ARIMA and ANN) combined. In doing this, we believe that it will be possible to anticipate tendencies for the labour market, on unemployment specific, offering valuable insights that could help these countries to be better prepared for what may come in a near future.

Scenarios to be proposed here are directly concerned and based on unemployment rates, considering a three-way of possible avenues of forecasting: Unemployment acceleration, stability, and deceleration. All future values depending heavily on the factual and available data for these countries. About the building of the dataset to be assessed and information retrieving process, is dedicated the next topic.

4.3.6. Unemployment rates variable and the building of the dataset.

In general, this research working dataset includes a baseline variable that is yearly unemployment rates from the five BRICS countries: Brazil, Russia, India, China, and South Africa. Some of these countries have data available in different levels like monthly and quarterly, as Brazil for example. Other as South Africa and India are more difficult to assess information with consistency in microlevels particularly with consistency from past years. Therefore, we decide to establish yearly rates as common indicator for all five countries considering this format have a better accessibility.

About timespan of data to be analysed we intent to work with as many as possible information available. Meaning from the earliest yearly data at disposal to the most recent ones, considering the momentum we are developing this research. Unemployment rates, for all five BRICS countries will be extracted from the World Bank repository of data, the Databank, which comprises information from International Labour Organization.

We decide to extract from World Bank due the existence of a package (`WDI`) and function available on R-Studio software, centring all data processing into one familiar and ease of use environment for our development. Given all the above defined, table 29 presents the variable to be analysed, from where they will be retrieved, and their natural form presented in the original database. Also, the timespan of information available for Brazil, Russia, India, China, and South Africa.

Table 29
Unemployment rates data information.
(Elaborated by the author).

Variable	Source	Aggregation	Periodicity	Timespan
Unemployment rates (UR)	World Bank – ILO	Weighted Average (%)	Annual	1991-2022

Notes: Unemployment refers to the share of the labour force that is without work but available for and seeking employment. Information is accessed and updated from ILOSTAT to World Bank dated to September 05, 2023. Available at: https://data.worldbank.org/indicator/SL.UEM.TOTL.ZS?view=chart

Data is mainly consistent in the timespans covered for all five countries. Only missing information is on the Russian unemployment rates by 2022. Therefore, we have in hand five timeseries, respecting each of the BRICS countries where for Brazil, India, China, and South Africa we have 32 observations (1991 up to 2022) each and for Russia we have 31 observations (1991 to 2021). We believe that in the methods to be applied missing data from Russia will not be harmful to the analyses, but if it is necessary, we may impute a median value considering the Russian data effectively available.

Considering the five BRICS countries, the main variable of interest and their respective timespans under analysis, descriptive statistics (mean, median, minimum, maximum, and standard deviation) for unemployment rates values from this research dataset are presented on the following table 30.

Table 30
Descriptive statistics for unemployment rates time series in each of the BRICS countries.
(Elaborated by the author).

Time series / Variable	Observations	Mean	Median	Minimum	Maximum	Standard deviation
UR in Brazil	32	9.44	9.44	6.03	13.90	2.22
UR in Russia	31*	7.29	6.54	4.50	13.30	2.41
UR in India	32	7.83	7.88	6.51	10.20	0.75
UR in China	32	4.03	4.48	2.37	5.00	0.79
UR in South Africa	32	21.90	20.90	19.50	29.80	2.52
UR in BRICS	159	10.10	7.96	2.37	29.80	6.48

Note: * Number of observations are referring to the unemployment rates at a given year within the country's timespan data. Russia length of information goes from 1991 to 2021 as other countries are from 1991 to 2022. All data, for five countries, were retrieved from World Bank/Databank on December 6, 2023.

First observation on the data retrieved considering the timespan covered being 1991 to 2022 for Brazil, India, China, and South Africa and 1991 to 2021 for Russia, and the presented descriptive statistics on table 30, it is possible to perceive that UR values are considerably higher in South Africa compared to the other four BRICS countries. On the opposite, China appears as the one in the group where unemployment is consistently low in comparison in all indexes, mean, median, maximum, and minimum values.

When observing values as a group we believe that the numbers presented on table 30 is solid. Mean and median values seem to represent well the five countries whereas the high standard deviation of 6.48 also captures well the disparities on the unemployment rates in South Africa comparing with the other countries in BRICS. Overall descriptive statistics captures well country dynamics, China as the most advanced economy tends to suffer less with unemployment while where the economic and intern social inequalities on South Africa implicates on a potential major problem on regard to unemployment.

From all the initial descriptive assessments and values of data presented on table 30, forecast will proceed next passing through the four defined methods: STL, ETS, ARIMA and ANN. Outputs of each method will have their accuracy checked to unveil which performs better and these results are the basis for the later presented scenarios for the unemployment rates in BRICS countries through the usage of the scenario forecasting.

Moving on for the next section empirical results and analyses for all the methodological proceedings defined to be used will be presented and discussed. Although we present the general premises for each method, we believe that they are better comprehended using real data application. Main purpose of the following section where BRICS dataset will be used to forecast unemployment rates for this group of countries.

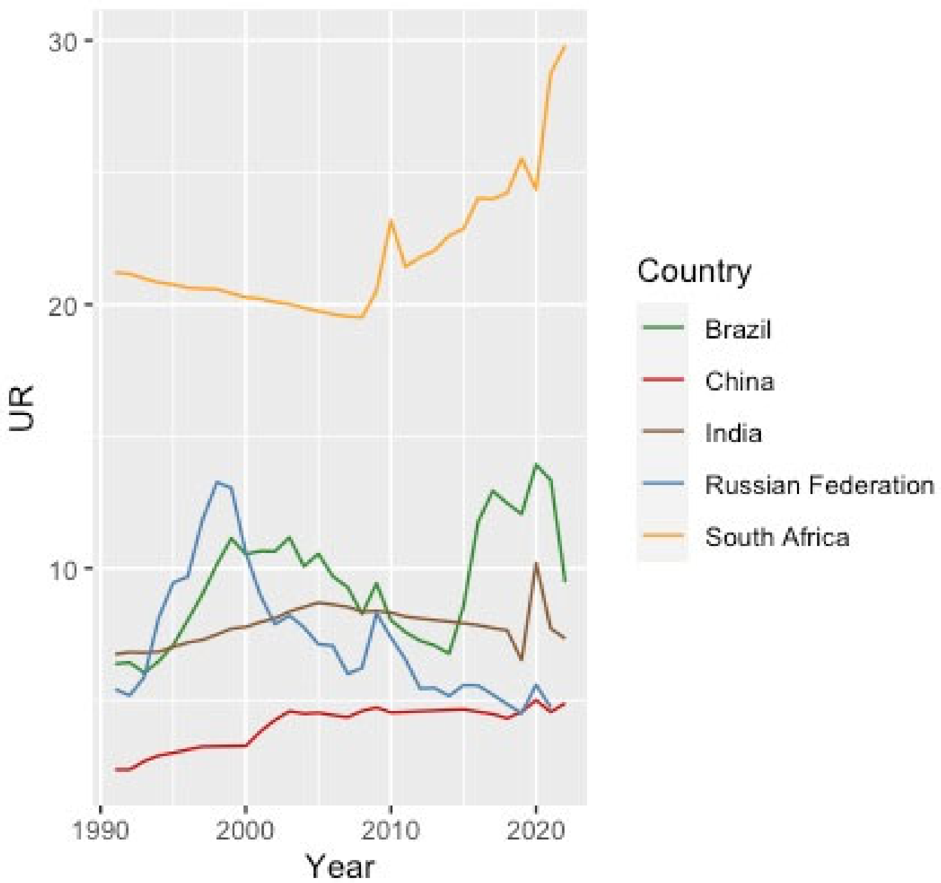
4.4. Results and discussions.

Presentation of the empirical results are following the early mentioned methods (Seasonal and Trend decomposition using Loess – STL, Exponential Smoothing Technique – ETS, and Autoregressive Integrated Moving Average – ARIMA, Artificial Neural Networks – ANN and Scenario forecasting) applying all of them at national level of data information about Brazil, Russia, India, China, and South Africa.

Data is corresponding to timeseries for the main target variable unemployment rates on yearly recurrency considering the five BRICS countries. Data will be analysed individually, country by country and aggregated, through the defined techniques presented on the previous section. Accuracy measurement for every modelling will be conducted to assess the performance of each method and produced outputs.

Before the STL, ETS, ARIMA and ANN applications, we first proceed with some initial assessments of the data. First conference about data behaviour will be made through visual presentation, considering the graphic timeseries for unemployment rates by each country on the BRICS. We use here each country full data available, meaning 32 observations for Brazil, India, China, and South Africa, 31 for Russia and 159 to all of them collectively. Following figure 24 presents unemployment rates behaviour for the five BRICS countries.

Figure 24
Unemployment rates in BRICS from 1991 to 2022.
(Elaborated by the author).



We present in the figures 24 a visual presentation that corroborates the values extracted from ILO and World Bank repository. All minimum and maximum indexes presented on table 30 are fitted on the plots and is noticeable that for every county unemployment peaks on 2020 and remaining high on the following years, representing the damaging effects from COVID-19 pandemic scenario (Kawohl & Nordt, 2020, Agrahari et al., 2021; Kumar et al., 2023).

Comparing the five lines which are respective to each the five BRICS countries, as presented on figure 24, we may observe as in a general observation Brazil, Russia, India, and China have a similar behaviour on their unemployment rates over the timespan covered. Brazil peaks higher than the other four on 13.90% of unemployment in 2020 while China has the lowest value on 2.37% in the early 1990s. On the opposite, South Africa separate itself from the others very clearly. Minimum unemployment on the country happening in 2008 at the rate of 19.50% is above all the maximum on the other BRICS countries whereas South African peak in 2022 is closing on 30% of the population not formally allocated on the labour force.

Alarming unemployment levels in South Africa since the political transition in 1994 is documented by Banerjee et al. (2008) that discusses how national labour supply interacted with stagnant labour demand implicating on unemployment rates increasing on the early 2000s and yet not being solved in recent years. Figure 24 lines shows that, at least visually, all five BRICS countries unemployment rates appears to be nonlinear over the years. This non-consistent behaviour, considering the type of data being analysed is expected and in line with similar studies (e.g., Chakraborty et al., 2020) as the one we are developing here, what may be an initial indication that we have a suitable data to proceed.

Assuming the nonlinearity of unemployment rates, as a second step to early checking on the data we proceed to test for stationarity. Augmented Dickey-Fuller test is performed for each of the unemployment timeseries previously presented on figure 24. Results of the test statistics and p-values are presented on the following table 31.

Table 31
Results from Augmented Dickey-Fuller tests for stationarity.
(Elaborated by the author).

Time series for Countries and BRICS	**UR**
Brazil	Test statistic: -2.48
	p-value: 0.38
Russia	Test statistic: -3.25
	p-value: 0.09
India	Test statistic: -1.98
	p-value: 0.58
China	Test statistic: -1.51
	p-value: 0.76
South Africa	Test statistic: 0.81
	p-value: 0.99
BRICS	Test statistic: -1.46
	p-value: 0.80

Note: Null hypothesis of ADF test is that the data being analysed has at least one unit root, or data is not stationary. We assume a statistical significance level on the threshold by 0.05 (or 5%) to check our results. P-values presented on the table fails to reject the null hypothesis of non-stationarity in all six timeseries, considering that the p-values are all > 0.05. Therefore, we have evidence to believe that our data has a unit root and, consequently, a stationary performance with time-dependent structure.

Dickey–Fuller test (1979) checks for the null hypothesis that presumes a timeseries as non-stationary or does not have a unit root in it. If this assumption fails, meaning that the timeseries is a non-stationary type, one of the suggestions that could be inferred is the existence of a form of time-dependent structure over time, as well does not have a constant variance over time, a common behaviour on real-life variables and data (Dickey & Fuller, 1979; Cheung & Lai, 1995; Mushtaq, 2011).

Results presented on table 31, are to be checked by the respective p-values for timeseries of unemployment in each of the five BRICS countries. Having as a parameter the significance level of 0.05, we fail to reject null hypothesis from ADF test considering all p-values obtained by country and for the BRICS group as well, meaning that our working unemployment data may be perceived as a stationary type and having, at least, a unit root existing in all six timeseries that will be later analysed.

Testing data for stationarity is particularly important for research where the underlying variables are based on timely distribution (Mushtaq, 2011). Even more for Datasets built on the collection of macroeconomic variables, as is the case for the timeseries being analysed in this study. This type of variable and timeseries tend to be naturally non-stationary and nonlinear in their essence (Chakraborty et al., 2020) due to the inherent fluctuations these indexes suffer depending on the country, their economic policies, and other particular characteristics.

A timeseries having some form of time-dependent structure without a constant variance over time in macroeconomic cases are indeed the expected as the data presented on table 31 confirms. Important to acknowledge that the usage of ARIMA modelling, for example, which may be used for tracking linear tendencies in stationary time series data, are useful for the purposes of this study (Cheung & Lai, 1995; Chakraborty et al., 2020), this justifies the initial assessment we perform for linearity and stationarity.

From hereafter we proceed to the specific application of the four statistical forecast methods defined: Seasonal and Trend decomposition using Loess – STL, Exponential Smoothing Technique – ETS, Autoregressive Integrated Moving Average – ARIMA, and Artificial Neural Networks – ANN, performing in each method analyses for every Brazil, Russia, India, China, South Africa, and for the BRICS as a group as well.

4.4.1. Seasonal and Trend decomposition using Loess – STL application.

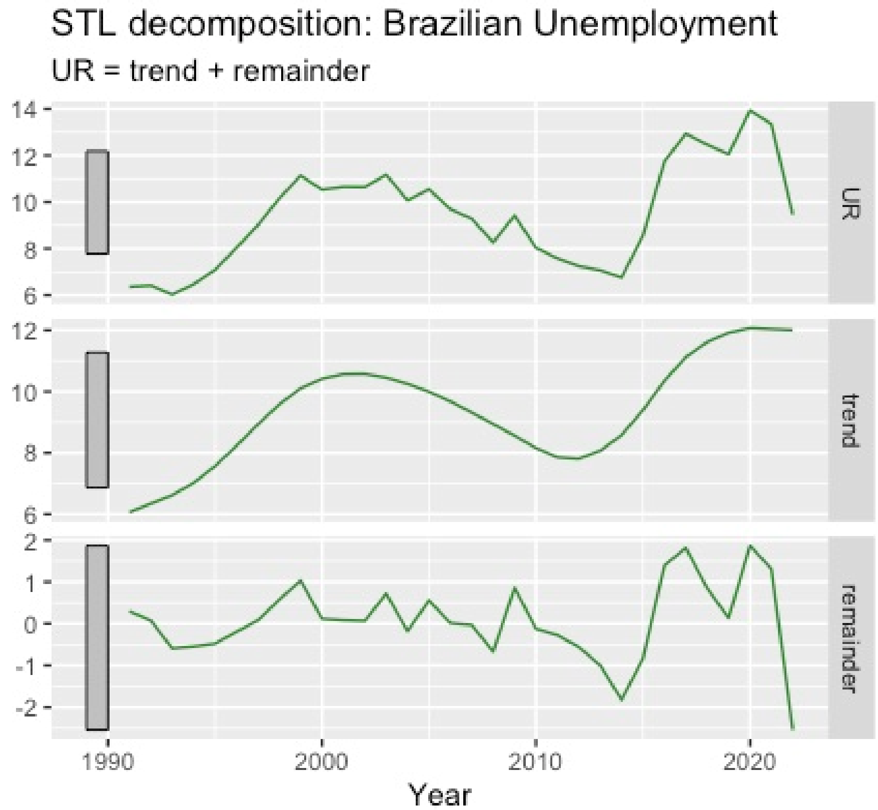
On this section we move to the application of STL statistical framework to produce forecasting for unemployment rates in BRICS countries. Each forecast will be made by country individually and after these we use the total number of observations to assess BRICS as a group. All proceedings will be following the steps presented on Hyndman & Athanosopoulos (2021).

Before the forecasting application in specific, we assess some STL features. Some caveats to be elucidate on first-hand. Usually, unemployment data are available on their seasonally adjusted format, to highlight variation due to the underlying state of the economy rather than the seasonal variation (Hyndman & Athanasopoulos, 2021). An increase in unemployment due to more people leaving school and seeking for a job is seasonality, while an increase in unemployment due to an economic crisis, as the COVID-19 pandemic, is non-seasonal.

It is from more interest to econometrics and business analyses to assess the non-seasonal variations. Therefore, if the purpose is to check for significant turning points in an observing series while reasoning to any changes in direction, then the most suitable data to be used is the trend-cycle component, rather than the seasonally adjusted information in variables that are already disclosed in this adjusted format (Hyndman & Athanasopoulos, 2021).

Opening application to be performed and presented are referring to the unemployment rates in Brazil over the timespan of data available going from 1991 to 2022. First proceeding is to decompose the Brazilian timeseries by using `STL` function. Following figure 25 presents the decomposition results to the data referring to Brazil.

Figure 25
STL application for unemployment in Brazil.
(Elaborated by the author).



Brazilian unemployment rate timeseries may be understood by an equation expressed as: $\text{Brazilian\_UR}_t = T_t + R_t$, seasonality is not present (all data are adjusted for yearly rates) and what composes these unemployment rates is a trend and residual values. Going back to the proposition of this method by Cleveland et al. (1990), we see a visual trend for Brazil that after some consistent increasing on unemployment after 2010s years, peaking in 2020, the tendency seems to indicate a decreasing trend. To check this statistically we employ the `feat_stl` function, to identify features on the STL decomposition. Outputs from this are presented on the following table 32.

Table 32
Features of STL decomposition for Brazilian unemployment rates.
(Elaborated by the author).

	trend_strength	*spikiness*	*linearity*	*curvature*	*stl_e_acf1*	*stl_e_acf10*
Brazil	0.82	0.00	6.32	-0.19	0.25	0.30

Two columns on the output are particularly important. On "trend_strength" we see the potential of the identified tendency of the data occur into subsequent values. Here we have an index by 0.82 of probability that the "curvature" found on the other column replicates on the future. Therefore, regarding Brazilian unemployment rates we have 82% of chance that in a foreseeable future some decline on unemployment may occur, more specifically by around -0.19 from previous observed values.

From these initial indicatives we move on to the forecast in specific. First step that will be adopted as a common practice for Brazil, the other four BRICS countries and when analysing them as a group is to divide data in portions, the training, and test sets of information. Training data is used to estimate the parameters in a forecasting method whereas test data is used to evaluate the accuracy of the method applied. Considering that the test portion is not used into the forecast process, the premise is on being apart it could provide a reliable indication of how well the model is expected to forecast new data (Hyndman & Athanosopoulos, 2021).

On the Brazilian as well into all future applications we are using an 80/20 proportion for the division on our datasets having as cutting point the year of information available. Therefore, for Brazil, India, China, and South Africa training portion considers unemployment rates from 1991 to 2016 (26 observations per country) and test set is from 2017 to 2022 (6 observations). Russia training set is from 1991 to 2016 as well, but the test set is from 2017 to 2021, as this is the last year of available Russian information.

Returning for the Brazilian forecast of unemployment, we first assess three simple forecasts proceeding to identify the most fitted one for later specified the best found into the final forecast using `STL` function. Mean, Naïve and Drift methods are checked into an adjusted model using the data for unemployment from Brazil following the above presented division of training and test datasets.

First, we apply a point forecast proceeding and later a distributional and interval prediction forecast, in this second moment we are already using the best method (if Mean, Naïve or Drift) with STL decomposition. On to implement these steps described above, once again, we are following the Hyndman & Athanasopoulos (2021) textbook and the `fable` package proceedings. Figure 26 presents the point forecast proceeding for unemployment in Brazil using the Mean, Naïve and Drift methods. After this visual presentation the table 33 shows the accuracy check for these three methods for the statistical identification of the best suited to proceed onto other forecast applications.

Figure 26
Point forecast for unemployment in Brazil.
(Elaborated by the author).

Table 33
Accuracy checking for Drift, Mean, and Naïve applications.
(Elaborated by the author).

Model	*RMSE*	*MAE*	*MAPE*
Drift	1.63	1.17	10.60
Mean	3.87	3.59	28.00
Naive	1.56	1.38	11.60

Mean and Naïve methods are outperformed when observing both figure 26 and table 33. On the table, MAE and MAPE have lower values on Drift in comparison with the other methods suggesting that as the most fitted method to proceed. Drift is a variation from the naïve method, the distinction is that the former allow that forecasts may fluctuate over time. These variations over time, the drift, are settled to be the average change projected considering the historical data (Hyndman & Athanosopoulos, 2021).

Using the drift method and the STL decomposition, we proceed to further forecast proceedings. Considering the training and test portions of data the function `hilo`, which converts the forecast distributions into intervals that presumes, by default, 80% and 95% of confidence of these prediction intervals returned (Hyndman & Athanosopoulos, 2021). We present these predicted and distribute intervals on the following figure 27. Three scenarios were forecast by the application on a bootstrapping procedure on the training portion.

Figure 27
Predicted intervals and distributional forecasts.
(Elaborated by the author).

Scenario 2 presented on figure 27 seems to project the most damaging unemployment rates in the future, peaking on 14% around the year 2024 while the lowest values appear to be on scenario 3 around 2023, where the projected unemployment is under 7%. Scenario 1 is the most adjusted on the real data when considering actual values available on the test portion of the data and comparing with figure 26 historical data. Table 34 presents the accuracy of this second forecast applied.

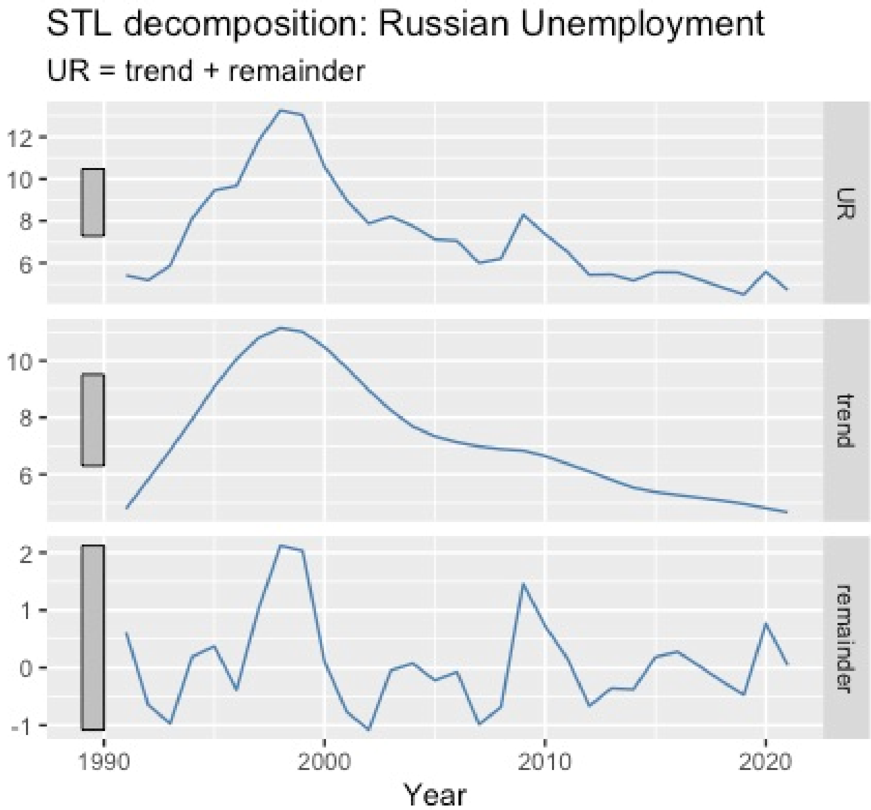
Table 34
Accuracy checking for the STLF forecast for Brazilian unemployment rates.
(Elaborated by the author).

Model	*winkler*	*percentile*	*CRPS*
STLF	7.21	0.93	0.92

STLF model as presented on table 34 is indeed an STL application where the F stands for the forecast applied on the Brazilian data. Still on table 34 we focus on the CRPS (Continuous Ranked Probability Score) (Gneiting & Katzfuss, 2014), that generalizes the MAE on the case of probabilistic forecasts and have lower values than all the measurement of errors presented on table 33, suggesting that this second forecast, the STLF, performs better than the first ones using Drift, Mean and Naïve methods.

Moving on for the next countries we replicate the same methods presented using Brazilian data. When other four countries results are reported we close on with an analysis for the BRICS as a collective. Russia is the following country to have its unemployment rates assessed. We remember that Russian availability of information is one-year short comparing with the other nations being observed, going from 1991 to 2021. Figure 28 presents decomposition plot extracted from `STL` function of data referring to Russia.

Figure 28
STL application for unemployment in Russia.
(Elaborated by the author).



Russian STL equation may be defined as: $\text{Russian\_UR}_t = T_t + R_t$, seasonality as in Brazilian case is not present unemployment rates are composed by a trend and residual values. Applying the `feat_stl` function we may have a statistical identification of the features on this STL decomposition. Results are presented on the following table 35.

Table 35
Features of STL decomposition for Russian unemployment rates.
(Elaborated by the author).

Russia	*trend_strength*	*spikiness*	*linearity*	*curvature*	*stl_e_acf1*	*stl_e_acf10*
	0.89	0.00	-7.10	-4.74	0.41	0.40

Column "trend_strength" on table 35 presents an index of 0.89 of probability that the tendency present on the data endures for future values. The "curvature" value found on -4.74 anticipates some accentuated decline on unemployment rates for Russia considering the historical data. It seems reasonable, considering that the peak of unemployment in the country are on the 1990s and since then levels of unemployed people remain declining.

Mean, Naïve and Drift methods are checked into an adjusted model using the data for Russian unemployment into a point forecast proceeding to unveil the most fitted method of these three for later proceed with the distributional and interval prediction using the STL decomposition forecast (STLF). Figure 29 presents the point forecast proceeding for unemployment in Russia. Following table 36 illustrates the accuracy check for three methods initially applied (the Mean, Naïve and Drift proceedings).

Figure 29
Point forecast for unemployment in Russia.
(Elaborated by the author).

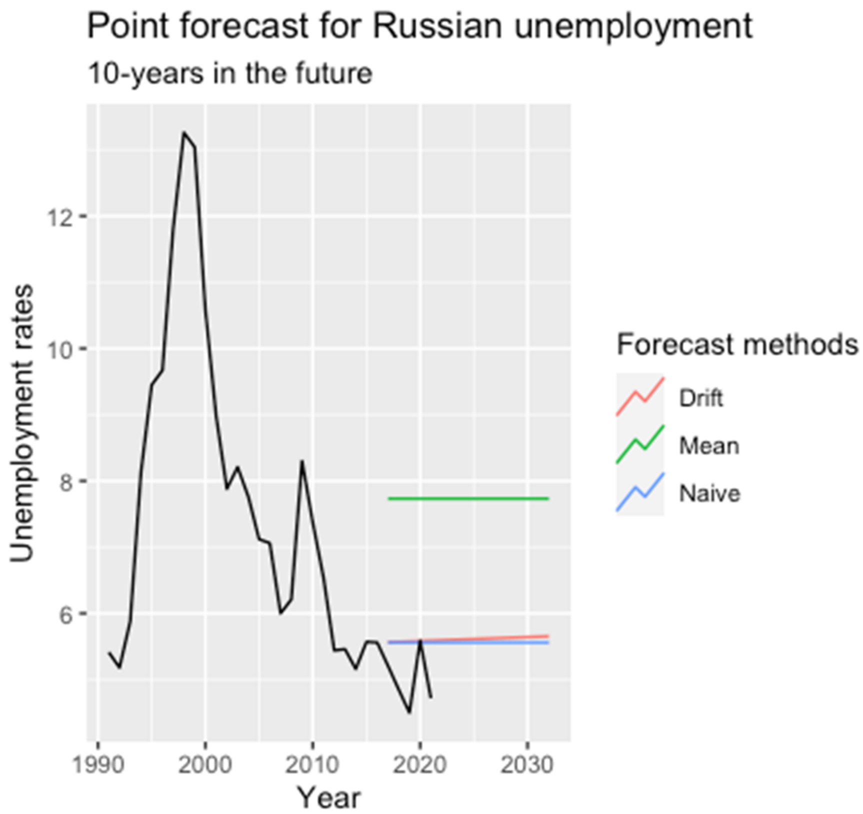


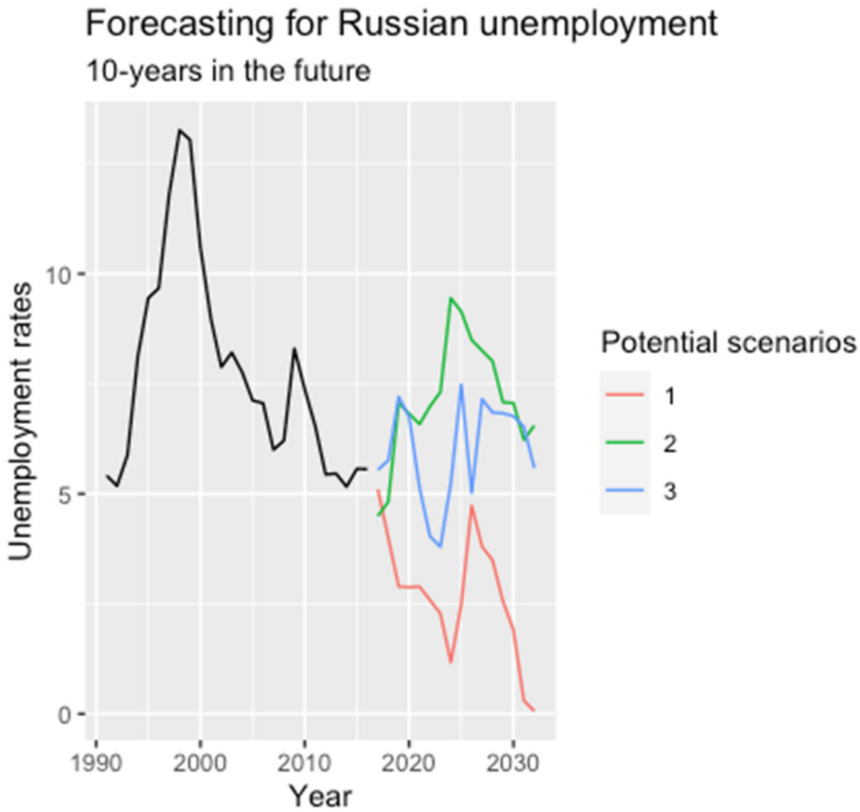
Table 36
Accuracy checking for Drift, Mean, and Naïve applications.
(Elaborated by the author).

Model	*RMSE*	*MAE*	*MAPE*
Drift	0.71	0.60	12.80
Mean	2.78	2.76	56.40
Naive	0.70	0.60	12.60

Table 36 results shows that Naïve method slightly outperforms the Drift whereas the Mean is the one with the most non-suitable results. Russian data therefore has a different best method in comparison with the Brazilian information presented on table 33 where the Drift has the best performance. Although the indexes being is not largely distinct, still we decide to proceed with Naïve, nonetheless. Setting all forecasts based on the values of last observation, as the Naïve method works, in some cases generate reliable results for many economic timeseries according to Hyndman & Athanasopoulos (2021).

Applying naïve method and the STL decomposition, we proceed to further forecast proceedings using `hilo` function to forecast distributions into intervals. These predicted and distributed intervals projected are presented on the following figure 30.

Figure 30
Predicted intervals and distributional forecasts.
(Elaborated by the author).



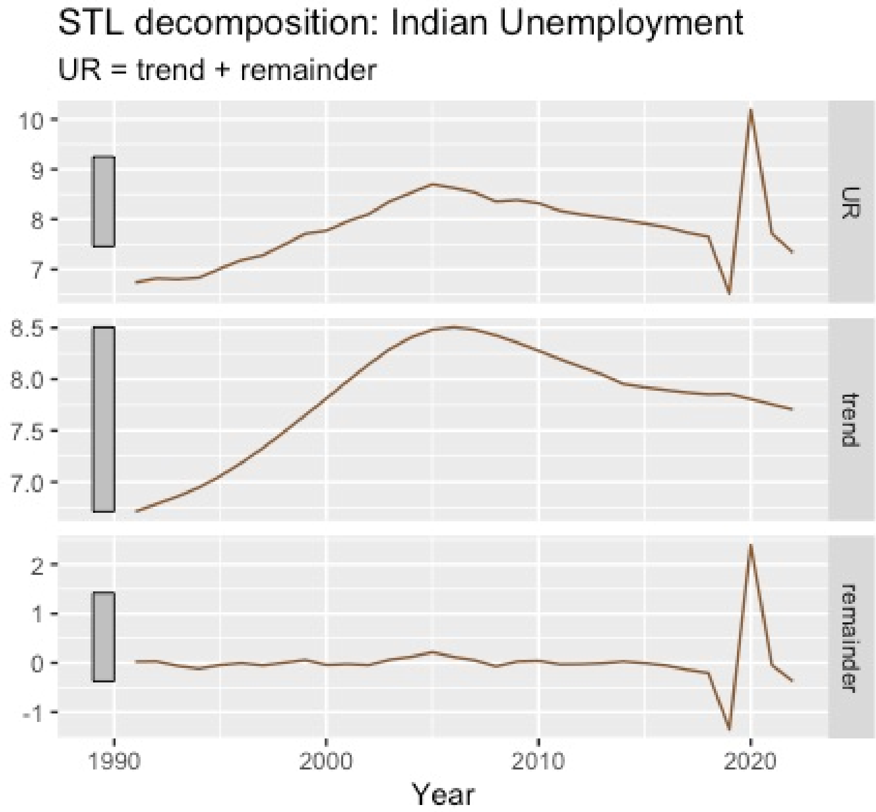
Scenario 1 from figure 30 is the one projecting that unemployment rates in Russia would remain decreasing, closing to 0% after 2030. Scenario 2 anticipates the opposite, suggesting that unemployment could peak close to 10% around 2025 whereas the blue line from scenario 3 suggests a more adjusted context for Russian unemployment with no highs as the one that factually happened on the 1990s and peaking at about 10% after 2030 as could happened on scenario 1. Table 37 presents the accuracy of this application.

Table 37
Accuracy checking for the STLF forecast for Russian unemployment rates.
(Elaborated by the author).

Model	*winkler*	*percentile*	*CRPS*
STLF	7.77	0.55	0.54

For Russia, similar as happened for Brazil, STLF have an overall better performance only in evaluation with the mean method, while the drift and naïve comparison depends on the metric being observed. This may be an indication for the further applications we intend to proceed through the ETS, ARIMA and ANN. Onto next country, figure 31 presents decomposition lines from `STL` function now on Indian data.

Figure 31
STL application for unemployment in India.
(Elaborated by the author).



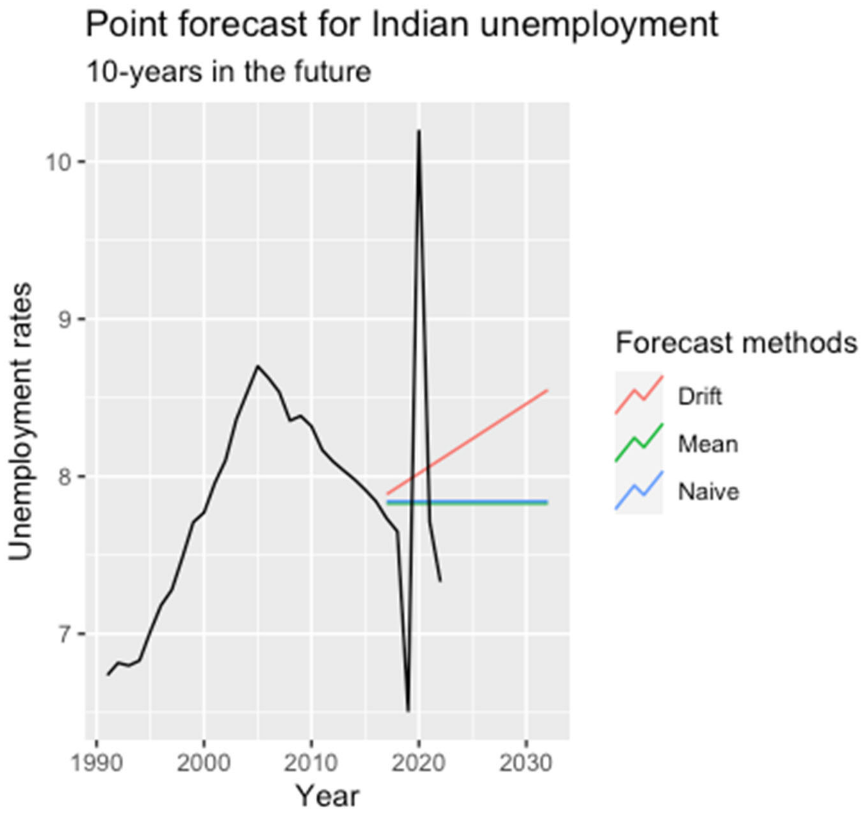
Similarly with the two already discussed countries, unemployment rates in India may be explained by: $\text{Indian\_UR}_t = T_t + R_t$. It could be visually perceived a somewhat constant trend on rates that when have a lowering on the end of 2010s years suffers a spike on 2020 potentially due to the COVID-19 pandemic.

Table 38
Features of STL decomposition for Indian unemployment rates.
(Elaborated by the author).

India	*trend_strength*	*spikiness*	*linearity*	*curvature*	*stl_e_acf1*	*stl_e_acf10*
	0.54	0.00	1.60	-2.29	-0.37	0.16

Results presented on the above illustrated table 38 extracted from `feat_stl` function, presents some interesting insights regarding the India unemployment rates. The country appears to have some tendency on decrease their unemployment levels higher than Brazil, here is on -2.29 while for Brazil is -0.19, however the probability of this happening is around 50%. This result suggests a twofold avenue for Indian labour market policies unemployment, a trend to decline if some serious efforts are taken into practice.

Figure 32
Point forecast for unemployment in India.
(Elaborated by the author).



Point forecast is applied into the Indian dataset replicating what was used for Brazil and Russia. Visual presentation is on above presented figure 32 and the three basic methods (Mean, Naïve and Drift) results are given on table 39, once again, are following the steps presented by Hyndman & Athanasopoulos (2021) textbook.

Table 39
Accuracy checking for Drift, Mean, and Naïve applications.
(Elaborated by the author).

Model	*RMSE*	*MAE*	*MAPE*
Drift	1.13	0.87	10.80
Mean	1.13	0.76	9.22
Naive	1.13	0.77	9.35

According with results on table 39, we have statistical evidence to believe that the most fitted method to proceed for the STL forecast is the Mean method, which in summary forecasts all future values as equal to the average of, in this case, the Indian historical data (Hyndman & Athanasopoulos, 2021). We the mean method combined with the decomposition from `STL` as well the `hilo` function to forecast distribution and predictive intervals thar are presented on figure 33 referring to Indian data information.

Figure 33
Predicted intervals and distributional forecasts.
(Elaborated by the author).

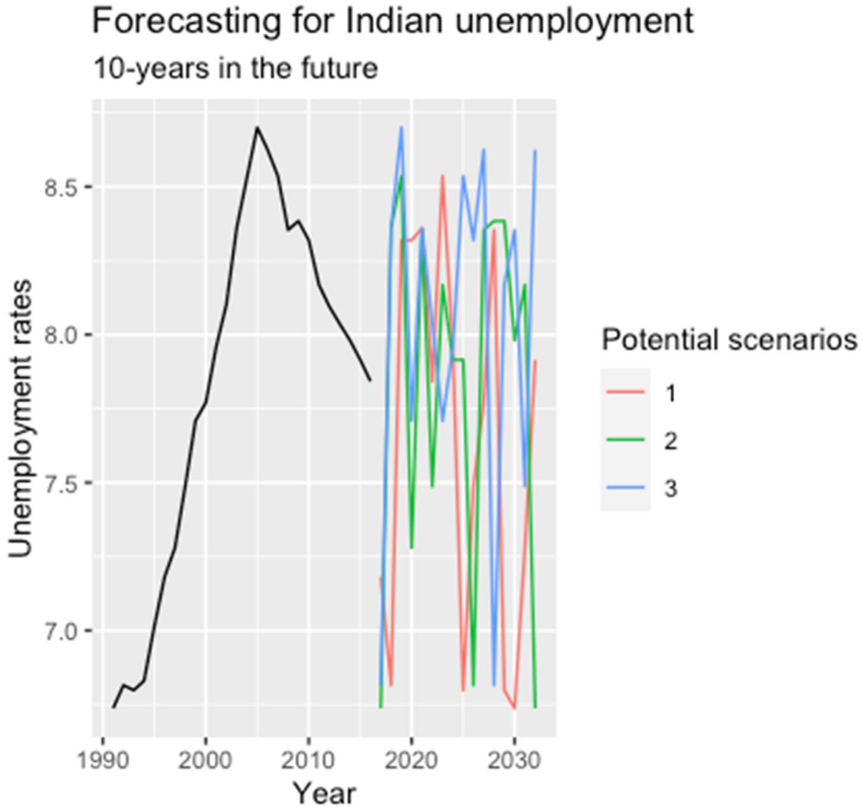


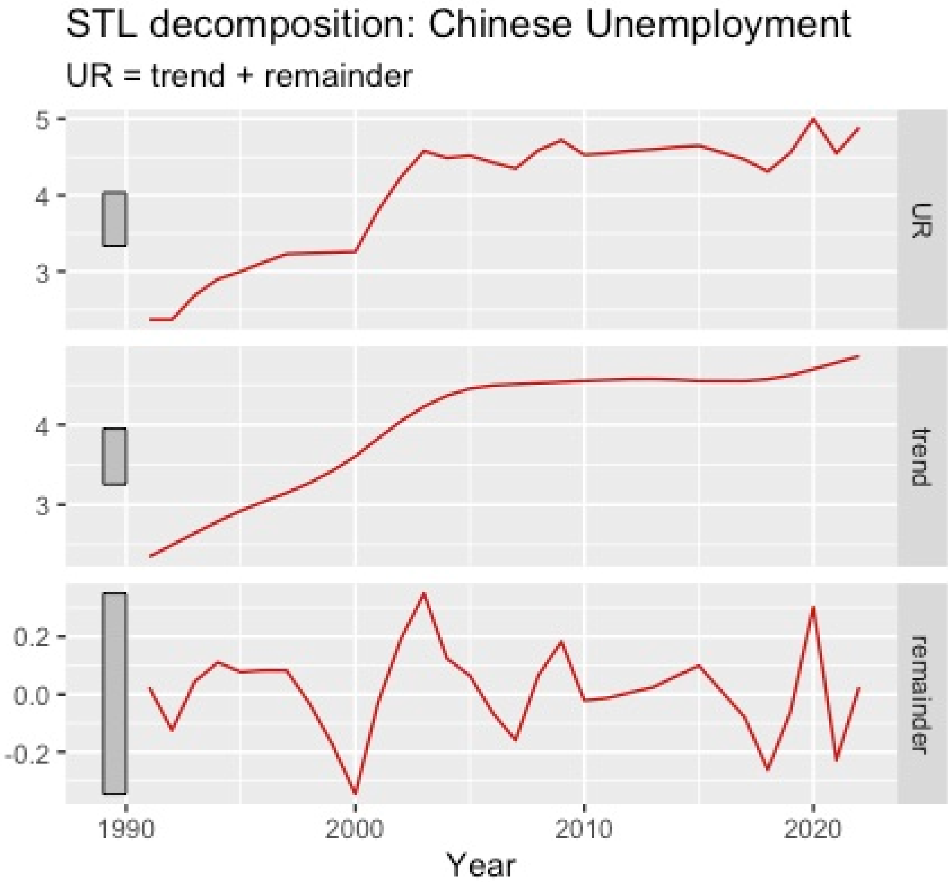
Figure 33 shows well how the trend strength obtained and presented on table 38 replicates well into intervals suggested for unemployment rates in India. All scenarios projected goes on consistent peaks both on high and low levels of unemployed people, replicating the 54% of the trend discussed earlier. Lines are mixed in a manner that even difficult the identification of which scenario would be a pessimist or optimist one. It is possible however to presume a consistency on these levels, considering that unemployment may not go on a lower of 6% neither a peak of 9% in a foreseeable future.

Table 40
Accuracy checking for the STLF forecast for Indian unemployment rates.
(Elaborated by the author).

Model	***winkler***	***percentile***	***CRPS***
STLF	13.40	0.66	0.66

STLF model for data from India appears as well to be the most adequate according to the results illustrated on table 40. Here, the case is similar with Brazilian data, where the Continuous Ranked Probability Score (CRPS), and here percentile also, presents a better performance from the before checked mean, drift and naïve method. On following figure 34 we start `STL` proceedings for Chinese dataset.

Figure 34
STL application for unemployment in China.
(Elaborated by the author).



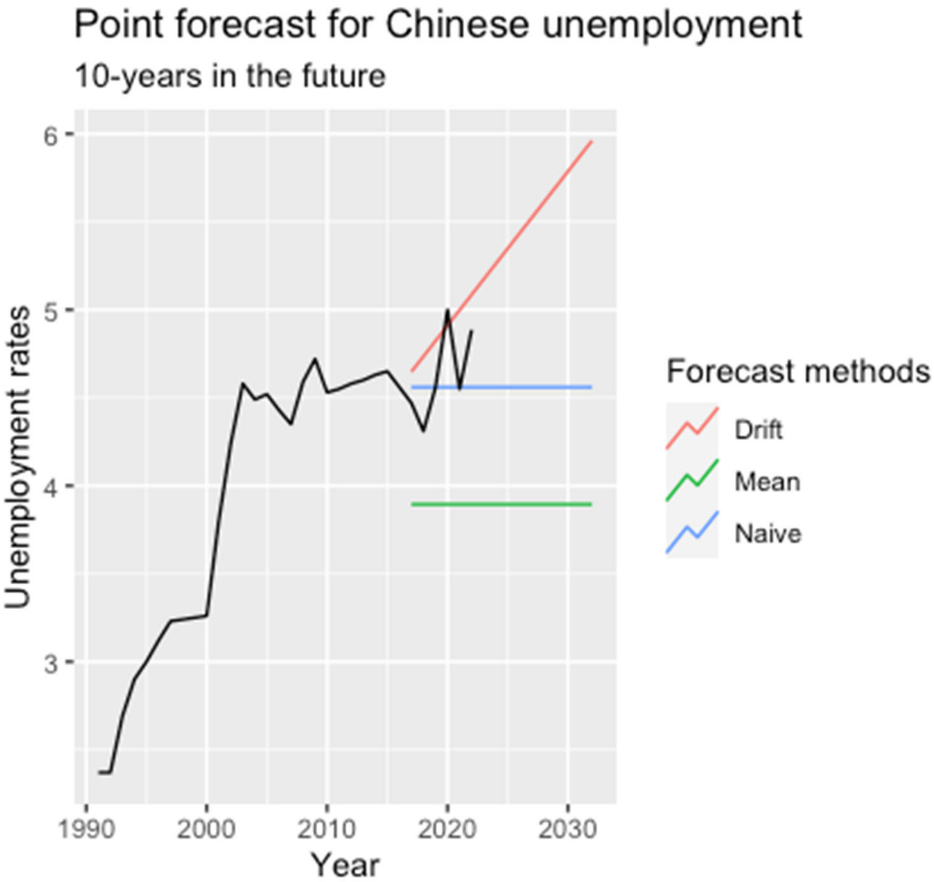
STL Chinese equation may be defined as: $\text{Chinese\_UR}_t = T_t + R_t$, seasonality again does not exist as the data is already adjusted. Using `feat_stl` function we have statistical of the features on this decomposition, results presented on table 41.

Table 41
Features of STL decomposition for Chinese unemployment rates.
(Elaborated by the author).

China	*trend_strength*	*spikiness*	*linearity*	*curvature*	*stl_e_acf1*	*stl_e_acf10*
	0.96	0.00	3.92	-1.67	0.23	0.41

A 0.96 on column "trend_strength" as presented on table 41 presents a strong suggestion that the tendency present on the data endures for future values. Curvature obtained as -1.67 anticipates some level of declining on unemployment considering Chinese historical data. Mean, naïve and drift methods are checked into a point forecast proceeding to unveil the most fitted method of these three. Figures 35 and table 42 checks.

Figure 35
Point forecast for unemployment in China.
(Elaborated by the author).

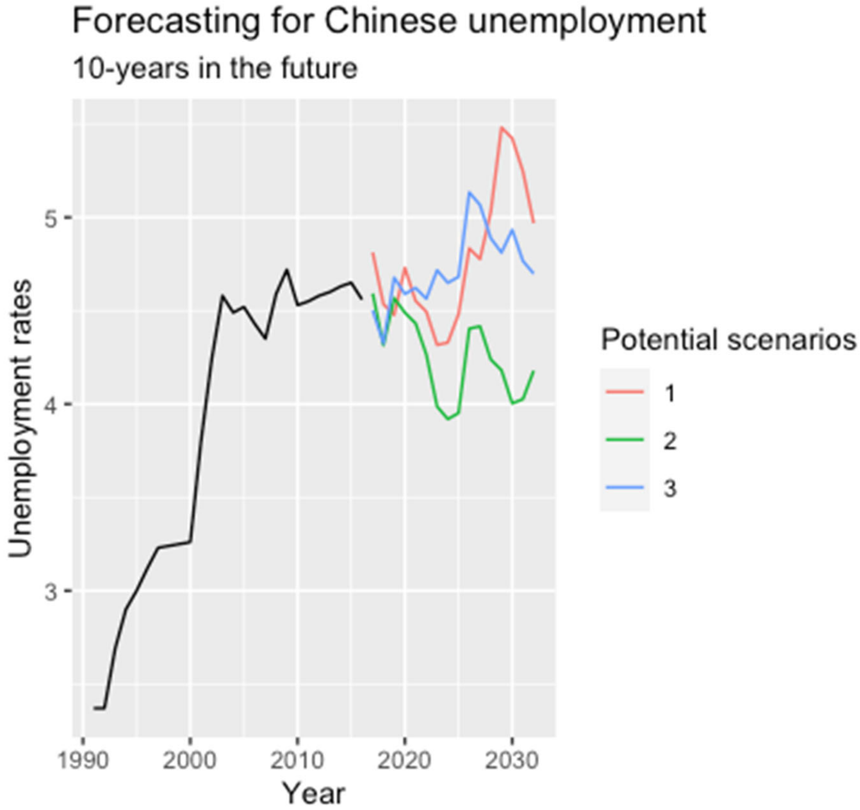


While on figure 35 are the lines for each of the three initial forecast proceedings, the following table 42 illustrates the accuracy check for these and how the Naïve method performs better than the other regardless the measurement of error selected. We remember that naïve establishes all forecasts based on the values of last observation on the training dataset. Using naïve method and the STL decomposition, we proceed to further forecast proceedings using the `hilo` in predicting distributions and intervals for unemployment. These values are presented on the following figure 36.

Table 42
Accuracy checking for Drift, Mean, and Naïve applications.
(Elaborated by the author).

Model	***RMSE***	***MAE***	***MAPE***
Drift	0.30	0.27	5.89
Mean	0.77	0.73	15.70
Naïve	0.25	0.19	3.92

Figure 36
Predicted intervals and distributional forecasts.
(Elaborated by the author).



Chinese potential scenarios are more like Brazil and Russia than the Indian projections. Again, we have three well distinct avenues for unemployment China, scenario 1 projects the worst-case for unemployment levels, suggesting a peak on early before 2030 by around 6%. On the opposite, scenario 2 presents the lowest possible closing on 4% as well in 2030. Scenario 3 is somewhat a mixed for the other two, crossing projections on both, at the beginning and the end of projections from scenario 3 but not lowering at any point the scenario 1. Overall, and in comparison, Chinese unemployment appear to be more stable than the other countries observed so far. Table 43 presents the accuracy of the STLF applied and presented on figure 36.

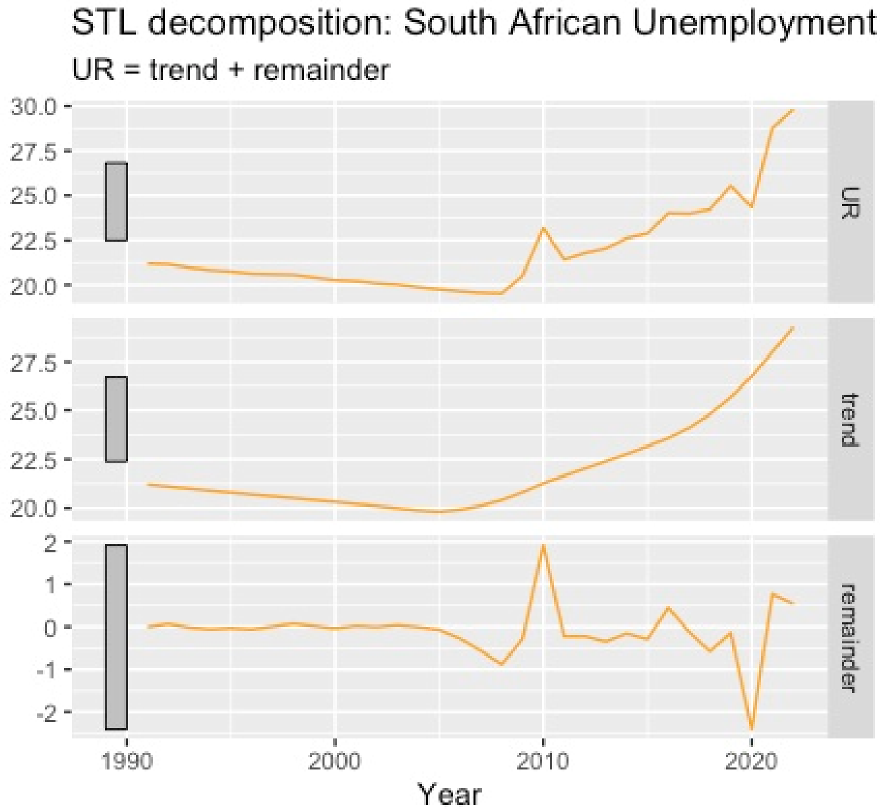
Table 43
Accuracy checking for the STLF forecast for Chinese unemployment rates.
(Elaborated by the author).

Model	*winkler*	*percentile*	*CRPS*
STLF	1.19	0.14	0.14

STL forecast for the Chinese data outperforms the other three forecast techniques including the naïve, previously suggested as the most adequate. Similar as happened on the point and distributional and interval projections of unemployment rate in Brazil, Russia, and India, on China the STLF method presents the best performance on the projected future values for unemployment.

We close on the BRICS countries individual assessments going for South African data. Figure 37 are presented the decomposition lines from `STL` function applied.

Figure 37
STL application for unemployment in South Africa.
(Elaborated by the author).



South African unemployment rates therefore may be as: $\text{SouthAfrican\_UR}_t = T_t + R_t$. Table 44 presents the output from the `feat_stl` function.

Table 44
Features of STL decomposition for South African unemployment rates.
(Elaborated by the author).

South Africa	*trend_strength*	*spikiness*	*linearity*	*curvature*	*stl_e_acf1*	*stl_e_acf10*
	0.94	0.00	10.60	8.65	-0.10	0.27

South African data is the one of the BRICS countries that presents a positive curvature. Not only is a positive value, is a relatively large one, by 8.65. A tendency of future values for unemployment in South Africa in a growing trend while trend strength on 0.94 indicates that the probability of the forecast values for number of unemployed people continues to accelerate, confirming some dysfunctionalities of labour market in South Africa (e.g., Nkoane & Seeletse, 2021). Figure 38 presents the point forecasting.

Figure 38
Point forecast for unemployment in South Africa.
(Elaborated by the author).

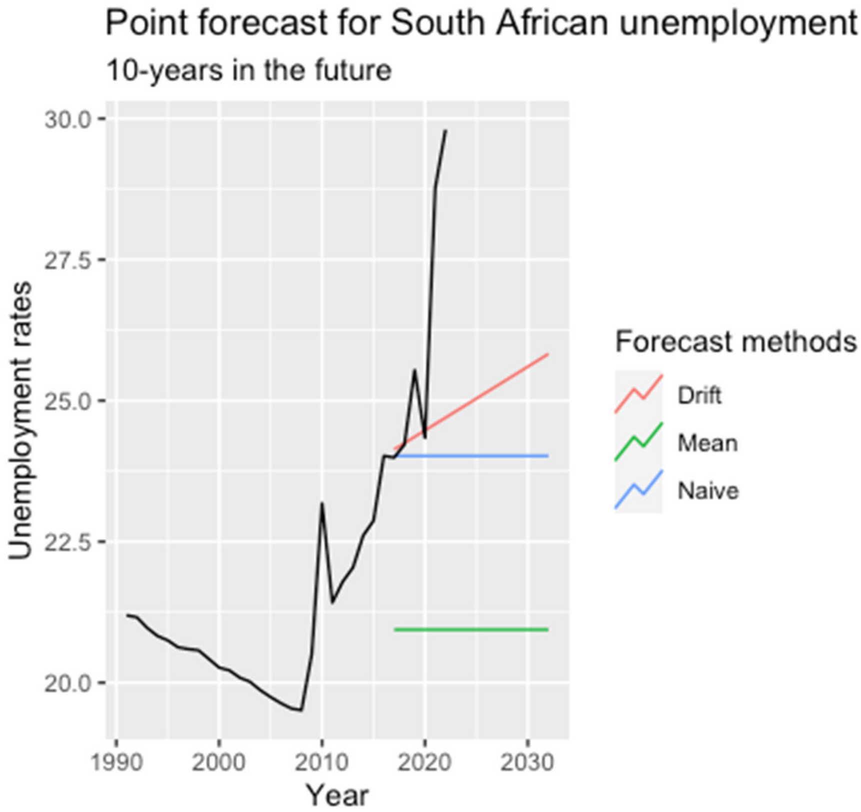


Table 45
Accuracy checking for Drift, Mean, and Naïve applications.
(Elaborated by the author).

Model	***RMSE***	***MAE***	***MAPE***
Drift	2.74	1.80	6.26
Mean	5.67	5.17	19.20
Naive	3.12	2.10	7.36

As figure 38 presents the point forecast for the South Africa, table 45 shows the accuracy check for the three initial methods to project unemployment rates in the future for the country. Results presented on table 45 suggest a clear most adequate method between the ones assessed, the Drift. As a variation from the naïve method, drift allows variations over time (Hyndman & Athanosopoulos, 2021) and as happened for Brazilian data, it will be the one used to proceed with the STL forecast.

Using drift method combined with the decomposition from `STL` function and `hilo` to forecast highs and lows future values for predictive intervals, on the following figure 39 are presented three scenarios to South African unemployment.

Figure 39
Predicted intervals and distributional forecasts.
(Elaborated by the author).

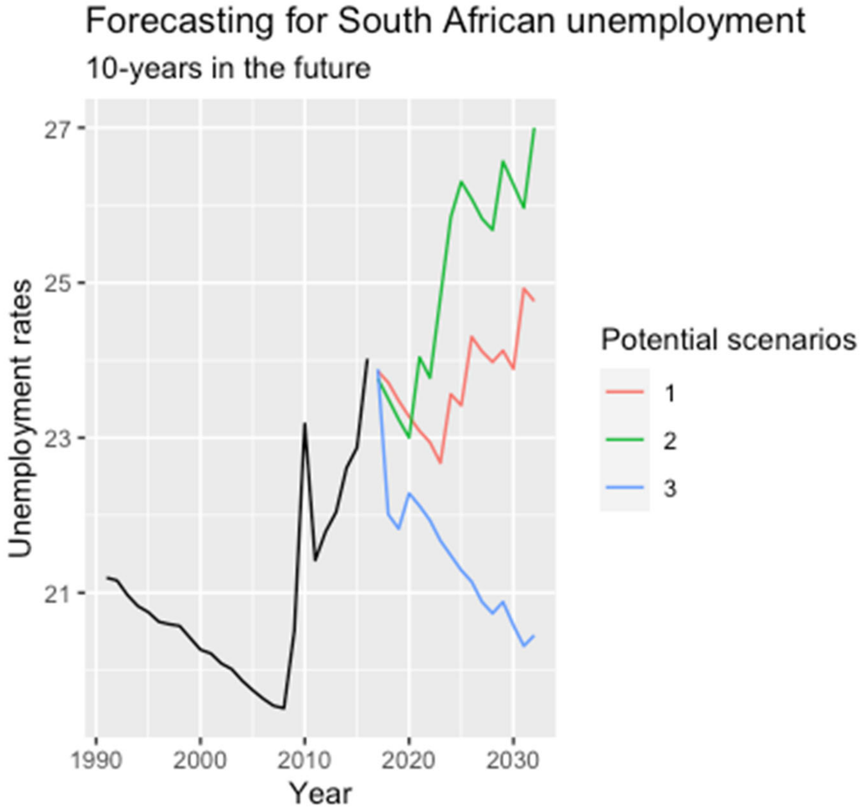


Figure 39 indeed illustrates that, generically, unemployment on South Africa tends to maintain in a high percentage. However, the scenarios projected by potential intervals of these rates could present a not extremely prejudicial path in comparison. Scenario 2 clearly projects a damaging future effect on labour market because unemployment, levels of it remaining consistently high and closing on 27% of people without a job. Scenarios 2 and 3 are, in comparison, less harmful, being the third, presented-on the blue line, the one where unemployment reaches under 21% after 2030.

There is a path to mitigate unemployment levels on South Africa. Scenarios projected could go better or worse depending on the measures adopted by the country to deal with this problem. Nkoane & Seeletse (2021) for example suggest state intervention on private companies to have a say on better allocation of unused labour force and this could be a fruitful manner indeed. Promotion of entrepreneurship activities may be another avenue to proceed with, qualifying and orienting people in propose and primordially maintain sustainable activities as we discuss on chapter 3 of this thesis could present beneficial as well. Nonetheless the policies to be applied, something beneficial must be implemented. Table 46 presents the accuracy check for South African STLF.

Table 46
Accuracy checking for the STLF forecast for Indian unemployment rates.
(Elaborated by the author).

Model	*winkler*	*percentile*	*CRPS*
STLF	23.00	1.79	1.77

STLF model using South Africa statistically presents as the most adequate according to the results illustrated on table 46, performing better than drift and the other methods used on forecasting. This result is in line with previous proceedings, where we have a better fit according to CRPS. Russian data is the exception where not necessarily the STL forecast performs better although it does not suggest a bad result for Russia but reinforces the proposition to perform other forecast techniques as well to analyse the BRICS collective, something we proceed next to close the Seasonal and Trend decomposing using Loess method.

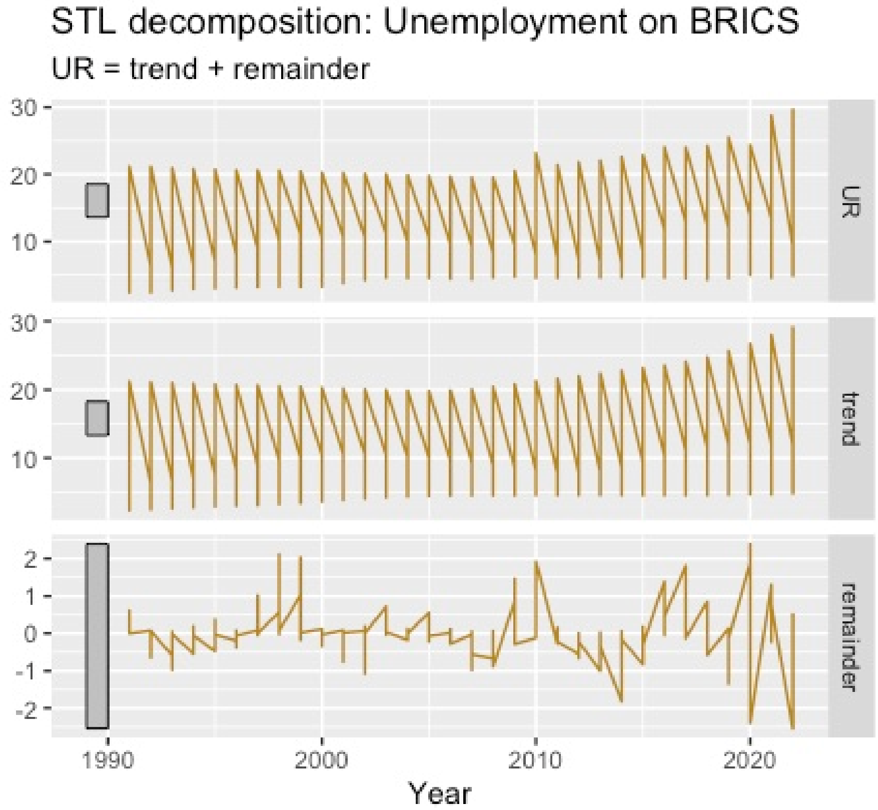
Although the steps to be performed when analysing countries within a group by STL are essentially the same as we applied for Brazil, Russia, India, China, and South Africa individually, as this is the final application using the method, we return to some characteristics about this technique to close the usage of this method. First, all countries data under analysis here are seasonally adjusted. Because unemployment rates, subject of interest here, already is usually disclosed on this format. Therefore, the "S", which stand for Seasonal, on the STL acronym are not necessarily primordial as may have noticed.

Second, BRICS countries will be analysed on the division of training data, for parameters estimative, and test data, to evaluate the accuracy of the STL. Once again, the proportional division of BRICS data will be of 80/20. Therefore, considering we have a total of 159 observations, 130 of these are for the training portion while the remaining 29 are the test dataset. Timespan covered is from 1991 up to 2022, despite the Russia having a year-short of information on collective we use the most recent data for the group.

Two types of proceedings are applied. First, we conduct a point forecast using three basic methods, Mean, Naïve and Drift, founding the most adequate of these we use this one to proceed with the STL forecast using the best of these initially assessed with the decomposition portion of the STL method to project future values for unemployment in BRICS based on predicted intervals and distribution of them through the `hilo` function. For each of these techniques we apply accuracy check to have a statistical definition for the one that performs better in comparison.

All stages are according to Hyndman & Athanasopoulos (2021) using `fable` R-Studio package. Figure 40 presents decomposition lines for initial STL application for unemployment in BRICS, assessing the trend ant the residuals for unemployment rates.

Figure 40
STL application for unemployment in BRICS.
(Elaborated by the author).



First line on figure 40 presents unemployment data fluctuations considering all the five BRICS countries, portraying the lowest levels of unemployed and China, where the rates are smallest in comparison, and the higher rates particularly present on South Africa. Trend line suggests that the tendency of this data is to somewhat replicate in the future historical information as retrieved from ILO and World Bank. Still on figure 40 we have an inconsistent line of residuals (remainder line), something that is expected given the particularities of each country participant on the group.

BRICS timeseries for unemployment rates then may be presumed by an STL equation with a trend, a residual portion and non-seasonality, expressed as: $\mathrm{BRICS\_UR}_t = T_t + R_t$. Considering the analysis of STL as proposed on Cleveland et al. (1990), we cannot envision a clear path on trend when countries are observed collectively, some highs and some lows are predicted similarly as it was already occurring in historical values available.

We statistically assess features of this STL decomposition applying `feat_stl` function. Outputs from this assessment are described on table 47 considering for each column on the table values for median, pondering all individual data by country in grouping them into the BRICS group.

Table 47
Features of STL decomposition for BRICS unemployment rates.
(Elaborated by the author).

	trend_strength	*spikiness*	*linearity*	*curvature*	*stl_e_acf1*	*stl_e_acf10*
BRICS	0.89	0.00	3.92	-1.66	0.23	0.29

As we analysed for each country, two columns are particularly important to discuss, the "trend_strength" and the "curvature" ones, on both we have some suggestion on how data may behave in predicted values for unemployment in the future of BRICS. Here we have an index by 0.89 of probability that the curvature on unemployment rates line may follow a declining pattern, by around -1.66. Meaning that exists an 89% of chance that in a foreseeable future some decline on unemployment may occur.

We further assess this potential of decline through the two-stage, point and distributional interval forecasts applied to the BRICS dataset. First step is the application of a point forecast proceeding using the three beforementioned simple forecast methods, mean, naïve and drift. Discovering which of these performs better, the most suitable will be later used on the STL forecast proceeding.

STL forecast modelling (STLF) will consider the best method of the three initially conferred and the decomposition portion from STL technique as visually presented earlier on figure 40, more specifically the components of trend and residuals, as seasonality is not present on the data. On to implement these steps described above, once again, we are following the Hyndman & Athanasopoulos (2021) textbook examples and the usage of `fable` R-Studio package.

Following table 48 presents the statistical accuracy check for the three methods, identifying which of them is the best suited to proceed with the STL forecast application.

Table 48
Accuracy checking for Drift, Mean, and Naïve applications.
(Elaborated by the author).

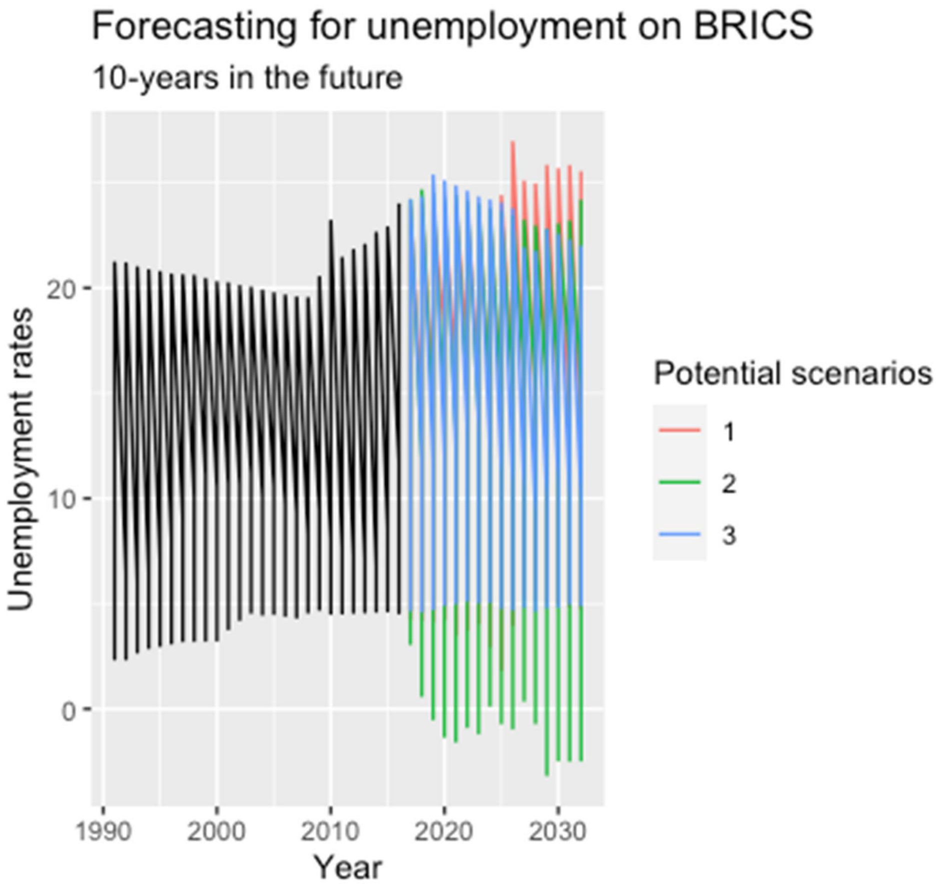
Error measurement / Model	**RMSE**	**MAE**	**MAPE**
BRAZIL			
Drift	1.63	1.17	10.60
RUSSIA			
Naive	0.70	0.60	12.60
INDIA			
Mean	1.13	0.76	9.22
CHINA			
Naive	0.25	0.19	3.92
SOUTH AFRICA			
Drift	2.74	1.80	6.26

We have mixed results as was already presumable when countries were being observed individually. Drift performs better for Brazilian and South African data; naïve method is the most suited for Russia and China whereas for the Indian dataset the model that indicates a minimisation on errors is the mean method. Mean method is simply the replication of the mean on historical data for each country; naïve is a replication as well but of the last-observed value (Hyndman & Athanosopulos, 2021).

Drift method has a slightly more complexity on it. It is a variation from the naïve method, the distinction between the two is that the former enables that projected values may fluctuate over time. These potential variations are known as the drift of the data and refers to the average change that occurred on historical data and potentially may also occur in future values (Hyndman & Athanosopoulos, 2021). We believe that this behaviour, a drift occurring, is more alike real-life data and consequently a tendency for better performance over forecasts. We proceed for the STL forecast using drift method.

Function `hilo`, converts predicted future values for unemployment into a range that goes from 80% up to 95% of confidence of these presumed values (Hyndman & Athanosopoulos, 2021). We present these predictions on the following figure 41 using potential scenarios from the potential values anticipated.

Figure 41
Predicted intervals and distributional forecasts for BRICS.
(Elaborated by the author).



To compose and present figure 41, we used the drift method and the STL decomposition to employ a STL forecast proceeding that predicts three possible scenarios for unemployment rates fluctuation on the BRICS countries. Future values forecast on this method in a general perception tends to have a similar pattern with the historical data available and retrieved from ILO and World Bank repositories. It seems that considering the 10-year in the future projected, the tendency is to repeat the data behaviour analysed over the historical timespan going from 1991 to 2022.

However, still on figure 41 it is possible to envision a better-projected scenario this being the one represented by the green line of scenario 2. In this scenario we perceive that values do not goes higher than the other lines (on red and blue) projected by scenarios 1 and 3. Scenario 1 seems to be particularly bad as values may go further 25% people outside of the potential labour force of the BRICS group after 2025. Scenario 3 are basically a repeat on historical data whereas on projection 2, best-case scenario, the minimal values may even reach the 0%, theoretically.

Following table 49 presents the accuracy of the second forecast applied using the STLF model. Diagnostic check is presented on the table similarly as the assessment showed on table 47, using median values for contemplation of data by country into a supranational proposition to BRICS.

Table 49
Accuracy checking for the STLF forecast for BRICS unemployment rates.
(Elaborated by the author).

Model	*winkler*	*percentile*	*CRPS*
STLF	7.56	0.67	0.67

Continuous Ranked Probability Score (CRPS) as presented on table 49 enable to reasonably believe that the STLF model is the one that offers most reliable values on forecast in comparison with earlier adopted methods (drift, mean and naïve). CRPS and percentile outperforms on minimising errors for the case of probabilistic forecasts (Gneiting & Katzfuss, 2014). CRPS on table 49 in specific is only higher than MAE for Russian and Chinese data only. With all presented in this topic, we believe to have offered a reasonable starting point on potential paths for unemployment in BRICS. Remains to be seen other techniques performance, continuing next with the ETS application.

4.4.2. Exponential Smoothing Technique – ETS application.

This topic covers the ETS framework onto a second technique to produce forecasting values for unemployment rates in BRICS countries following the initial STL proceeding. Here, similarly as the performed on the previous topic, each forecast will be made country by country and after these individual assessments we conclude with a BRICS collective analysis. Once again, all proceedings will be following Hyndman & Athanosopoulos (2021) textbook.

We apply STL before the Exponential Smoothing Technique because the first one offers some insight on how to proceed with the latter. There are three potential methods within the overall application of a ETS proceeding: Simple exponential smoothing, that uses weighted average, flat parametrisation, and more simplistic procedures; methods with trend; methods with seasonality; and these methods may even be combined depending on the type of data under analysis (Hyndman & Athanasopoulos, 2021).

On STL application it was identified that seasonality is not present in the working dataset of unemployment on BRICS. Neither for countries individually nor for their information grouped. Therefore, ETS with seasonality may not be adequate for our data. Simple exponential smoothing method on the other hand could work well if country's information being analysed does not show an eminent trend on it.

The trend approach has some caveats on it that are worth of clarification. First, there is the Holt's (2004) proceeding, which is most used on simple exponential method. This application presumes a constant and linear trend that goes indefinitely into the future. However, exists some empirical evidence from real-life data suggests that this procedure tend to over-forecast, especially on relatively longer forecast horizons (Hydman & Athanosopoulos, 2021). Second trend application derives from Gardner & McKenzie (1985) proposition, including a parameter into Holt's equation to dampens the linear trend presumed in a manner to flat the trend-line of the intended forecast horizon.

First step in our ETS application will be the identification of which between these methods performs better considering Brazil, Russia, India, China, and South Africa data characteristics. Having the best one discovered, we proceed with the estimation process of our ETS statistical frameworks, letting the function `ETS` from the R-Studio package `fable` to automatically identify the best suited model referring to each unemployment data being analysed.

A basic ETS model may be denoted as $ETS(A, A, A)$, where the first letter represents the error term, that could be A for additive or M for multiplicative; the second letter represents the trend term, and the third letter represents the seasonality term. All of these respecting the dataset used on the modelling procedure.

Basic equations for additive error, trend, and seasonality are as follows:

Level equation (L): $y_t = l_{t-1} + b_{t-1} + s_{t-m}$;

Trend equation (B): $b_t = \alpha \times (l_t - l_{t-1}) + (1 - \alpha) \times b_{t-1}$;

Seasonal equation (S) is: $s_t = \gamma \times (y_t - l_{t-1} - b_{t-1}) + (1 - \gamma) \times s_{t-m}$.

Where y_t is the predicted value at time t, l_t is the level, b_t is the trend and s_t is the seasonal component. α is the smoothing parameter for the trend and γ the smoothing parameter for the seasonality. The term m represents the number of seasons in a cycle if they exist. It is not our intention to proceed with calculation manually as `fable` package and `ETS` function gives the forecast values automatically, but it is worth to describe how parameters found in each model may be inserted into mathematical proceedings.

From all the above specifications we move for the opening application that is made considering unemployment rates for Brazil in the timespan ranging from 1991 to 2022. First proceeding as described earlier is the identification of which model (simple exponential smoothing - SES, Holt's linear or Holt's dampened) performs better.

Outputs for this assessment is presented on table 50.

Table 50
Accuracy from potential ETS models: Brazilian dataset.
(Elaborated by the author).

Model	*RMSE*	*MAE*	*MAPE*
Damped	5.09	4.49	35.90
Holt	8.12	7.00	57.00
SES	4.04	3.61	28.30

Despite the metric for error minimisation selected, table 50 presents that for Brazilian unemployment rates the simple exponential smoothing tends to perform better comparing with the other potential options. This is an indication that the trend on unemployment for Brazil may not be substantial into our intended 10-year forecast horizon.

Therefore, we move for ETS estimation parameters using the SES technique and these outputs are as following: Model $ETS(A, N, N)$, $\alpha = 0.99$, initial states $l(0) = 6.36$, $\sigma^2 = 1.08$. Smoothing parameter determines the weight given to the most recent observation in predicting the next value. A value by 0.99 as the one found here being close to 1 suggests that the model heavily relies on the most recent data.

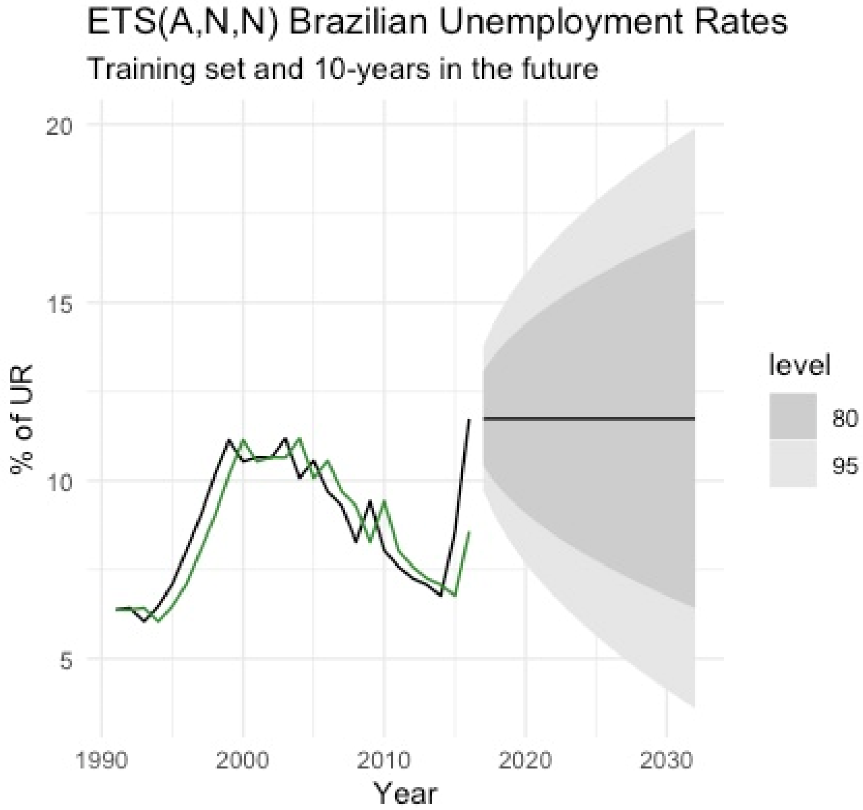
Also, σ^2, or estimated variance of the error term, on 1.08 offers an indication of the volatility of the residuals in the Brazilian ETS model. Overall, Brazilian model suggests that the unemployment series has an additive error term, which indicates that deviations from values predicted will potentially be added linearly whereas high value of α suggests a significant emphasis on the most recent observations in predicting future values.

On particular to the forecast of these potential values for unemployment rates considering the Brazilian dataset, we present these on the sequence. Some clearance on the ETS forecast proceeding before the presentation that will be applied on Brazil as well on the other countries data. Once again, we use the training portion of data, similarly as used on STL application following an 80/20 proportion of total observation for each country. We rely on that separation to later assess the accuracy on projected values.

For Brazilian and other BRICS countries analysis we start with a visual presentation of historical data behaviour with projected indexes and after that we assess these predicted values on specific comparing the projected with the test portion of data to assess accuracy. Following figure 42 presents the ETS forecast application for unemployment rates in Brazil.

As we trim the Brazilian data training portion on 2016 year, there is on figure 42 a presentation from forecasts for the period 2017-2032. Also, it is presented on the figure by the green line a plot for one-step-ahead fitted values alongside the factual data retrieved from 1991-2016. Large value of α obtained for Brazil (0.99) is reflected on the also large adjustment in estimated future values projected on figure 42. Brazilian unemployment rates, with a 95% of confidence could go lower than 5% while the high could near 20% after 2030.

Figure 42
ETS forecast application for Brazil.
(Elaborated by the author).



On specific of these predicted and projected intervals shown on figure 42, they are calculated, similar as in the STL application, through R-Studio function `hilo`. Predicted intervals shows a significant uncertainty for future values of unemployment rates over the 10-year forecasting. To assess how misleading these intervals could be, we compare the real values from 2017 to 2022 with ETS forecasted values at the same period.

Following table 51 presents this comparison of values. Ideally a good projected value would fall within the 80% range interval area as when moves up to 95% the intervals are larger enabling to more values, as presented on figure 42. All information on table 51 is considering Brazilian unemployment indexes as extracted from World Bank and ILO considering the 6-years timespan covering 2017 up to 2022.

Table 51
Comparison of real data and projected intervals: Brazilian ETS forecast.
(Elaborated by the author).

Country	Year	Real UR	Forecast UR (80%)	Forecast UR (95%)
Brazil	2017	12.90	[10.40, 13.07]	[9.70, 13.77]
	2018	12.50	[9.85, 13.62]	[8.86, 14.62]
	2019	12.00	[9.43, 14.04]	[8.21, 15.26]
	2020	13.90	[9.07, 14.40]	[7.67, 15.81]
	2021	13.30	[8.76, 14.71]	[7.18, 16.29]
	2022	9.46	[8.48, 15.00]	[6.75, 16.72]

Table 51 shows that projected values on the interval that has 80% level of confidence are within the real unemployment indexes. Meaning that we have a reasonable forecast of Brazilian unemployment rates comparing the test set and our projections. This result offers some confidence to the values projected beyond 2022, indicating that we have proposed a reasonable interval for the future of labour market in the country. CRPS and percentile indexes were around 0.88 and 0.89 respectively attesting the accuracy of the Brazilian $ETS(A, N, N)$ model, performing statistically well on minimising errors about the predicted values following this modelling.

We move forward with the Russian data from hereafter. Complete Russia unemployment rates covers the timespan ranging from 1991 to 2021. The caveat is on the training and test portion of this data, being the first from 1991 to 2016 and the second one observation short than other BRICS countries, from 2017 up to 2021. We proceed for the best model for ETS available assessing simple exponential smoothing - SES, Holt's linear and Holt's dampened. Results for this assessment is presented on table 52.

Table 52
Accuracy from potential ETS models: Russian dataset.
(Elaborated by the author).

Model	*RMSE*	*MAE*	*MAPE*
Damped	3.33	1.86	37.70
Holt	7.99	4.79	96.70
SES	1.94	1.61	33.00

RMSE, MAE and MAPE metrics for error minimisation considering the Russian dataset as presented on table 52 suggests the SES technique as the one that will potentially perform better in comparison with the other ETS possibilities. Again, as happened for Brazil, this may be an indication that any trend for unemployment in Russia could not be very influential 10-years into the future.

Presuming the SES as most suited technique, we move for ETS estimation parameters and outputs are as following: Model $ETS(M, N, N)$, $\alpha = 0.99$, initial states $l(0) = 5.30$, $\sigma^2 = 0.02$. ETS model for Russia has a multiplicative error, differently as in Brazil as the error was and additive type (model there was $ETS(A, N, N)$). Smoothing parameter for data from Russia is the same from Brazil, being as 0.99, suggesting that Russian model also relies on the most recent observation.

Estimated variance of the error term here is lower, $\sigma^2 = 0.02$, offering a potential low volatility of residuals in the Russian ETS model. This type of model, with a multiplicative error indicates that any eventual deviations from values predicted into the future will be added multiplicatively, while the high α values define a significant relevance on the most recent data of unemployment on future values projections.

We deepen on the Russian analysis with a visual presentation of historical data behaviour and projected values covering the test portion of data and beyond on the following figure 43, that presents the ETS forecast application for unemployment rates in Russia.

Figure 43
ETS forecast application for Russia.
(Elaborated by the author).

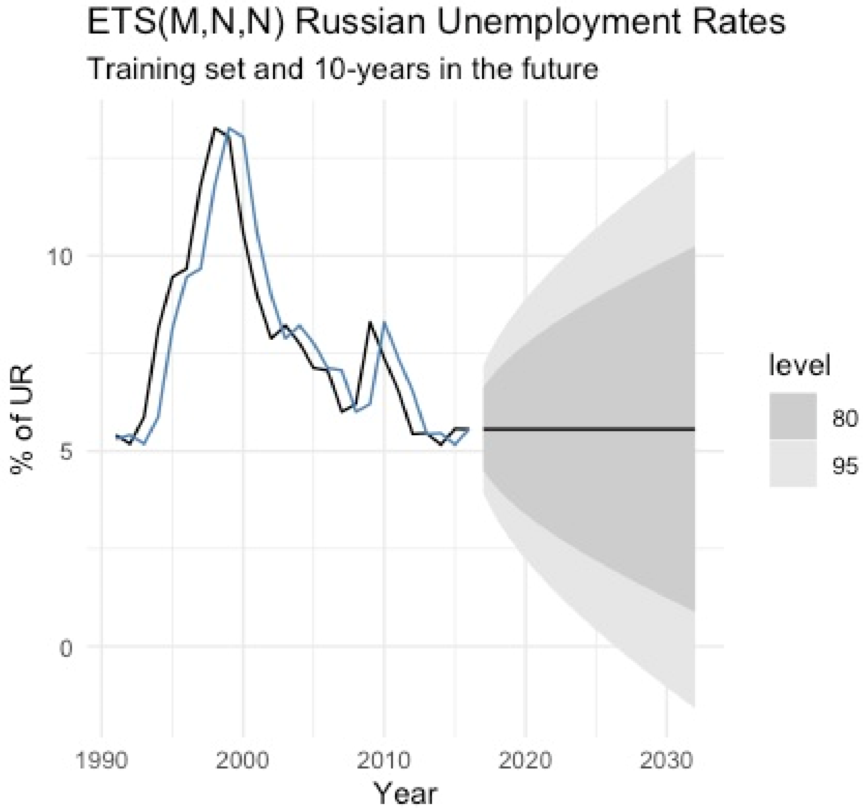


Figure 43 follows a similar presentation as the one referring to Brazilian data on figure 42. Here, the blue line on the plot as well refers to the one-step-ahead fitted values in line with data Russian unemployment rates from 1991-2016. High value of α as well is reflected on the adjustments for future values projected on 80% and 95% confidence interval. However, as unemployment in Russia in comparison with Brazil is relatively lower, statistical projection of these rates could go below 0% on the second-half of 2025 years which in real-life is not feasible.

Projected values on figure 43 indicates some high uncertainty on potential future values for unemployment rates, may be as high as 15% and as low as 0%. We assess these higher and upper estimation on the `hilo` function once again. The following presented in the table 53 presents a conference of real-data and forecasted values similar as we have done for Brazilian case, comparing the 2017-2021 timespan respecting the limitation of available information for Russia.

Table 53
Comparison of real data and projected intervals: Russian ETS forecast.
(Elaborated by the author).

Country	Year	Real UR	Forecast UR (80%)	Forecast UR (95%)
Russia	2017	5.21	[4.48, 6.63]	[3.92, 7.20]
	2018	4.85	[4.03, 7.08]	[3.22, 7.89]
	2019	4.50	[3.68, 7.43]	[2.68, 8.43]

	2020	5.59	[3.38, 7.74]	[2.22, 8.89]
	2021	4.72	[3.10, 8.01]	[1.81, 9.30]

Estimated values considering the confidence interval by 80% covers well the real data of unemployment rates in Russia. Therefore, it is reasonable to assume that we have in hand a fitted forecasted values for Russian data. Complementing the figure 43 when analysing the intervals to compose table 53, we perceive those values under 0% are all on the 95% level of projection and starting on 2027.

On opposition, in 80% level of projections values close to 0% but does not reach it up to 2032, minimum value projected is on 0.88 in this year. Statistical accuracy checks through CRPS, and percentile indexes found around 0.46 and 0.47 respectively attests a better accuracy for Russian $ETS(M,N,N)$ model, performing satisfactorily well on minimising errors, improving the SES technique, and outperforming Holt's linear model and its variation with the damping parameter.

Moving forward, Indian unemployment rates again have all possible data available, meaning 32 observations from 1991 to 2022, similar as Brazil and the future assessed Chinese and South African dataset and differently from Russian data. Considering this complete range of data, we check for best model to proceed with ETS. Table 54 presents this conference for Indian application.

Table 54
Accuracy from potential ETS models: Indian dataset.
(Elaborated by the author).

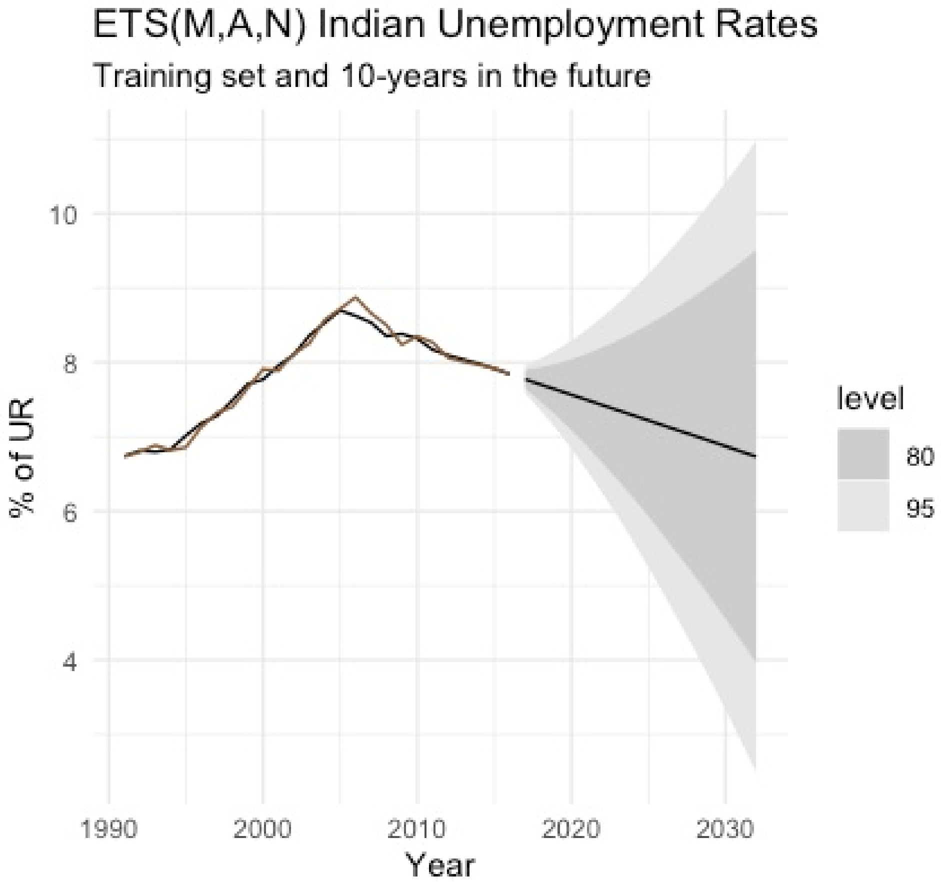
Model	***RMSE***	***MAE***	***MAPE***
Damped	1.79	1.38	18.10
Holt	2.04	1.50	19.40
SES	1.21	1.03	13.20

Once again, simple exponential smoothing seems to be most adequate. It happens in the third of three assessed countries. Four metrics for error minimisation presented on table 54 indicates SES for a better performance. Simplicity it seems, at least so far, very useful on our type of data. On ETS model estimation, the `ETS` function presents the following parameters: First, the model is different from Brazil and Russia, being an $ETS(M,A,N)$. The "A" on the second term suggests a potential trend on Indian data.

Other parameters found by `ETS` beyond the model type are the $\alpha = 0.99$ and a $\beta = 0.56$. Beta in this case is present because the trend on Indian model, different from Brazil and Russia where no clear trend is identified. Estimated smoothing coefficient on β represents the slope on the available information of India. A slope by 0.56 is relatively large, which may suggest that Indian tendencies could change rapidly even slightly (Hyndman & Athanasopoulos, 2021). To acknowledgment initial states are $l(0) = 6.68\ and\ b(0) = 0.05, \sigma^2 = 0.002$.

Indian historical data from the beginning of the timespan covered up to the end of training dataset and projected values by ETS forecast application for unemployment rates is presented on the following figure 44.

Figure 44
ETS forecast application for India.
(Elaborated by the author).



Forecasted values for India appears to have a sensible and decreasing trend as presented on figure 44. However, this trend as we inferred earlier is related with a wide prediction interval, which in a manner reflects well the variation in Indian historical values. As well this corroborates the high value of α and the presence of a slope in the data as suggested by $\beta = 0.56$.

This data behaviour could be a suggestion that a damped parameter inclusion could improve the forecast with data from India, however as the `ETS` usage suggested SES to proceed, as presented on table 54, we decide for the simple exponential smoothing, and we check if this was indeed a reliable indication or not moving forward with our higher and upper estimation presented on figure 44 and illustrated on the following presented table 55.

Table 55
Comparison of real data and projected intervals: Indian ETS forecast.
(Elaborated by the author).

Country	Year	Real UR	Forecast UR (80%)	Forecast UR (95%)
India	2017	7.73	[7.65, 7.89]	[7.58, 7.96]
	2018	7.65	[7.47, 7.93]	[7.35, 8.05]
	2019	6.51	[7.28, 7.98]	[7.10, 8.16]

	2020	10.19	[7.09, 8.04]	[6.83, 8.29]
	2021	7.71	[6.87, 8.11]	[6.55, 8.44]
	2022	7.33	[6.65, 8.20]	[6.24, 8.60]

In general, the estimated values by 80% of level of confidence suits well the real-life data for unemployment in India. There is however an outlier on 2019, where from World Bank and ILO Indian percentage of people outside of labour force was on 6.51% while the lowering side of the projection goes by 7.28% in the same year. This had some form of compensation on the following year, given that by 2020 real unemployment was on 10.19% which surpasses even the higher projection on 95% column (8.29% in that year).

Indian unemployment has indeed some peaks and lows on unemployment as we already noticed analysing the country historical data. Nonetheless, even considering these two potential outliers on the data, we have most projections falling within the range of probabilities of 80% level when comparing estimation and real data for unemployment in India. Statistical accuracy checking by CRPS and percentile being around 0.63 gives some confidence to believe that Indian $ETS(M, A, N)$ model performs well on forecasting.

Next country to be analysed is China in their unemployment covering the 1991 to 2022 timespan. Chinese complete dataset as well goes by 32 observations, with no missing information within this sample. From these values we proceed to identify best suited model between Holt's linear, Holt's damped and SES to use on ETS forecasts. Table 56 presents this initial assessment.

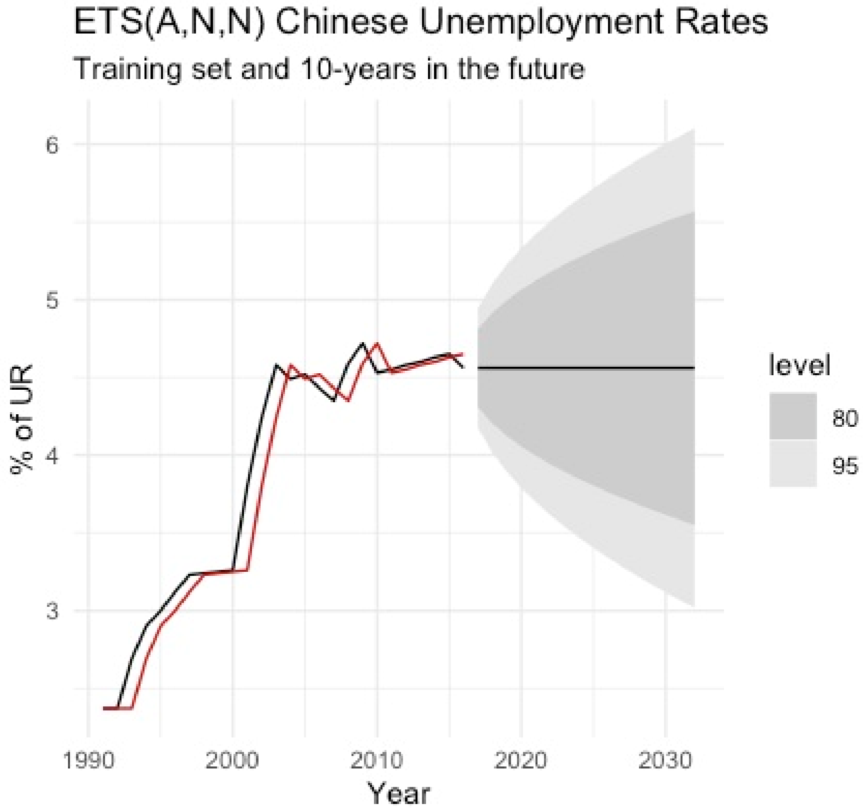
Table 56
Accuracy from potential ETS models: Chinese dataset.
(Elaborated by the author).

Model	*RMSE*	*MAE*	*MAPE*
Damped	0.82	0.63	14.00
Holt	1.36	1.14	25.10
SES	0.27	0.21	4.48

Table 56 presents that the simple exponential smoothing - SES again outperforms on minimising errors both Holt's procedures. Numbers on the Damped proceeding is not far from SES but we one more time move with the simple exponential method. About ETS model coefficient estimation, `ETS` function gives these parameters: Chinese model is an $ETS(A, N, N)$, the same that was found on Brazilian data. $\alpha = 0.99$ there is no suggestion of a trend as happened on India, the initial states are $l(0) = 2.37$ and error variance is on $\sigma^2 = 0.03$.

Visual presentation of Chinese historical data is presented next on the figure 45. Initial values up to the trimming point of the training portion of data is showed with fitted values in the red line and projected intervals forecasted values are described in the right-side of the figure.

Figure 45
ETS forecast application for China.
(Elaborated by the author).



Chinese presentation on figure 45 is similar with Brazil and Russia. Here, the fitted values of unemployment on the red line reflects well the historical real-data and presents a slight crescendo at the end of the data on the year 2016. Chinese model having a high value of α is confirmed on the intervals projected for unemployment into the future. People unemployed on China could range at above 6% after 2030 and a minimum around 3 in this same period, suggesting some level of uncertainty on the projections.

These projections on the interval range from 80% and 95% of confidence is assessed by the `hilo` function. One more time we compare the real-life data available on test set and the upper and higher values forecasted. Table 57 presents these comparisons replicating what we apply for the Brazilian, Russian, and Indian data information.

Table 57
Comparison of real data and projected intervals: Chinese ETS forecast.
(Elaborated by the author).

Country	Year	Real UR	Forecast UR (80%)	Forecast UR (95%)
China	2017	4.47	[4.31, 4.81]	[4.17, 4.94]
	2018	4.31	[4.20, 4.91]	[4.01, 5.10]
	2019	4.56	[4.12, 4.97]	[3.89, 5.22]
	2020	5.00	[4.05, 5.06]	[3.79, 5.33]
	2021	4.55	[3.99, 5.12]	[3.70, 5.42]
	2022	4.88	[3.94, 5.17]	[3.61, 5.50]

Estimated values considering the confidence interval by 80% covers all the real unemployment rates for China over the timespan from 2017 up to 2022. Henceforward, it is quite realistic then to presume that our projected values for Chinese unemployment levels in the future are reasonable to be believed. Accuracy conference of the $ETS(A, N, N)$ Chinese model confirms that these projections are indeed statistically reasonable, as we have values for CRPS, and percentile very low in comparison, being of 0.14 index for both measurements. Meaning that the potential errors of estimated values are reasonably low in practical ways.

Fifth of the five BRICS countries to be checked individually is South Africa. One more time, African dataset comprises 32 observations that ranges from 1991 to 2022 of available information of country's unemployment rates. Assessment of which technique (simple exponential smoothing - SES, Holt's linear or Holt's dampened) performs better is made and presented on the following table 58.

Table 58
Accuracy from potential ETS models: South African dataset.
(Elaborated by the author).

Model	*RMSE*	*MAE*	*MAPE*
Damped	7.86	6.31	23.90
Holt	8.49	6.58	25.00
SES	5.27	4.63	17.40

RMSE, MAE and MAPE indexes presented on table 58 as well as happened for the other four BRICS countries suggests that simple exponential smoothing technique outperforms the other possible formats to proceed with our ETS application. Results may be an indicative that South African unemployment rates does not have a clear tendency for the future, or if exists, may not be a substantial one considering the intended 10-year forecast horizon.

Moving for ETS estimation parameters using the SES technique, we let `ETS` function from `fable` package to fit the best suited model and these are outputs obtained: Model $ETS(M, N, N)$, as happened for Russia, an $\alpha = 0.96,$ and an initial state by $l(0) = 21.17$. This confirms the significantly higher values of unemployment in South Africa in comparison with other countries from BRICS. The estimated variance of the error term, which here is a multiplicative type, is of $\sigma^2 = 0.001$.

Smoothing parameter presented on α is still a high number by 0.96 but when comparing with the other countries analysed is the lowest. Meaning that the most recent observation has a significant weight on the smoothing process but the historical of elevated unemployment in South Africa carries some importance as well. Residuals on the other hand are not so volatile, given the low estimated values by sigma being equal to 0.001.

With these parameters and analysis in hand we proceed for the visual presentation of South African historical data behaviour on regard to unemployment and the projected values that may occur in the future. Following figure 46 presents the training portion of data from south Africa and the ETS forecast application for the country's unemployment.

Figure 46
ETS forecast application for South Africa.
(Elaborated by the author).

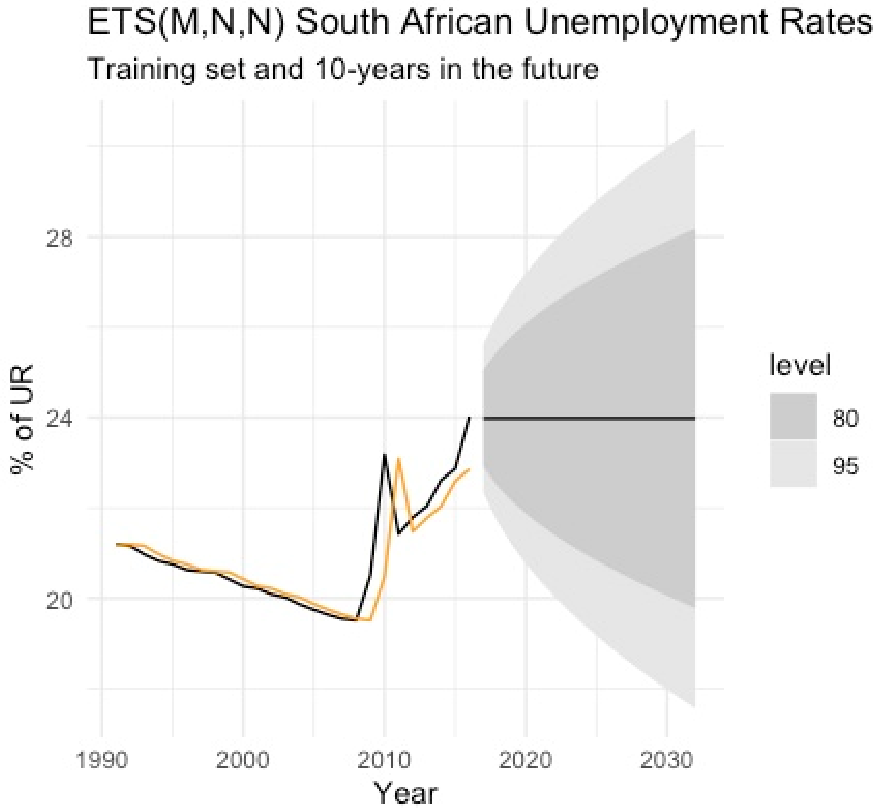


Figure 46 illustrates how already high was unemployment in South Africa at the beginning of our working dataset and how these rates growth on the second half of 2000s years and even more after that. The signs of the projected values after 2016 as well may be portraying a worrisome future. Value of α obtained by 0.96 is reflected on the potential adjustments in estimated future values projected and presented on figure 46. South African unemployment rates could go as high of 30% and in a better prospective scenario as low as 17% after 2030.

About these predicted values on the projected intervals shown on figure 46 we used the function `hilo` on the calculation of it. Projected intervals suggest a level of uncertainty in the future values for unemployment rates over the 10-year forecast horizon. To assess it we compare real values from South Africa test dataset from 2017 to 2022 and the ETS forecasted values in the same period. The following depicted table 59 presents this comparison.

Table 59
Comparison of real data and projected intervals: South African ETS forecast.
(Elaborated by the author).

Country	Year	Real UR	Forecast UR (80%)	Forecast UR (95%)
South Africa	2017	23.99	[22.90, 25.05]	[22.33, 25.63]
	2018	24.22	[22.48, 25.48]	[21.68, 26.27]
	2019	25.54	[22.15, 25.80]	[21.18, 26.77]

2020	24.34	[21.87, 26.08]	[20.76, 27.20]
2021	28.77	[21.63, 26.33]	[20.38, 27.57]
2022	29.80	[21.40, 26.55]	[20.04, 27.91]

Results on table 59 presents that from 2017 to 2020, real-life unemployment rates fall within the projected values that have 80% level of confidence. However, in 2021 and 2022 the real values are not covered even on the 95% level of projections, considering that the unemployment on these years were above the expected. Probably this came as consequence from COVD-19 pandemic effect, where unemployment levels have grown in South Africa and all over the world. This spike complies with the alfa smoothing parameter found of 0.96, suggesting that these most recent values are relevant but historical data have a significance as well.

Overall, our results suggest a feasible fit for our forecasted values. When the numbers are not covered by the projection, the slightly lower on the smoothing parameter obtained on α somewhat compensates. For the values projected beyond 2022 we have some suggestion to believe that projected results will range well what could happen in real-life and offering some reliable insights on the theme. Statistical assessment by CRPS (1.72) and percentile (1.74) indexes attests a good accuracy of the South African $ETS(M, N, N)$ model, minimising well the potential errors on predicted values.

We replicate the ETS modelling similar as applied for Brazil, Russia, India, China, and South Africa individually to assess the complete BRICS data collectively. As for the countries' case, we use the proportional division of data by 80/20. 130 observations are on the training portion and the remaining 29 are the test dataset. Timespan covered is from 1991 up to 2022, exception is in Russia' case that goes from 1991 to 2021.

We check for the best suited technique from Simple Exponential Smoothing - SES, Holt's linear and Holt's damped proceeding to continue with ETS method (Gardner & McKenzie, 1985; Holt, 2004; Hyndman & Athanasopoulos, 2021). For all countries we have discovered that SES performs better than the other possibilities when comparing them, minimising errors on potential predicted values by ETS. This may suggest that recent trends on unemployment does not necessarily will be carried into future values.

Considering that Holt's linear method assumes the premise of values in the future continuing the trend from historical data, growing or decaying indefinitely; and the damped parameter is inducted to flat this continuous linear pattern (Hyndman & Athanasopoulos, 2021), it seems acceptable that recent values of unemployment, that are largely influenced by the COVID-19 effects implicating on jobs losses all around the world, does not necessarily will be carried into any future prospection.

We presume that on analysing BRICS, SES technique will be the most suited as well, considering that when error measurement was assessed for all countries individually this was unanimously indicated as the best one to proceed on all checking's for RMSE, MAE and MAPE, suggesting that simple exponential smoothing technique is the one with lowest indexes implying that when forecasting using SES eventual errors on predictions will be minimised. Table 60 presents our country-by-country metrics.

Table 60
Accuracy from ETS models: BRICS dataset.
(Elaborated by the author).

Error measurement / Model	RMSE	MAE	MAPE
BRAZIL			
SES	4.04	3.61	28.30
RUSSIA			
SES	1.94	1.61	33.00
INDIA			
SES	1.21	1.03	13.20
CHINA			
SES	0.27	0.21	4.48
SOUTH AFRICA			
SES	5.27	4.63	17.40

SES may be the best suited technique to proceed, however, each member of the BRICS has its own ETS model. We apply `ets` function from `fable` R-Studio package to automatically identify this and results are summarized on the table 61:

Table 61
ETS parameters and model estimation: BRICS dataset.
(Elaborated by the author).

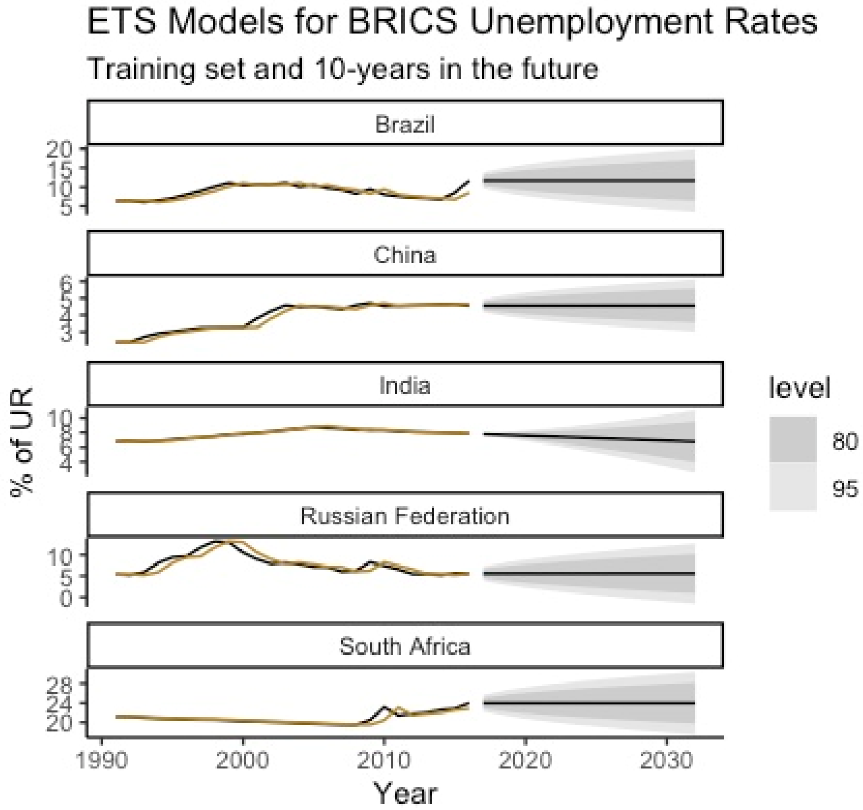
Country	*Model*	*Sigma2 (σ^2)*	*AIC*	*AICc*	*BIC*
BRAZIL	ETS(A, N, N)	1.08	90.60	91.70	94.40
RUSSIA	ETS(M, N, N)	0.02	94.20	95.30	98.00
INDIA	ETS(M, A, N)	0.0001	-31.00	-28.00	-24.70
CHINA	ETS(A, N, N)	0.39	4.100	5.19	7.87
SOUTH AFRICA	ETS(M, N, N)	0.001	72.30	73.40	76.10

Table 61 presents that For Brazil and China, we have the $ETS(A, N, N)$; for Russia and South Africa the $ETS(M, N, N)$ whereas for India, ideal model would be the $ETS(M, A, N)$. Table indicates also their own estimated error variance, presented by sigma and AIC, AICc and BIC statistical indexes, the ones that indicates the models with best potential performance.

Assuming SES technique as orientation and using the `hilo` function to predict and project estimated values for unemployment in a 10-year forecast horizon for BRICS countries, the following figure 47 presents the complete analysis of our ETS application. The following presented at the figure 47 comprises the real historical data of unemployment rates for each of the five BRICS countries and on the golden line respective to every nation their fitted values.

At the end of the year that composes the training set portion of data we move for the predicted values that this unemployment may evolve in the future, considering projection within 80% and 95% of confidence level for the forecasts.

Figure 47
ETS forecast application for BRICS countries.
(Elaborated by the author).



Comparisons could be made for test portion of real data and forecasted values and when data is not available, from 2022 and beyond we believe to have offered a solid projection for unemployment. We assess for the accuracy of our Exponential Smoothing forecast through statistical measurements that are presented on the following table 62. To comparison purposes, we present the ETS results and the STL forecast results as well, to assess which of these two implemented so far, using the same data, performs better.

Table 62
Accuracy checking for the ETS and STL forecasts: BRICS unemployment rates.
(Elaborated by the author).

Country	*Forecast Model*	*winkler*	*Percentile*	*CRPS*	*Forecast Model*	*winkler*	*Percentile*	*CRPS*
BRAZIL		7.21	0.93	0.92	**ETS (A, N, N)**	7.35	0.89	0.88
RUSSIA		7.77	0.55	0.54	**ETS (M, N, N)**	5.57	0.47	0.46
INDIA	**STLF**	13.40	0.66	0.66	**ETS (M, A, N)**	18.00	0.63	0.63
CHINA		1.19	0.14	0.14	**ETS (A, N, N)**	1.39	0.14	0.14
SOUTH AFRICA		23.00	1.79	1.77	**ETS (M, N, N)**	26.40	1.74	1.72

Some interesting outcomes on country analysis are presented on table 62. In a broader sense, we have some results that led us to believe that when comparing STL and ETS forecasts, the second one has a better performance. For South Africa and Russia, ETS minimises errors on all measurements (winkler score, percentile and CRPS). Percentile and CRPS indexes are best on ETS for Brazil and India whereas Winkler score is better on STL forecast for these two countries and for China as well.

Diagnostic checks for median values when observing the BRICS collective is slightly better on Exponential Smoothing method considering a winkler = 7.35, percentile and CRPS = 0.63; while in Seasonal and Trending decomposition using Loess method the values are on 7.72 for winkler and 0.67 for both percentile and CRPS scores. The better performance of one method in comparison with the other even though it exists, is not abnormal as well not unanimous for all countries.

Results obtained and compared offers some suggestion to continue with our intention on the assessment of other methods, beyond STL and ETS, as well onto an eventual combination of these techniques. When all assessments are performed, the objective is to have a factual best fitted modelling from its results propose more reliable and solid scenario for future unemployment in BRICS countries. Next forecast to be executed is the ARIMA modelling, presented on the next topic.

4.4.3. Autoregressive Integrated Moving Average – ARIMA application.

Autoregressive models using moving averages, as the ARIMAs, tend to complement and even improve predictive power of forecasts given that directly uses past forecast errors in a regression-like model, rather than the absolute past values of this same target variable. Comparing with exponential smoothing models, these are based on a description of trends and, if exists, seasonality in the data while in ARIMA modelling the purpose is to describe the autocorrelations in the data (Hyndman & Athanasopoulos, 2021).

Some important concepts for clarification before proceeding with the ARIMA applications by BRICS countries, namely: stationarity and differencing. A stationary timeseries is one whose properties do not depend on the time at which the series is observed, meaning that there is no trend or seasonality largely affecting the values in this series (Hydnan & Athanasopoulos, 2021).

We assessed stationarity earlier and considering our working dataset we may presume that our data follows a stationary behaviour. Or it has, at least, a unit root existing in it considering the variable of analysis (unemployment) and the timespan covered. It is reasonable to presume that unemployment rate on a time t should be influenced by what occurred at time $t - 1$ and so on.

In theory, labour market effects perdures, for better or worse, depending on economic scenario, public policies, and many other factors exerting some form of influence on unemployment levels. Understanding about how much or if any of this influence affects unemployment though is much more complex than analysing the unemployment variable itself on isolation as we intend to do on this chapter.

Therefore, even we had previously statistical evidence to suggest our data as stationary, we do not disregard that non-stationarity may be the case. In fact, not all data have this stationary behaviour. In some cases, the trend and seasonality's within the data are not so relevant than a p-value number may suggest. For these situations, differencing process comes as a solution to, if necessary, turn a non-stationary timeseries into a stationary one computing the first differences (Hyndman & Athanosopoulos, 2021).

Differencing can help stabilise the mean of a time series by removing changes in the level of a timeseries, eliminating or reducing trend and seasonality on the data being analysed. To identification of stationarity and, consequently, to notice if differencing is necessary to be performed in a dataset, time plot of this data in the manners of ACF and PACF plots are useful to visually perceive the pattern of data over a timespan.

For stationary timeseries, as we have a first suggestion is the case for the data we are working with, ACF plot will drop to zero relatively quickly while the ACF of non-stationary tends to present a slow decreasing behaviour. Nonetheless, we decide to not completely rely on p-values found and presented at the beginning of this section, even they are suggesting stationarity on our timeseries we use the visual presentation to have a more complete perception of the data beyond p-values. If the visual indicates the opposite, we perform differencing before select an ARIMA model and to forecasts.

Considering all the above explained, for each Brazil, Russia, India, China, and South Africa analyses, we first present the timeseries of unemployment behaviour and its respective ACF and PACF plots. Evidently, visual presentation on isolation may be misleading, therefore, we use them as a first suggestion for the ideal parameters to our ARIMA models and complement the analysis using `ARIMA` function available on the previously referenced `fable` package on software R-Studio.

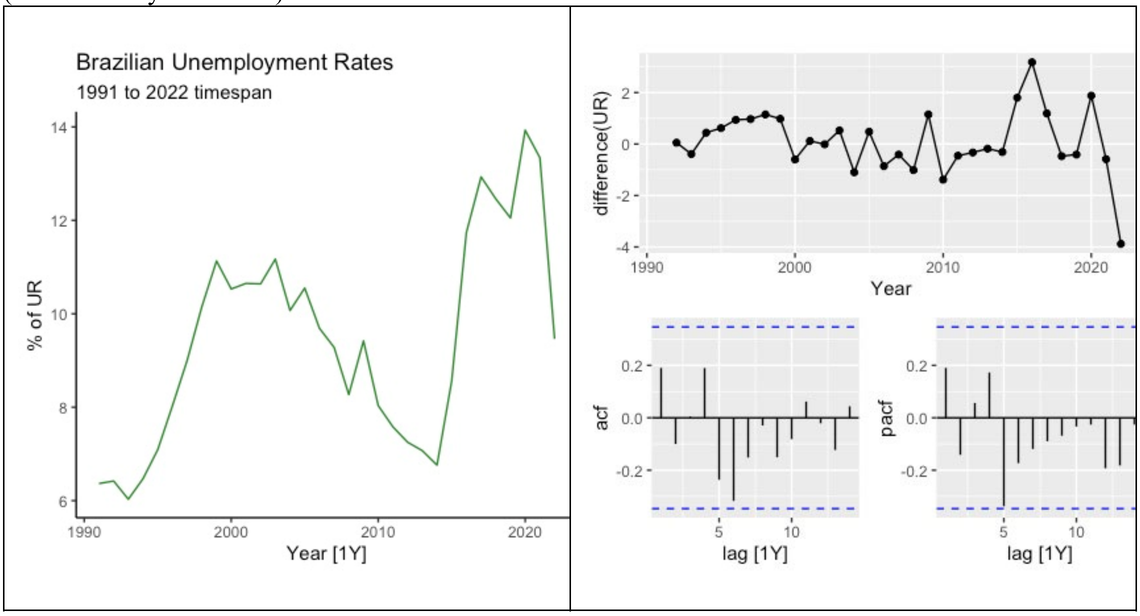
`ARIMA` function enables the identification of appropriate parameter of (p, d, q), classical form of ARIMA models. With the parameters selected and the visual presentation we may select the 'best' fitted model which is obtained by the lowest AIC indexes for the observer dataset. Having the fittest model in hand it is possible to proceed for forecasts on our intended 10-year in the future horizons.

Still about the `ARIMA` function usage, we follow the steps proposed on Hyndman & Athanasopoulos (2021) and the default application of Hyndman & Khandakar (2008) algorithm, combining unit root tests, minimisation of the AICc and MLE to retrieve the most suited ARIMA model considering the unemployment information for Brazil, Russia, India, China, and South Africa in their respective timespan of data available.

We are complying with the seven steps proceedings by Hyndman & Athanasopoulos (2021) textbook are followed and we recommend the checking of this book's chapter 9 for further information. As our first proceeding of analysis is visual, we use the complete dataset considering full timespan of data available for Brazil, comprising unemployment rates from 1991 to 2022.

Following figure 48 presents the pattern of unemployment variable in the first assessed country: Brazil.

Figure 48
Brazilian unemployment rates, differenced data, ACF and PACF plots.
(Elaborated by the author).



As we did it for STL and ETS applications, we proceed from hereafter analysing individually each of the five BRICS countries in sequence to close out this topic performing an assessment of the collective of countries as a group. First analysis on Brazilian dataset considers as starting point the above-presented figure 48. Important to clarify that next proceedings will be, as in the other methods, using the training portion from the complete dataset, meaning that following the 80/20 proportion the training set is Brazilian unemployment rates from 1991 to 2016 and test set comprises 2017 to 2022.

Right-side of the figure 48 presents real unemployment rates for Brazil and it appears to show some level of non-stationarity with a tendency on growing on the mid of 2010s years, peaking in 2020 and declining after. Presuming this non-stationarity, we proceed with a first differentiation of unemployment data, that now presents the pattern presented on the left-side within the same figure and still considering the complete dataset (1991 up to 2022).

If presuming that a differencing procedure need to be performed, Brazilian data is set to be transformed into a stationary type. Still observing figure 48, PACF plot is suggestive of an AR(6) model, when the tails starts to decay; meaning an initial potential model an ARIMA(6,1,0). The ACF suggests an MA(2) model; suggesting an alternative candidate an ARIMA(0,1,2).

With these two visual indicatives, we test models for both ARIMA(6,1,0) and ARIMA(0,1,2) along with two automated `ARIMA` function model selections, one that uses the function default, the stepwise procedure, and one that searches for a larger model (we name this as "search" model).

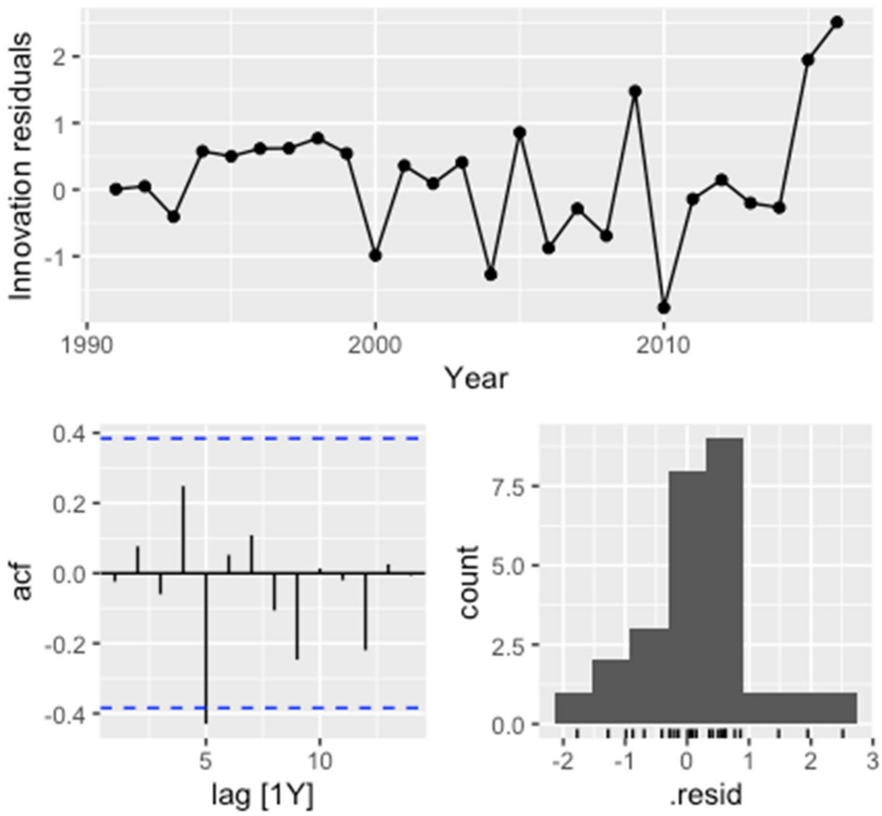
Outputs for all these testing are presented next on table 63. Remembering that at this point we are using the training portion of Brazilian data, contemplating 26 observations from 1991 to 2016 about unemployment rates.

Table 63
Fitting the best ARIMA model: Brazilian dataset.
(Elaborated by the author).

Country	***Model test***	***Orders***	***sigma2 (σ^2)***	***AIC***	***AICc***	***BIC***
BRAZIL	arima610	ARIMA(6,1,0)	0.81	75.30	81.90	83.90
	arima012	ARIMA(0,1,2)	1.00	75.10	76.20	78.70
	stepwise	ARIMA(3,0,0) w/mean	0.85	78.10	81.10	84.40
	search	ARIMA(3,0,0) w/mean	0.85	78.10	81.10	84.40

Suggestions inferred from visual presentations on figure 48 are for two models that could be fitted by the `ARIMA` function. ARIMA(0,1,2) however, is the one presenting the best fit, with better indexes on all indicators the AIC, AICc and BIC parameters. Stepwise and search models have identical values on all criteria presented in table 63 and performs better than ARIMA(6,1,0) but the model with no AR terms and two MA remains the one with a better fit in comparison.

Figure 49
ACF plot of the residuals from the Brazilian ARIMA(0,1,2) model.
(Elaborated by the author).



Therefore, we assume as best model for Brazilian data an ARIMA(0,1,2). Using this best fitted model, we produce a new ACF plot, considering residuals of our ARIMA(0,1,2). This presentation is on above-presented figure 49 that plots ACF and residuals illustration from the ARIMA(0,1,2).

Plots presented on figure 49 seems to indicate that all autocorrelations are well contained inside the threshold limits of ACF plot while residuals appear to be behaving like white noise. Residuals being white noise suggests there is indeed autocorrelation on the data, which is primordial for ARIMA method. Also, indicates the data is homoscedastic, and usually, but not necessarily, normally distributed (Huang, 2015). Box–Pierce and Ljung–Box test statistics, sometimes known as ‘portmanteau’ tests (Box & Pierce, 1970; Ljung & Box, 1978), are performed for examination of the white noise presumption. Results for both tests are presented next on table 64.

Table 64
White noise tests on the residuals ARIMA(0,1,2) model for Brazilian unemployment rates.
(Elaborated by the author).

Country	*Model*	*Test*	*p-value*
BRAZIL	ARIMA(0,1,2)	Ljung-Box	0.89
		Box-Pierce	0.90

Large p-values showed on table 64, way above the usual statistic threshold of 0.05, for both Box-Pierce and Ljung-Box tests enables the rejection of null hypothesis on both tests. Therefore, we have statistical evidence along with the visually presented on figure 49, to believe that residuals of this model, have a white noise behaviour. Meaning that our ARIMA(0,1,2) Brazilian model for unemployment rates is well suited. Before the forecasts, following table 65 presents the parameters of the Brazilian ARIMA(0,1,2) model. These parameters are relevant to compose the basic ARIMA equation.

Table 65
Parameter statistics of model ARIMA(0,1,2) for Brazilian unemployment rates.
(Elaborated by the author).

Model	*Parameters*	*Estimate*	*Standard error*	*Statistic*	*p-value*
ARIMA(0,1,2)	MA1	0.37	0.27	1.35	0.19
	MA2	0.24	0.18	1.31	0.20

Results from table 65 that considers the ARIMA(0,1,2) model for Brazilian unemployment rates, indicates an equation that may be presented as the following and accordingly with the basis equation presented on Hyndman & Athanasopoulos (2021) for ARIMA modelling:

$$Brazilian\_UR_t = \varepsilon_t + 0.37_{t-1} + 0.24_{t-2}$$

Without overextending on mathematical explanations, our ARIMA(0,1,2) model equation suggests that the current value of unemployment in Brazilian timeseries is influenced by its own past values up to 2 lags of moving averages. Coefficients for each MA terms indicate the strength and direction of the influence by each lag while the constant term represents the baseline level of this timeseries. White noise error term (ε_t) captures the variability in the data that is not explained by the lagged values.

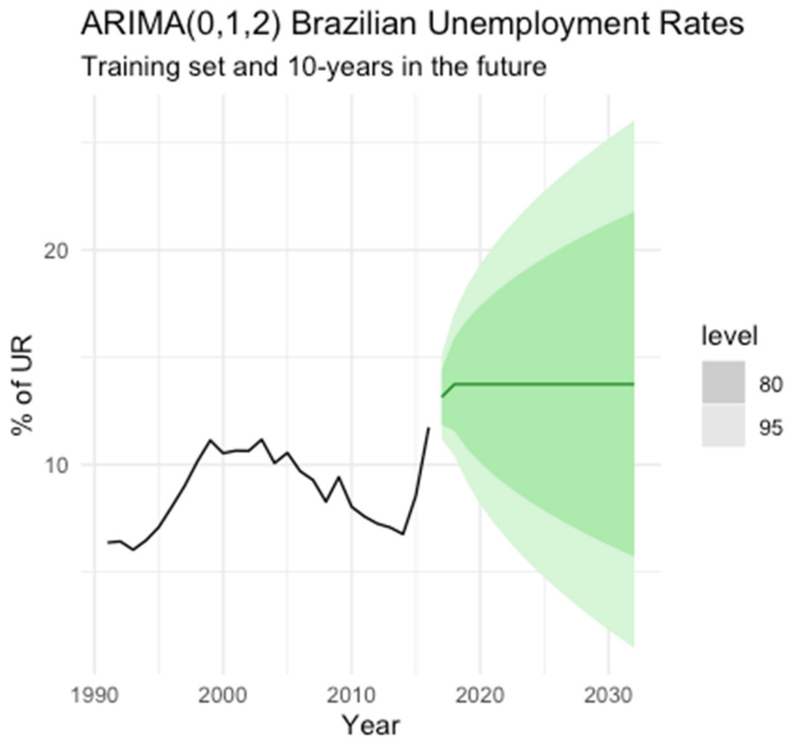
With all these assessments taken into consideration up to this point, we may finally proceed for the forecasts for Brazilian unemployment rates in a 10-year into the future horizon. Once again, as in STL and ETS application, forecasts are produced considering the training portion of the data to compare predicted values with the available real information at disposal on the test set. Therefore, we have a number-based comparison and a visual presentation. Number comparisons are presented on the following table 66.

Table 66
Forecasting Brazilian unemployment rates with ARIMA(0,1,2) model: Values comparison.
(Elaborated by the author).

Country	*Years*	*Real UR*	*ARIMA projected UR distribution*	*ARIMA UR by mean*
BRAZIL	2017	12.93	(13, ±1.00)	13.10
	2018	12.46	(14, ±2.90)	13.70
	2019	12.05	(14, ±5.50)	13.70
	2020	13.93	(14, ±8.10)	13.70
	2021	13.34	(14, ±11.00)	13.70
	2022	9.46	(14, ±13.00)	13.70

Table 66 presents a reasonable projection of values comparing with the real unemployment rates on the 2017 to 2022 years timespan of data. When projecting by mean, values projected through ARIMA application are relatively close with the real unemployment rates in a broader sense exception made for 2022, where the mean of 13.70 is way above unemployment that occurred in that year which was by 9.46.

Figure 50
ARIMA(0,1,2) forecast for Brazilian unemployment rates.
(Elaborated by the author).



Comparing projected mean values presented on the table 66, largest discrepancy between forecast and real data is indeed in 2022, overestimating unemployment by around 4%. This happened due some political modifications made on Brazil regarding to unemployment measurements by IBGE (*Instituto Brasileiro de Geografia e Estatística*), where the previous on-charge President intended to report a low number of people without a job due political interests on an electoral year in the country.

Visual presentation of the potential future of unemployment in Brazil is on the above-presented figure 50, comprising training portion of data and after the cutting point (2016 year), all projected values and intervals. Brazilian unemployment as presented on figure 50 could go as low as under 5% after 2025 year and high as above 25% after 2030. It appears that the unemployment curve in the future may repeat historical data considering the interval range projected on 80 and 95 percentage of confidence. Situation could go poorly peaking high above 20% as well a very good scenario may suggest an under the median historical unemployment.

Overall, number of Brazilians without a job could be relatively stable if public policies and labour efforts are taken to practice diminishing the potential growing to maintain under 10% of people outside the labour force. Statistical accuracy for ARIMA(0,1,2) model is the following: Winkler score = 9.67, percentile = 1.03 and CRPS = 1.02.

Moving on for the next country from the BRICS group, we proceed with Russian dataset analyses. For the behaviour of unemployment variable in Russia we use the complete collection of information available, that for this country specific ranges from 1991 to 2021 year. Following figure 51 presents the pattern on Russian unemployment, first differencing of this variable, ACF and PACF plots.

Figure 51
Russian unemployment rates, differenced data, ACF and PACF plots.
(Elaborated by the author).

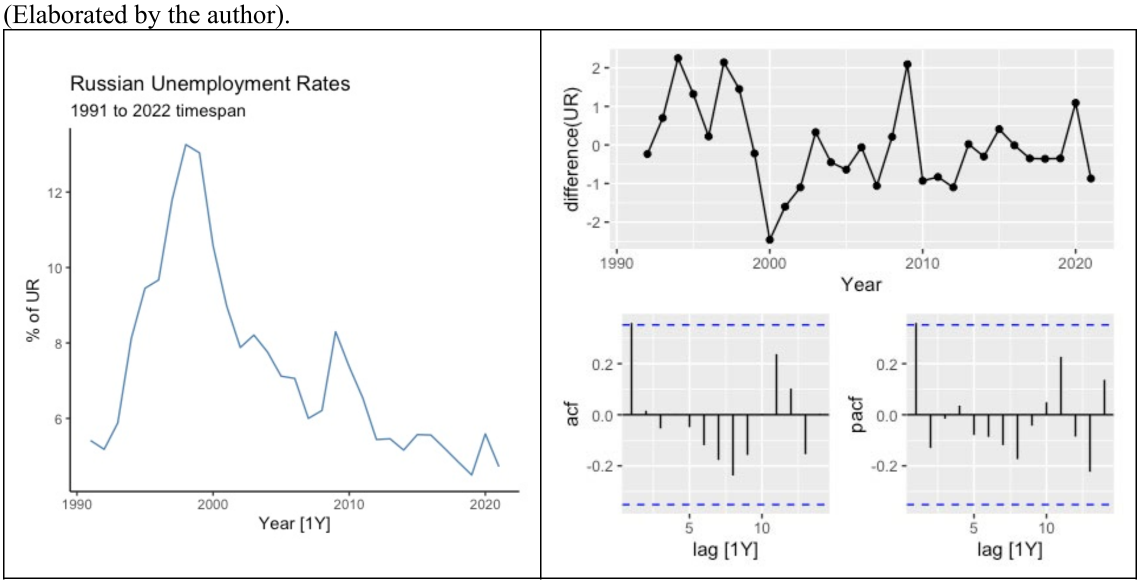


Figure 51 presenting Russian real unemployment rates on the right-side seems suggestive for a non-stationary behaviour. It is perceived a growing to a peak before the 2000s years an accentuate decline after this peak and some lowering tendency after 2015 and before the 2020 COVID-19 pandemic year. Presumed non-stationarity enables the attempt of first differentiation of Russian data and on the left-side on figure 51 unemployment appears to have a more “regular” and stationary behaviour.

Figure 51 PACF plot is suggestive at maximum an AR(3) model, when the tails shows first signs of continuous decaying; suggesting a model as ARIMA(3,1,0), whereas the ACF plot on the same figure may indicate an MA(2) model, using a similar inferential analysis. An alternative candidate could be then an ARIMA(0,1,2). We test for both visual suggestions and the two automated `ARIMA` function options, one using stepwise proceeding and other tat does not use it.

Outputs for these testing are presented on the following table 67 and from hereafter we are focusing on the training part of Russian data that comprises 26 observations from 1991 to 2016 to later test for accuracy of forecasts using there the test part of this data, that ranges from 2017 up to 2021 values of unemployment rates in the country. We reinforce that this one-year shortage of information is due data availability when we proceed the retrieving of it on World Bank and ILO databases.

Table 67
Fitting the best ARIMA model: Russian dataset.
(Elaborated by the author).

Country	*Model test*	*Orders*	*sigma2 (σ^2)*	*AIC*	*AICc*	*BIC*
RUSSIA	arima012	ARIMA(0,1,2)	1.15	78.70	79.80	82.40
	arima310	ARIMA(3,1,0) w/mean	1.20	80.60	82.60	82.50
	stepwise	ARIMA(2,0,0) w/mean	1.03	81.50	83.40	86.50
	search	ARIMA(2,0,0) w/mean	1.03	81.50	83.40	86.50

Russian dataset similarly as the Brazilian data as presented here on table 63 suggests the ARIMA(0,1,2) as the potentially better fit. AIC, AICc and BIC are all lower on this model although the error variance is the second highest. Stepwise and search models have identical values on all criteria but as we define, following Hyndman & Athanosopoulos (2021) the AICc as main parameter, we assume as the best model for Russian data the ARIMA(0,1,2).

A new ACF plot is presented on the following figure 52, now considering residuals from ARIMA(0,1,2). Figure 52 plots presents all autocorrelations within unemployment data are contained inside the threshold limits of ACF plot and residuals, at least visually, suggests a white noise behaviour. We test the for the white noise assumption statistically using again Ljung-Box and Box-Pierce test statistics.

Results for both statistical tests are presented on table 68 which is presented in detail after the presentation of figure 52. Sequential table and figure, 68 and 52 respectively, are referring to analysis regarding Russian data to complement the before presented at table 67 and following a similar pattern of results presentation that we have used for the Brazilian dataset. We are organizing text and visual presentations in a manner to ease the readiness of this chapter.

Figure 52
ACF plot of the residuals from the Russian ARIMA(0,1,2) model.
(Elaborated by the author).

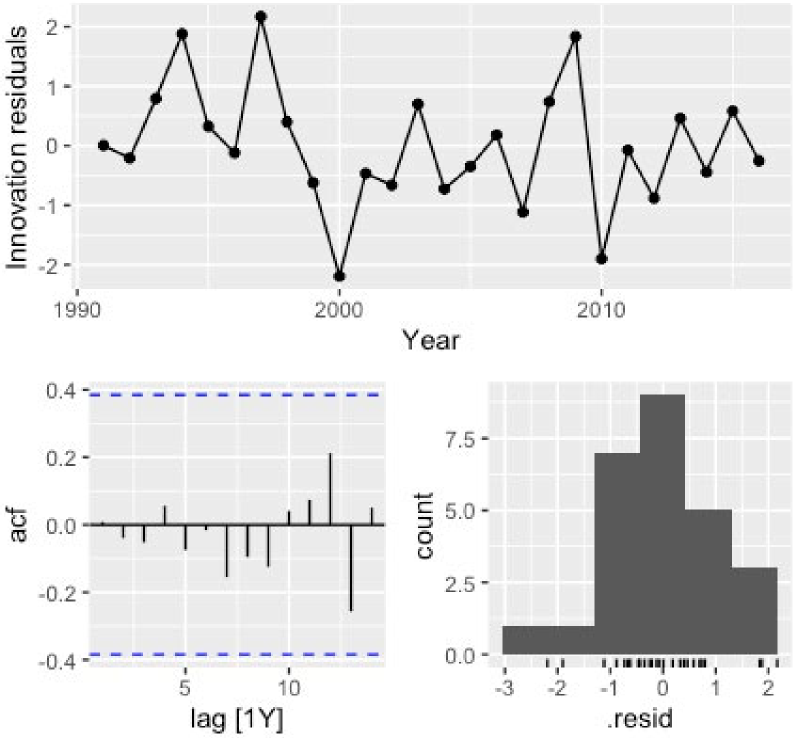


Table 68
White noise tests on the residuals ARIMA(0,1,2) model for Russian unemployment rates.
(Elaborated by the author).

Country	*Model*	*Test*	*p-value*
RUSSIA	ARIMA(0,1,2)	Ljung-Box	0.96
		Box-Pierce	0.97

The p-values presented on table 68 being above 0.05 Box-Pierce and Ljung-Box tests enables the rejection of null hypothesis for both assessments. We have statistical and visual evidence as showed on figure 75 for believe that residuals of defined Russian model are indeed white noise, indicating that ARIMA(0,1,2) is a fitted modelling to proceed with forecasts for unemployment rates in Russia.

Before to forecasts, following table 69 presents the parameters to build the ARIMA equation respective to the Russian ARIMA(0,1,2) model. As the definition suggests we have no autoregressive coefficients for this model, however the moving average and the differencing orders are relevant. Results from table 69, which considers the ARIMA(0,1,2) Russian model for unemployment rates, are used to write the basic equation presented on the sequence.

Table 69
Parameter statistics of model ARIMA(0,1,2) for Russian unemployment rates.
(Elaborated by the author).

Model	*Parameters*	*Estimate*	*Standard error*	*Statistic*	*p-value*
ARIMA(0,1,2)	MA1	0.49	0.21	2.36	0.02
	MA2	0.09	0.20	0.48	0.63

$$Russian\_UR_t = \varepsilon_t - 0.49_{t-1} - 0.09_{t-2}$$

An ARIMA model, as the one we found for unemployment data from Russia, without AR components is often referred to a Moving Average (MA) or an Integrated (I) model. Absence of AR terms means that current values of unemployment in this timeseries is not directly dependent on past values. ARIMA(0,1,2) relies on differencing (I) moving average (MA) components, that are sufficient to capture temporal patterns and dependencies in Russian data.

Finally, we have a suited model to proceed for forecasts to unemployment rates in Russia for our 10-year future horizon. Forecast is performed using training portion of Russian data and we later compare the predictions with the real unemployment rates from 2017 up to 2021 in Russia's case. First presentation is about the values projected and these are presented and compared on the following table 70.

Table 70
Forecasting Russian unemployment rates with ARIMA(0,1,2) model: Values comparison. (Elaborated by the author).

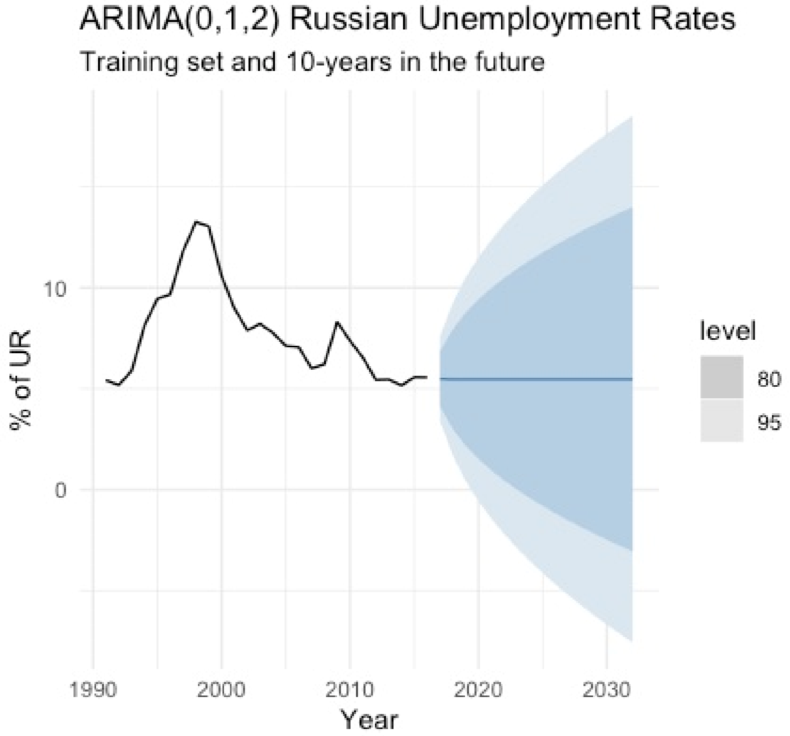
Country	*Years*	*Real UR*	*ARIMA projected UR distribution*	*ARIMA UR by mean*
RUSSIA	2017	5.21	(5.5, ±1.20)	5.49
	2018	4.85	(5.5, ±3.70)	5.47
	2019	4.50	(5.5, ±6.60)	5.47
	2020	5.59	(5.5, ±9.50)	5.47
	2021	4.72	(5.5, ±12.00)	5.47

Table 70 presents a constant value projected from 2018 and after by 5.47% of unemployment. A low number but somewhat feasible considering the overall low levels of unemployed people in Russia since the beginning of this ongoing century. Interval distribution of values however expand the possible outcomes of unemployment to dampen the naïve presumption that these rates would remain the same over the forecast horizon.

Projected unemployment rates for Russia on the following figure 53 have a resembling pattern as the one we present when performing the ETS method. Using ARIMA we found that unemployment in Russia could close to 0% and peak above 15% after 2030. This visual presentation may be deceiving as unemployment very unlikely could go so low and considering historical data does not present any suggestion to reach as high as occurred in the 1990s years. That is the reason we perform the number comparison, presented on table 70, and to not incur on misleading inferences. Accuracy checks for ARIMA(0,1,2) model is the following: Winkler score = 9.55, percentile and CRPS are both by 0.63.

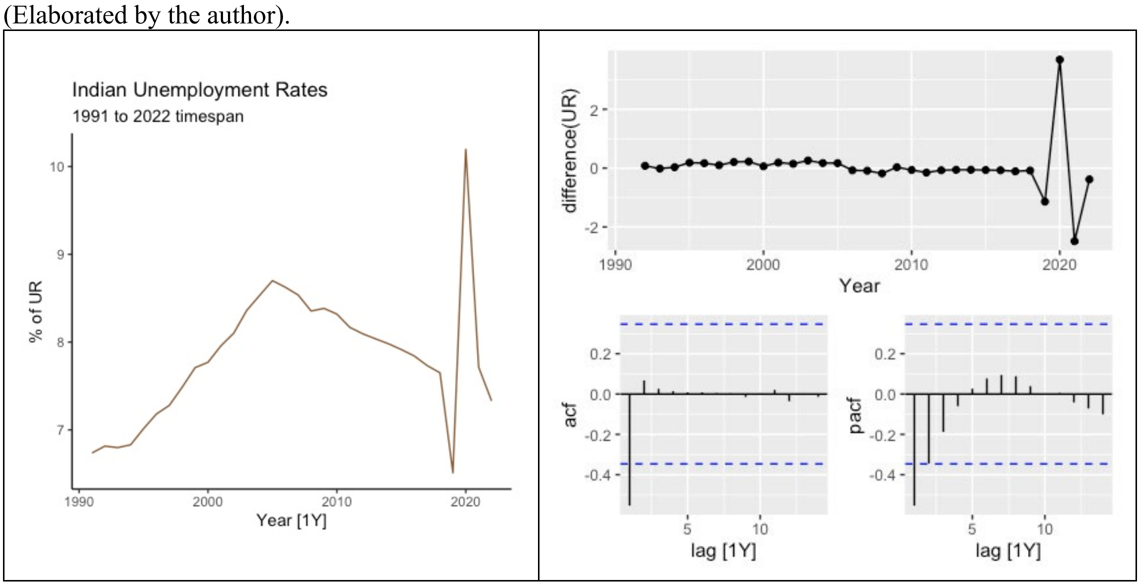
Visual presentation of potential future of unemployment in Russia is presented on figure 53 that shows training portion of data and the projected possible from 2017 up to 2032.

Figure 53
ARIMA(0,1,2) forecast for Russian unemployment rates.
(Elaborated by the author).



Next country to be assessed is India. Indian full dataset has the same number of observations as Brazil, 32, respective to unemployment rates from 1991 to 2022. Figure 54 similarly as the Brazilian and Russian cases presents the variable natural behaviour, first differentiation of it as well ACF and PACF plots.

Figure 54
Indian unemployment rates, differenced data, ACF and PACF plots.
(Elaborated by the author).



Indian real unemployment rates on a first impression seems to be fit with a non-stationarity pattern. From the initial point of data available (1991) to around 2015 data unemployment seem to consistently grow to an also relatively constant decline after it and peaking high on the COVID-19 pandemic years. On the side of figure 54 we see a clear distinction on the target variable, that when first differencing procedure is applied seems way more stationary although the significant peak in 2020 is still captured.

PACF plot also presented on figure 54 indicates an potential AR(2) model which is suggestive for an ARIMA(2,1,0) type of modelling. ACF plot also indicates some number around 2, offering some room to we infer that an MA(2) model being ARIMA(0,1,2) could be suited as well. We assess again the two indications from visual analyses and the automated ARIMA function options, with and without stepwise proceeding.

Results from these assessments are presented on the following table 71. Going forward with Indian examination of unemployment rates we turn attention for the training portion of country's available data as a benchmark for our Indian ARIMA model that we later check for fitting using the test set of India's data. Best model to proceed with is selected considering the results presented on the next table.

Table 71
Fitting the best ARIMA model: Indian dataset.
(Elaborated by the author).

Country	*Model test*	*Orders*	*sigma2 (σ^2)*	*AIC*	*AICc*	*BIC*
INDIA	arima210	ARIMA(2,1,0)	0.009	-41.80	-40.70	-38.20
	arima012	ARIMA(0,1,2)	0.01	-38.10	-37.00	-34.40
	stepwise	ARIMA(0,2,1)	0.01	-40.20	-39.70	-37.90
	search	ARIMA(0,2,1)	0.01	-40.20	-39.70	-37.90

Table 71 suggests that our inferred ARIMA(2,1,0) is the one having the better fit by 1.0 point, comparing AICc with both stepwise and search models. On all the analysed models we have lower values on sigma2 even considering the ones that, in comparison, could potentially offer more misleading on forecasts. Considering that AICc in ARIMA(2,1,0) is slightly better, this is the selection as best model for Indian data and the one used for the new presentation of unemployment variable on the following figure 55.

Presented next on the figure 55 plots presents a new residuals behaviour and the new ACF illustration appears to show that autocorrelations within unemployment variable are contained inside the threshold limits of the plot. Residuals presentation both in the line and on the histogram appears to indicate white noise behaviour. To confirm if white noise is indeed happening, we perform Ljung-Box and Box-Pierce test statistics to have a statistical confirmation about it. Results for these tests are presented next on table 72.

Figure 55
ACF plot of the residuals from the Indian ARIMA(2,1,0) model.
(Elaborated by the author).

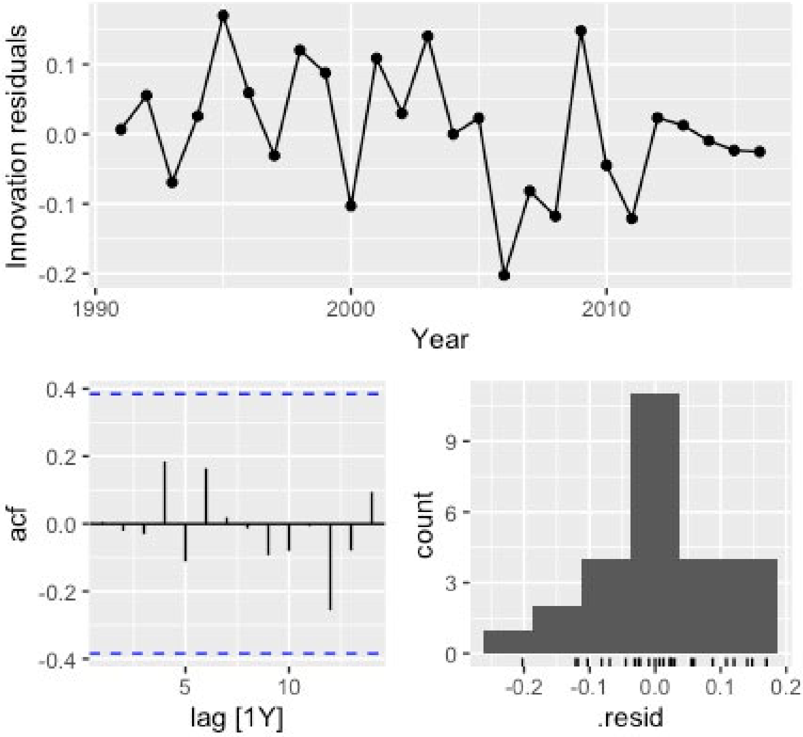


Table 72
White noise tests on the residuals ARIMA(2,1,0) model for Indian unemployment rates.
(Elaborated by the author).

Country	*Model*	*Test*	*p-value*
INDIA	ARIMA(2,1,0)	Ljung-Box	0.98
		Box-Pierce	0.98

High p-values presented on table 72, as happened on both Brazil and Russia, enables the rejection of null hypothesis for Ljung-Box and Box-Pierce parameters. Therefore, we have beyond the visual suggestion on figure 55, statistical evidence to believe that our Indian dataset have white noise behaviour and that ARIMA(2,1,0) is a well fitted model to proceed with forecasts.

On table 73 we present the parameters of the ARIMA equation for the Indian ARIMA(2,1,0) model. Again, we use the Hyndman & Athanasopoulos (2021) basic equation to fit one considering the data we are working so far. Equation could be written as its following after the presented on the following table.

Table 73
Parameter statistics of model ARIMA(2,1,0) for Indian unemployment rates.
(Elaborated by the author).

Model	*Parameters*	*Estimate*	*Standard error*	*Statistic*	*p-value*
ARIMA(2,1,0)	AR1	0.55	0.19	2.85	0.00
	AR2	0.20	0.19	1.04	0.30

$$Indian\_UR_t = \varepsilon_t + 0.55_{t-1} + 0.20_{t-2}$$

ARIMA(2,1,0) model is effective in capturing long-term dependencies in the timeseries, meaning that Indian current unemployment rates is directly related to its past values through the autoregressive terms. We have 2 AR terms and no moving averages, distinction on both models is that for India exists a clear need for differencing, as was already noticeable on visual presentation from figure 55. Premise to believe that our model is suited is that primary dependencies within Indian data are well-captured by autoregressive relationships.

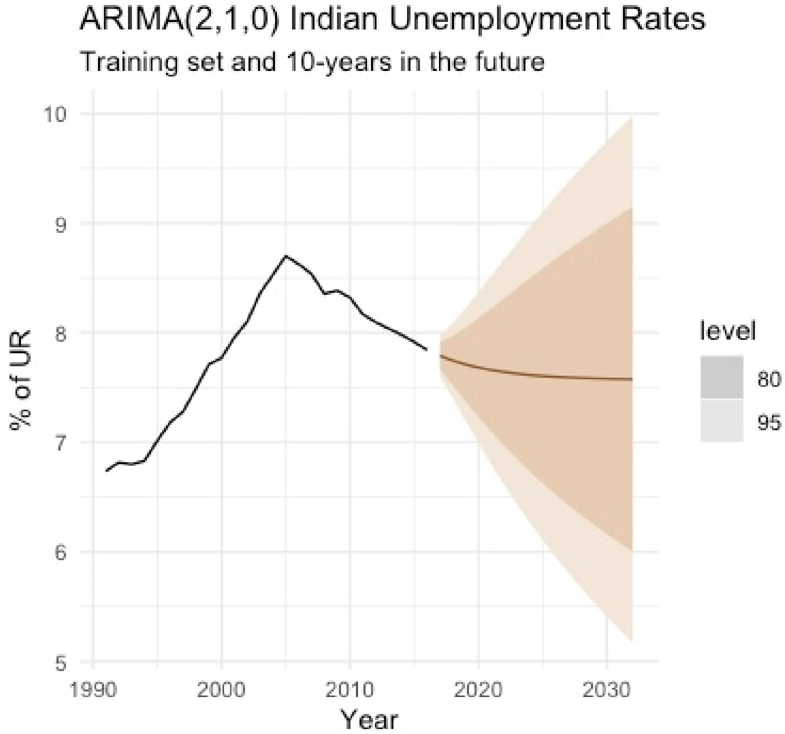
Moving forward with the ARIMA(2,1,0) model we proceed with the forecasts of unemployment rates in India. Prospection of values are performed on training portion for comparison purposes with real data available on the test set of Indian data. This examination of potential unemployment is presented on the following table 74, following a similar approach used on Brazil and Russia.

Table 74
Forecasting Indian unemployment rates with ARIMA(2,1,0) model: Values comparison.
(Elaborated by the author).

Country	*Years*	*Real UR*	*ARIMA projected UR distribution*	*ARIMA UR by mean*
INDIA	2017	7.73	(7.8, ±0.00)	7.79
	2018	7.65	(7.7, ±0.03)	7.74
	2019	6.51	(7.7, ±0.07)	7.71
	2020	10.19	(7.7, ±0.12)	7.68
	2021	7.71	(7.7, ±0.19)	7.66
	2022	7.33	(7.6, ±0.28)	7.64

Table 74 presents very assertive projected values when comparing with the real unemployment rate for India on the 2017-2022 period. Overall perception is that all forecasted values, even by the mean, are well approximate with the rates retrieved from World Bank and ILO for the country. Only disparity is in 2020, more than 2% underestimation on the predicted and factual rate. This is the peak of unemployment levels as presented on figure 54 where COVID-19 pandemic effects were massive on the Indian labour market. Despite that, all forecasts seem well effective. Projections for the potential future unemployment in India are presented on the following figure 56.

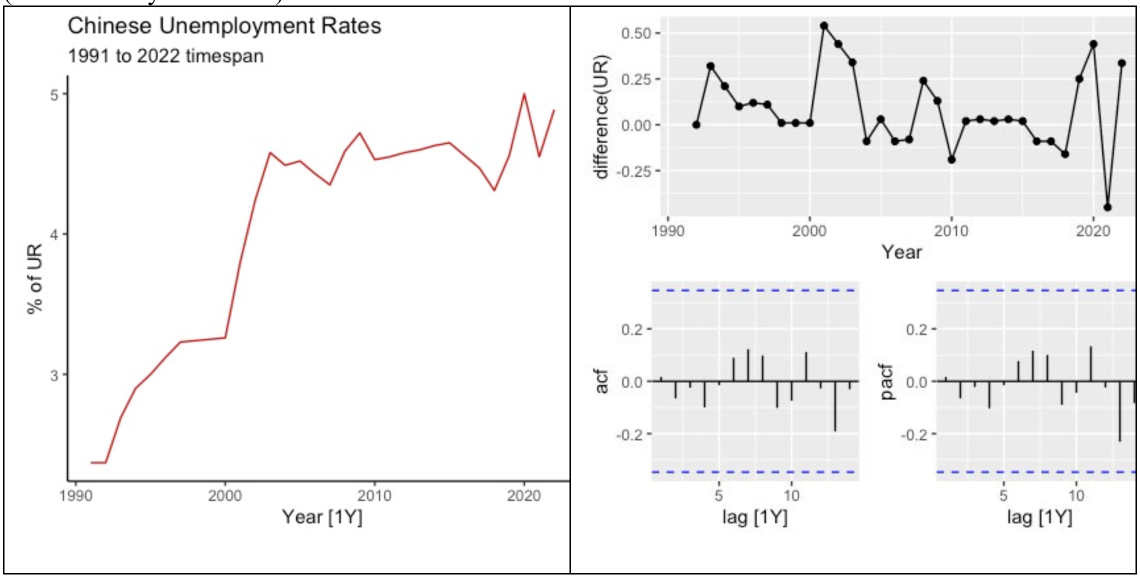
Figure 56
ARIMA(2,1,0) forecast for Indian unemployment rates.
(Elaborated by the author).



Considering the comparison of test data and the projected values for unemployment rates presented on table 74, we have some reason to believe that numbers projected and presented on figure 56 may very well happen for Indian unemployment in our 10-year intended forecast horizon. ARIMA(2,1,0) forecast suggest that could go as low as 6% to 5% and high as around 9% or 10%. Accounting that comparison on table 74 suggests a feasible projected values and that we have a low estimated error variance, we believe to have indeed a fitted model and, consequently, a good forecast for India's case. Accuracy checks for the model is: Winkler score = 17.90, percentile and CRPS are both by 0.63.

Fourth of the five BRICS countries to be analysed is regarding the Chinese unemployment data. Complete dataset, as in Brazil and India has 32 observations of this variable, ranging from 1991 to 2022 year. Following figure 57 presents the real unemployment in the country, first differencing of the data, ACF and PACF illustrations.

Figure 57
Chinese unemployment rates, differenced data, ACF and PACF plots.
(Elaborated by the author).



Chinese unemployment rates apparently have no sign of being stationary, having an increasing and decreasing of these numbers in a not necessarily evident pattern. First difference of unemployment data presented on the right side of figure 80 shows a more stationary behaviour of the data. PACF and ACF plots are not very clear on potential terms to be used on our ARIMA modelling but we infer nonetheless considering what could be perceived.

For PACF we suppose a possible AR(4), meaning an ARIMA(4,1,0) and from ACF we presume an MA(6), and ARIMA(0,1,6). As visual suggestions are not so evident as happened especially for Russia and India data, probably the `ARIMA` functions with or without stepwise procedures will fit the best model for Chinese data. Table 75 presents the results of the assessments and inferred models and applications from `ARIMA` function.

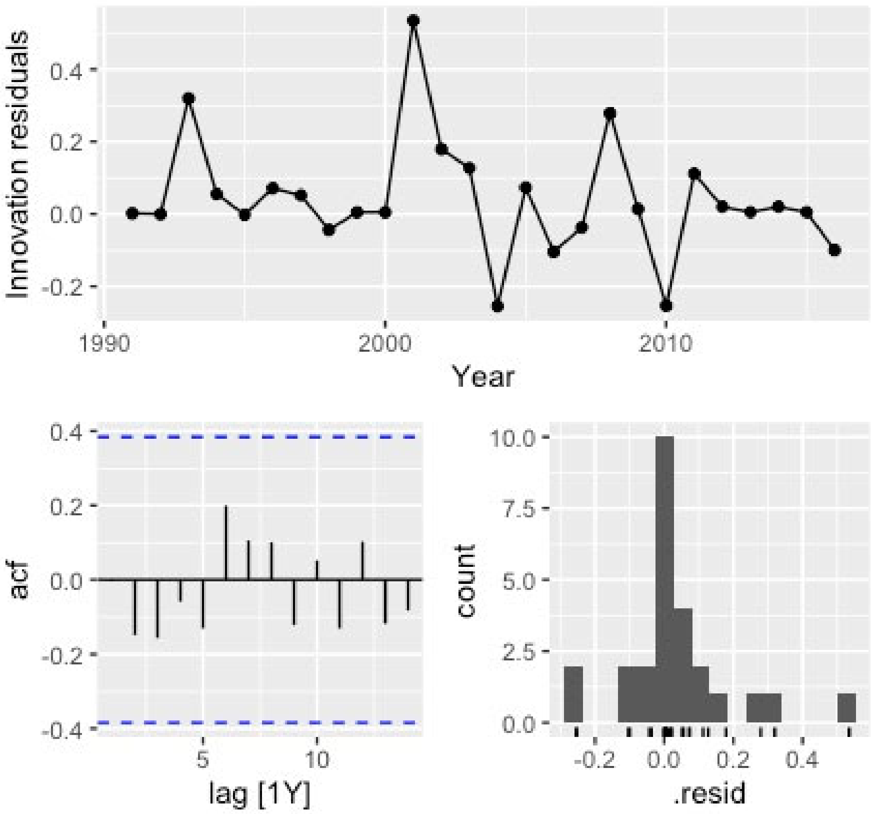
Table 75
Fitting the best ARIMA model: Chinese dataset.
(Elaborated by the author).

Country	*Model test*	*Orders*	*sigma2 (σ^2)*	*AIC*	*AICc*	*BIC*
CHINA	arima410	ARIMA(4,1,0)	0.03	-8.74	-5.59	-2.65
	arima016	ARIMA(0,1,2)	0.03	-5.22	1.37	3.31
	stepwise	ARIMA(0,1,1) w/ drift	0.02	-14.90	-13.70	-11.20
	search	ARIMA(1,1,0)	0.03	-14.30	-13.70	-11.80

As we presumed, visual deductions from figure 57 were not very adequate on minimising errors according to our results presented on table 75. Stepwise and search models perform better than the ones we speculate, being the potential best fit the ARIMA(1,1,0) obtained not using the stepwise procedure. AICc on stepwise and search tests are the same we use as second criteria to select the best one the BIC index, once again, following the Hyndman & Athanasopoulos (2021) proposition, suggesting the ARIMA(1,1,0) as the best to proceed.

Therefore, using ARIMA(1,1,0) we plot the residuals of this model, a new ACF plot for Chinese unemployment data on the following figure 58.

Figure 58
ACF plot of the residuals from the Chinese ARIMA(1,1,0) model.
(Elaborated by the author).



New ACF plot presented on figure 58 indicates that all autocorrelations of Chinese unemployment are within the threshold limits of the plot. First ACF presentation already suggest this, but at this new one we have a more solid illustration. Both portraits of residuals are indicating a white noise behaviour of data, especially the line on the top of the figure 58. We use Ljung-Box and Box-Pierce tests to check if white noise presumption is correct. Results about its are presented next on the following table 76.

Table 76
White noise tests on the residuals ARIMA(1,1,0) model for Chinese unemployment rates.
(Elaborated by the author).

Country	*Model*	*Test*	*p-value*
CHINA	ARIMA(1,1,0)	Ljung-Box	0.98
		Box-Pierce	0.98

Elevated p-values showed on table 76 permit to reject null hypothesis for both Ljung-Box and Box-Pierce tests. Indicating that Chinese data indeed has white noise behaviour and confirming the fitting of ARIMA(1,1,0). Table 77 presents the parameters for ARIMA equation for the Chinese model ARIMA(1,1,0). Having in hand the AR coefficient we present the basic ARIMA mathematical equation on the sequence.

Table 77
Parameter statistics of model ARIMA(1,1,0) for Chinese unemployment rates.
(Elaborated by the author).

Model	*Parameters*	*Estimate*	*Standard error*	*Statistic*	*p-value*
ARIMA(1,1,0)	AR1	0.48	0.17	2.86	0.00

$$Chinese\_UR_t = \varepsilon_t + 0.48_{t-1}$$

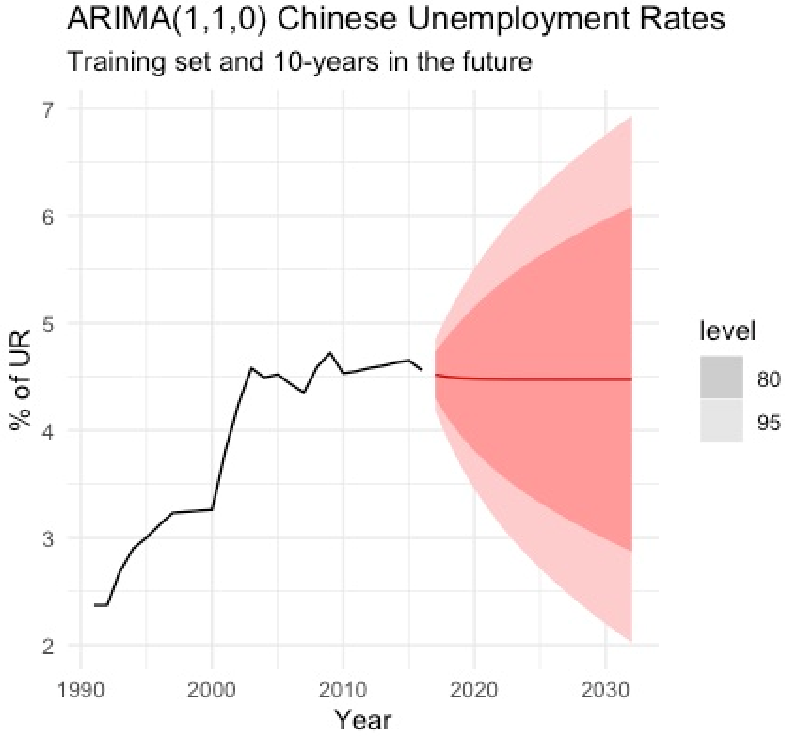
ARIMA$(1,1,0)$ model is effective in capturing long-term dependencies for the Chinese data, similar as the Indian case we have the suggestion that current rates of unemployment in China is directly influenced to past values and well captured by the autoregressive term. With a model identified we may proceed for the forecasts. Following table 78 as performed for the other three countries analysed so far, presents real unemployment and projected values for Chinese data.

Table 78
Forecasting Chinese unemployment rates with ARIMA(1,1,0) model: Values comparison.
(Elaborated by the author).

Country	*Years*	*Real UR*	*ARIMA projected UR distribution*	*ARIMA UR by mean*
CHINA	2017	4.47	(4.5, $\pm$0.03)	4.52
	2018	4.31	(4.5, $\pm$0.09)	4.50
	2019	4.56	(4.5, $\pm$0.18)	4.49
	2020	5.00	(4.5, $\pm$0.28)	4.48
	2021	4.55	(4.5, $\pm$0.38)	4.48
	2022	4.88	(4.5, $\pm$0.48)	4.48

Table 78 presents good projected values in comparison with real unemployment rates for China on the timespan covered by the test portion of country's data. There are no large discrepancies from actual values measured by World Bank and ILO and the ones we forecast using the ARIMA$(1,1,0)$ model. Considering the assertively predicted values as well the low estimated error variance of the model we have reason to believe that Chinese unemployment in our 10-year forecast horizon is well suited. Visual presentation about it is illustrated on the following figure 59.

Figure 59
ARIMA(1,1,0) forecast for Chinese unemployment rates.
(Elaborated by the author).



Similar as happened for Russian data the plot presented on figure 59 is similar with the one projected when using the ETS technique for forecast on the previous topic of this thesis. Overall, comparing with the other four BRICS countries, unemployment in China is significantly low and the projection for the future is that this tendency may endure. People outside the Chinese labour force could reach, at maximum around 7% after 2030 but in the opposite could low to around 2%.

Policies regarding labour market on China apparently have been working well providing positive results. We found some evidence to continue with them or to slightly improve them, to mitigate the potential worst-case scenario, that is still a good one comparing with the BRICS colleagues. Historical data for unemployment in the country does not show a peak above 6% and we have room to believe that this will not happening in the future. ARIMA(1,1,0) accuracy check for the forecasts is the following: Winkler score = 1.79, percentile and CRPS are both around 0.16.

To close out on individual assessments, we move to the fifth BRICS country, the South Africa. Again, for the African analysis the proceedings are similar with the ones produced for Brazil, Russia, India, and China, we present first a visual illustration about unemployment in the country considering the complete dataset, with 32 observations, on the full covered time that ranges from 1991 to 2022. Figure 60 presents the behaviour of unemployment in South Africa.

Figure 60
South African unemployment rates, differenced data, ACF and PACF plots.
(Elaborated by the author).

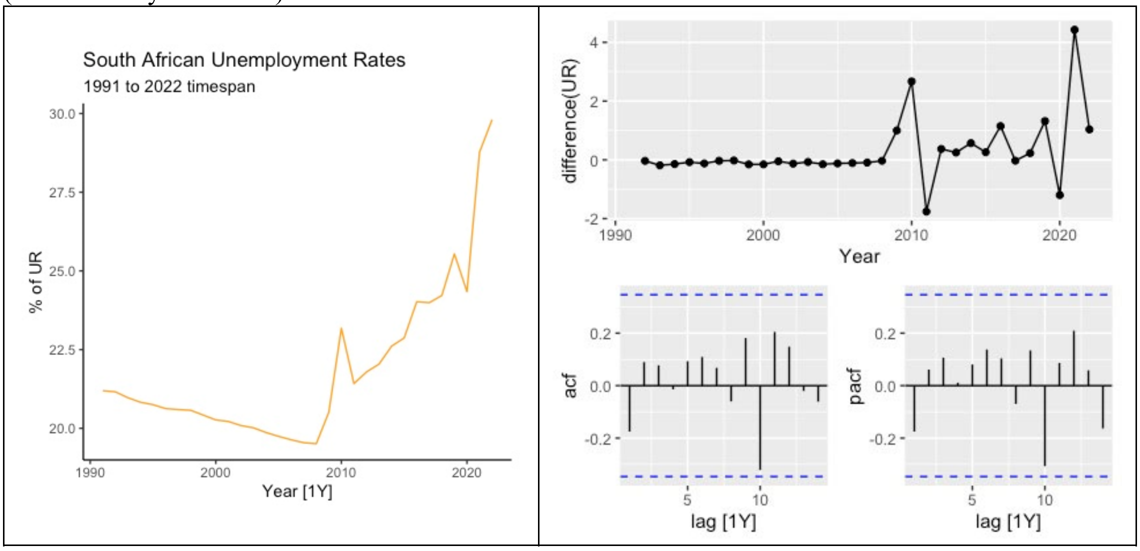


Figure 60 confirms all conferences previously made that indicates South African unemployment as the highest within BRICS countries. Real unemployment rates South Africa when low is still very high in comparison, being minimal, under 20.0%, on the final years of the first decade of 2000s years and peaking close to 30.0% after the COVID-19 pandemic effects. Data appears to be non-stationarity having a relatively constant growing pattern. Differencing the data, as presented on the left-side of figure 60. Appears to minimise variance and generate some level of stationarity.

Assuming the data after the differencing process presented on figure 60, PACF plot is suggesting for AR(6) model, which indicates an potential model like ARIMA(6,1,0). ACF plot, visually, suggests something around MA(3), indicating an alternative candidate an ARIMA(0,1,3). Therefore, we test models ARIMA(6,1,0), ARIMA(0,1,3) and the two automated `ARIMA` selections, one using the stepwise procedure, and one that searches for an alternative model not using the stepwise.

Outputs for the checking about these four potential fitted models are presented next, on the following table 79. As in figure 60 we present the complete range of observations for South African unemployment, from hereafter on further analyses we are using the training portion of country's data, now using 26 observations from 1991 to 2016.

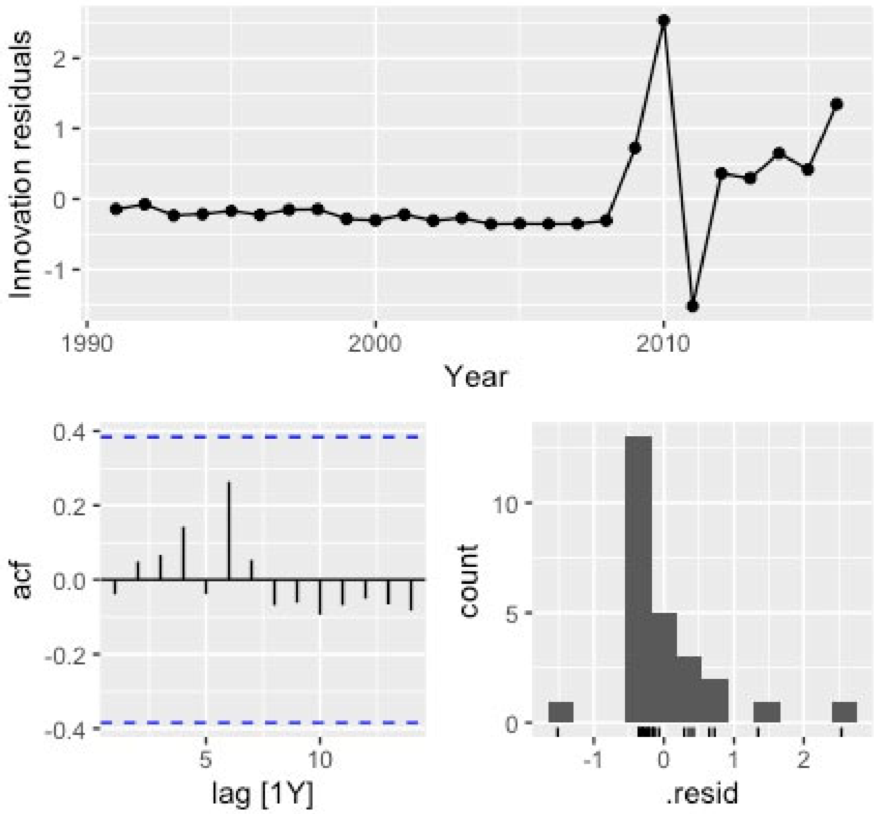
Table 79
Fitting the best ARIMA model: South African dataset.
(Elaborated by the author).

Country	*Model test*	*Orders*	*sigma2 (σ^2)*	*AIC*	*AICc*	*BIC*
SOUTH AFRICA	arima610	ARIMA(6,1,0)	0.60	66.10	72.70	74.60
	arima013	ARIMA(0,1,3)	0.59	62.70	64.70	67.50
	stepwise	ARIMA(1,0,0) w/mean	0.55	63.30	64.40	67.10
	search	ARIMA(1,0,0) w/mean	0.55	63.30	64.40	67.10

Stepwise and search tested models are fitted and they perform slightly better than the inferred ARIMA(0,1,3), minimising AICc values by 0.30 in comparison. AIC, BIC and sigma2 as well have better indexes on the ARIMA(1,0,0), suggesting this as the one having the best fit.

From table 79 results we adopt as best model for South African data an ARIMA(1,0,0). Having this defined, we move to produce a new ACF plot, that differently from the one presented on figure 60 is considering residuals of ARIMA(1,0,0). This presentation is next, on figure 61.

Figure 61
ACF plot of the residuals from the South African ARIMA(1,0,0) model.
(Elaborated by the author).



New ACF plot presented on figure 61 shows that all autocorrelations are well contained inside the threshold limits from the blue dotted line. Residuals presented in the line format at top of figure and histogram at the bottom-right, seems to indicate that these residuals apparently ae behaving as white noise. Ideally, residuals should be white noise, meaning that within data exists significant autocorrelations, homoscedastic, and normally distributed (Huang, 2015). We check for white noise presumption using the portmanteau tests of Box–Pierce and Ljung–Box statistics. Results for both tests are presented next on table 80.

Table 80
White noise tests on the residuals ARIMA(1,0,0) model for South African unemployment rates.
(Elaborated by the author).

Country	*Model*	*Test*	*p-value*
SOUTH AFRICA	ARIMA(1,0,0) w/mean	Ljung-Box	0.83
		Box-Pierce	0.84

Both p-values presented on table 80, for both tests, are considerably above of 0.05, enabling to reject null hypothesis on both tests. Null hypotheses rejected we have statistical evidence attesting that residual of this model, have a white noise behaviour and ARIMA(1,0,0) model for South African for unemployment rates is well fitted. Before to usage of this model into forecasts, table 81 presents the parameter of the model followed by the ARIMA equation for South Africa data.

Table 81
Parameter statistics of model ARIMA(1,0,0) for South African unemployment rates.
(Elaborated by the author).

Model	*Parameters*	*Estimate*	*Standard error*	*Statistic*	*p-value*
ARIMA(1,0,0) w/mean	AR1	0.85	0.11	7.79	< 0.00
	Constant	3.03	0.12	24.40	< 0.00

$$SouthAfrican\_UR_t = 3.03 + 0.85B_{t-1} + \varepsilon_t$$

Having one autoregressive parameter, ARIMA(1,0,0) model equation suggests that the current value of unemployment in South Africa is significantly influenced by past values up to 1 lag. This justifies the endurance of elevated indexes of unemployed people being consistently high, it already was in this form at the beginning of data retrieving (1991) and potentially will remain somewhat alike on future values predicted.

Still on the equation, coefficients for B indicates the strength and direction (here a positive one by 0.85) of the influence by each lag while the constant term represents the baseline level of this timeseries. White noise error term (ε_t) captures the variability in the data that is not explained by lagged values and B.

Details about South African data presented so far give us the basis to proceed for the forecasts for the country's unemployment rates in a 10-year into the future horizon. Presentation of the predicted values follows the same format from Brazil, Russia, India, and China, a number comparison using the test portion of data and the forecasted numbers, presented on table 82, and later a visual presentation about those predicted unemployment rates is depicted on figure 62.

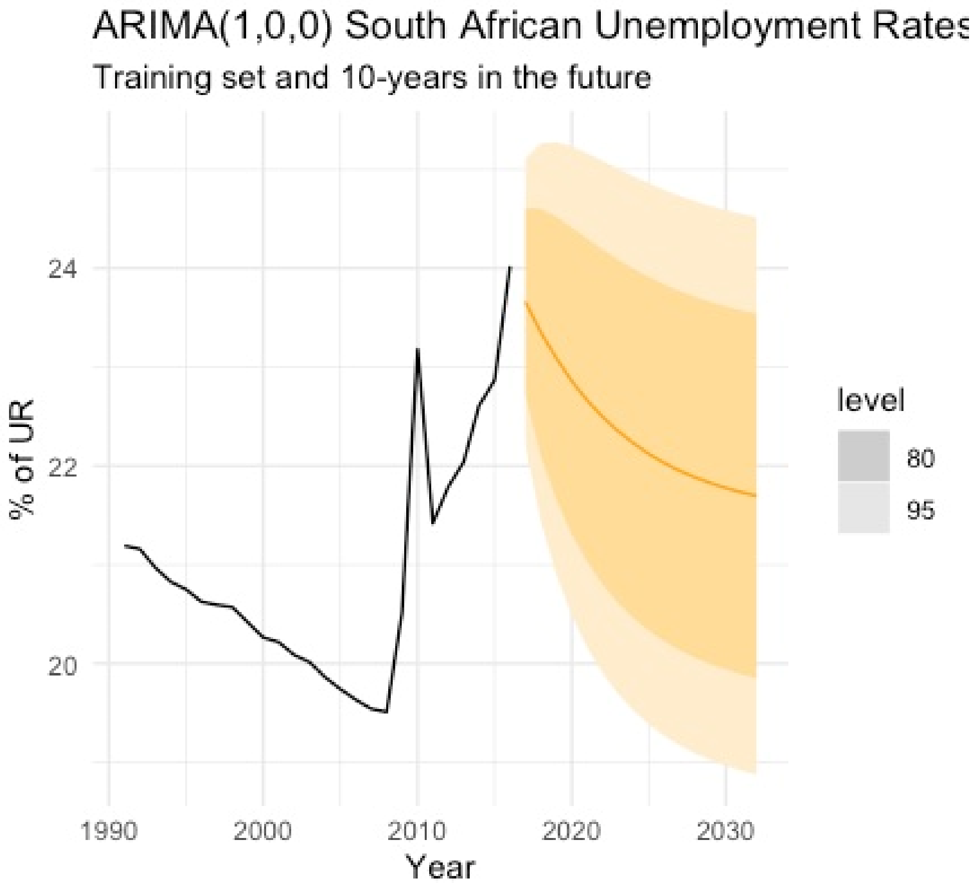
Table 82
Forecasting South African unemployment rates with ARIMA(1,0,0) model: Values comparison.
(Elaborated by the author).

Country	*Years*	*Real UR*	*ARIMA projected UR distribution*	*ARIMA UR by mean*
SOUTH AFRICA	2017	23.99	(24, $\pm$0.55)	23.70
	2018	24.22	(23, $\pm$0.95)	23.40
	2019	25.54	(23, $\pm$1.20)	23.10
	2020	24.34	(23, $\pm$1.50)	22.90
	2021	28.77	(23, $\pm$1.60)	22.70

	2022	29.80	(22, ±1.70)	22.50

Results on table 82 offers a mixed analysis on how effective forecasted values using the ARIMA(1,0,0) may be for unemployment rates in South Africa. 2017 and 2018 especially and, in comparison, 2019 as well, values predicted by the method are relatively close to the real unemployment percentage in the country according to World Bank and ILO repositories. Starting on 2020, not coincidently the COVID-19 pandemic year, and after real and forecasted values are way apart, by more than 6% on 2021 and 7% in 2022.

Figure 62
ARIMA(1,0,0) forecast for South African unemployment rates.
(Elaborated by the author).



Forecasting unemployment in South Africa with the ARIMA(1,0,0) model could be efficient up to the point that some unexpected economic shock occurs, as the pandemic. Before the Coronavirus outbreak, predicted and real values were close, but after that the projections underestimated the damaging effects that COVID-19 already imposed and to what extension this may endure in the future. South African unemployment as presented on figure 62 could peak around 25% and go as low as under 19% after 2030. Indicatives seems to suggest that unemployment curve in the country's future may somewhat repeat historical data, remaining the highest on unemployed people inside BRICS group.

As a group, indeed, this could present as an urge to cooperate, countries in BRICS that suffers less with labour market dysfunction could offer some help to mitigate South

African problem. As well, inside efforts by the Africans itself must be taken urgently in pace to help population's wellbeing and, consequently, better economic performance. Statistical accuracy checks for ARIMA(1,0,0) South African model is the following: Winkler score = 61.60, percentile = 2.59 and CRPS = 2.58.

As we perform for STL and ETS methods, after the assessment of Brazil, Russia, India, China, and South Africa individually, we perform another analysis now considering the five as a group to have a general perspective about the BRICS. As done for the countries in separate, we use the proportional data by 80/20, on training and test datasets. Timespan of observed values of unemployment rates ranges from 1991 to 2022, Russia has one-year shortage of data, but this does not affect overall analysis.

As it was noticeable when performing the ARIMA model fit country by country, each of them, according to their own data characteristics, have its specific models, some having AR terms with no MA whereas other have only MA term, with autoregressive parts. Some of the five countries have their the d part on the ARIMA(p, d, q) form as relevant, while for other it does not exist.

Complementing to these specificities regarding each of the five countries unemployment variable behaviour, we observe thoroughly each ACF and PACF plots. From these visual presentations we infer early indicatives about potential terms that AR and MA terms could assume and, therefore, what could be the ARIMA model ideal for that country under analyses.

These country-level features add more complexity on the intent to assess countries as a group. Taking all the above into consideration, we rely on `ARIMA` function from `fable` R-Studio package to automatize the process of selecting the best model to be used into the forecasts. As we referred earlier, the default of the function uses the stepwise procedure on identifying best models and to expand the range of possibilities, we test for obtain the best suggested model applying and not applying stepwise, testing models that we name "step" and "search", respectively.

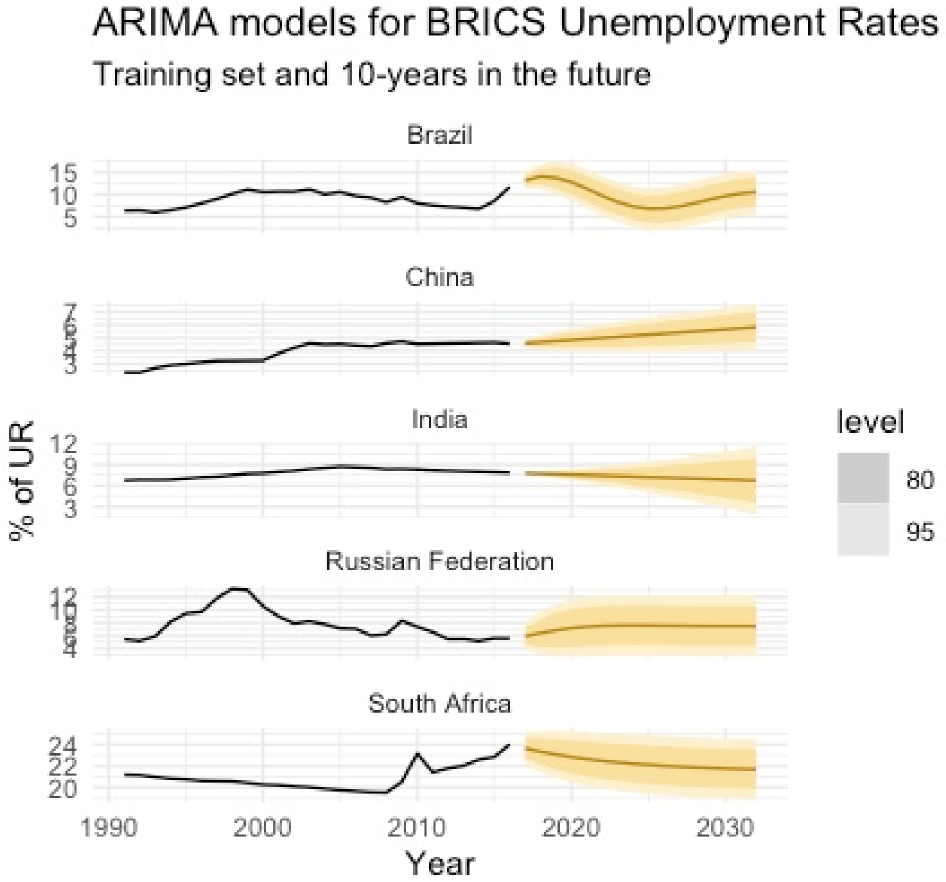
Results for these proceedings check for the ARIMA modelling are presented on the following table 83. We emphasize what was already illustrated before, for some countries, in some cases, the defaults of the method are not necessarily the best fit we have obtained. This collective BRICS assessment does not surpass the individual country's analysis when the interest is oriented for one of the five countries in specific.

Table 83
Fitting the best model from ARIMA function: Stepwise and search selections.
(Elaborated by the author).

Country	*"step"*	*"search"*
BRAZIL	ARIMA(3,0,0) w/mean	ARIMA(3,0,0) w/mean
RUSSIA	ARIMA(2,0,0) w/mean	ARIMA(2,0,0) w/mean
INDIA	ARIMA(0,2,1)	ARIMA(0,2,1)
CHINA	ARIMA(0,1,1) w/drift	ARIMA(1,1,0)
SOUTH AFRICA	ARIMA(1,0,0) w/mean	ARIMA(1,0,0) w/mean

Presented in table 83 are the two models fitted by `ARIMA` function using and not using stepwise proceedings. In general, if we use or not the stepwise procedure the model fit is the same for every country except for Chinese data, where we highlight in the table that the "search" model of ARIMA(1,1,0) for China is the one presenting best results on errors minimisation. Figure 63 presents the projected values for unemployment considering the models presented on table 83.

Figure 63
ARIMA forecast application for BRICS countries.
(Elaborated by the author).



All automatically identified models, on "step" and "search" columns of table 83 are found to have a white noise behaviour confirmed on both Ljung-Box (1970) and Box-Pierce (1978) statistics returning large p-values indexes, every one of them being above 0.05 usually adopted threshold. Which is an important result, considering residuals being white noise are essential on the presumed autocorrelations within data when applying ARIMA-based forecasts.

Overall, when we compared forecasted values and real numbers, we have solid results. Attesting the ARIMA technique on providing reasonable estimations. Accuracy of the method were performed through Winkler score, percentile, and Continuous Ranked Probability Score. Results for ARIMA specific are illustrated on table 84, we remind you to compare following table 84 with the previously presented accuracy measurements table that contains same statistic parameters (Winkler score, percentile and CRPS), for STL and ETS methods.

Table 84
Accuracy checking for ARIMA forecasts: BRICS unemployment rates.
(Elaborated by the author).

Country	***Forecast Model***	***winkler***	***Percentile***	***CRPS***
BRAZIL	**ARIMA(0,1,2)**	9.67	1.03	1.03
RUSSIA	**ARIMA(0,1,2)**	9.55	0.63	0.63
INDIA	**ARIMA(2,1,0)**	17.90	0.63	0.63
CHINA	**ARIMA(1,1,0)**	1.79	0.16	0.16
SOUTH AFRICA	**ARIMA(1,0,0)**	61.60	2.59	2.59

ARIMA method does not necessarily outperforms neither STL nor ETS methods for none of the countries. In fact, performs worse than both considering Brazilian, Russian, and South African data whereas for India and China results are similar on both percentile and CRPS while the STL forecast performs better in these two countries when observing the Winkler score. For further specification on the indexes we reinforce, compare tables that presents diagnostic measurements.

Diagnostic checks for median values when forecasting by the ARIMA method is slightly outperformed by Exponential Smoothing proceeding by Winkler score, which was of 7.35 there and is 7.39 here. Percentile and Continuous Ranked Probability Score from ARIMA are worst if comparing with both ETS and STL forecast procedures. For ARIMA the median values are about 0.79 for percentile and 0.78 on CRPS (it was 0.63 for both in ETS and 0.67 also for both in STL).

ARIMA being not the one with overall better results of the three ones we applied so far is a little surprising, considering that this is one of the most recurrently used methodology on forecasting purposes. ARIMA a deceitful method at all, outperforming STL and ETS with Brazilian data and producing good results for India is a good indicative of the methodological procedure as a good one even not being the better when analysing the BRICS as a collective.

As we presume when defining the premises of this chapter, it is not reasonable to presume that one method would perform better in all scenarios for all countries and with every data specificity. The reasoning on applying different methodologies to identify where performs better or worse than other seems even more justifiable when we close on the ARIMA usage and how it complements the before produced STL and ETS forecasts. At last, we have one more method to assess before to proceed with scenario forecasting and propositions, the Artificial Neural Networks. Next topic covers the ANN method before we present an extra one combining the four used.

4.4.4. Artificial Neural Networks – ANN application.

Considering the intention to forecast unemployment rates for Brazil, Russia, India, China, and South Africa into a 10-year horizon of a unique macroeconomic variable, we use lagged values of unemployment as the initial inputs of our neural networks. Due the usage of lagged values of unemployment, this indicates a Neural Network Autoregression (NNAR) model (Hyndman & Athanosopoulos, 2021).

`NNETAR` function from `fable` R-Studio package fits the ideal $\mathrm{NNAR(p,k)}$ model in an automatized process, where p are the lagged inputs and k the number of nodes in the hidden layer level of the network. All these procedures respecting country's data specificities. For every of the five BRICS countries, we are not predefining neither p nor k, they will be selected according to the optimal number of lags based on the AIC criteria for a linear $\mathrm{AR(p)}$ model.

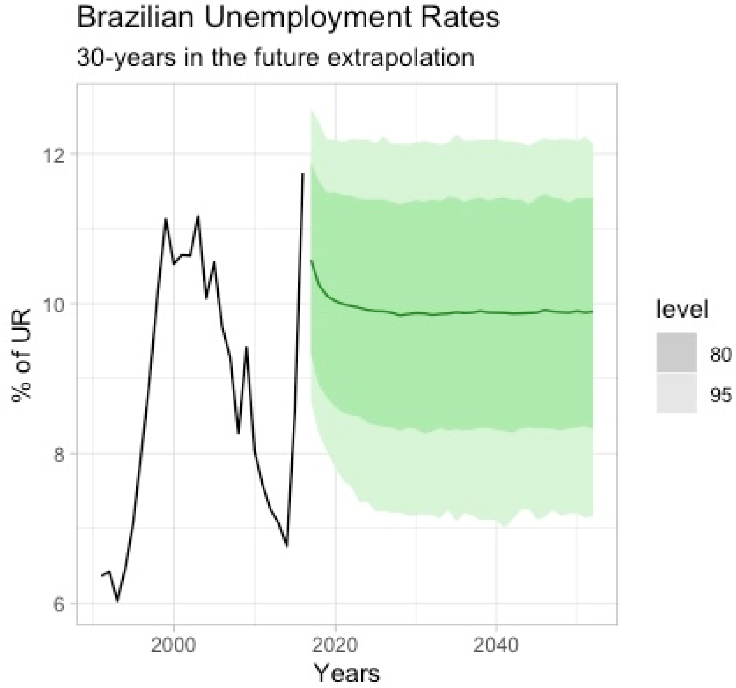
Also, we presume data to be non-seasonal, considering what returns from the previous application of STL, ETS, and ARIMA methods as well the pre-assessments of data made earlier. On producing a UR_{t+1} projection we use the historical data of unemployment for every country to be analysed and for UR_{t+2} the previously estimated forecast is added with historical data and this procedure is repeated up to the end of the 10-year intended forecast horizon.

From these defined proceedings we move for individual application country by country to latter assess them collectively as the BRICS. As done for the other methods, once again we are using training and test portions of country's data. For Brazil, first data to be analysed, we have the complete dataset ranging from 1991 to 2022, meaning 32 observations. Training test is 80% of the complete data, 26 observations from 1991 to 2016 and test set is the remaining information available, that goes from 2017 to 2022.

Using `NNETAR` on the training portion of Brazilian data, function identifies a model $\mathrm{NNAR(1,1)}$, indicating that data for Brazil has 1 lagged input to the network and 1 node in the hidden layer. A $\mathrm{NNAR(1,1)}$ model is a neural network with the last observation available used as main input for forecasting purposes using one neuron on the hidden layer. Although we do not have negative values on the retrieved data, on fitting this model we use a square root transformation to ensure that forecasted values stay positive in the future.

We perform an extrapolation by 30-year forecasts for unemployment in Brazil to understand how and if the $\mathrm{NNAR(1,1)}$ model captures well the cyclicality on the future when comparing with historical data available. To all effects our main forecast is still on the 10-years horizon, the illustrated-on figure 64 is just an initial assessment on how well the suggested model from `NNETAR` function deals with Brazilian data.

Figure 64
Forecasts from a Brazilian NNAR(1,1) model: 30-years extrapolation.
(Elaborated by the author).



From unemployment rate in 2016, last available information on the training set portion of Brazilian data, and 1 neuron in the hidden layer we produce the illustrative forecast presented on figure 64. In the extrapolated future, data appears to be somewhat around 10% of people outside labour force, considering mean values, while when observing the intervals projected, from 2017 and after, cyclicity apparent on historical data seems to repeat in the future.

Asymmetry of cycles appears to be well captured by Brazilian NNAR(1,1) model, increasing flows of unemployment are steeper than the decreasing part of historical data. NNAR(1,1) for Brazil has an average of 20 networks, each of which is a 1-1-1 network with 4 weights options were linear output units and the $\boldsymbol{\sigma}^2$ is estimated as 0.02.

Though we may presume that NNAR(1,1) model is reasonable considering the unemployment in Brazil, differently than the default on STL as well the ETS and ARIMA methods, neural networks are not substantiated on well-defined stochastic models, given that the technique is iterative in its essence (Hyndman & Athanasopoulos, 2021). This inherent process does not enable a straightforward prediction of intervals for the resultant forecasts, similar as the ones we produced using ARIMA or through the `hilo` function on ETS applications.

Nonetheless, prediction intervals can still be retrieved using simulation processes, where future sample paths are generated using bootstrapped residuals from the available data. Simulating these sample paths, we may build knowledge of the distribution for all future values based on the fitted neural network (Hyndman & Athanasopoulos, 2021), respecting our 10-year intended forecast horizon.

We proceed with a simulation of 3 possible future paths for unemployment rates in Brazilian data. Each of these paths covers the next 10 years after last observation available on the dataset (from 2016) and possible tracks unemployment in the country could evolve in the future. As the learning process is iterative, every 3 paths produced could generate different results on each iteration. We create and present one iteration and for illustrative purposes, we present potential paths on the following figure 65.

Figure 65
NNAR(1,1) forecast for Brazilian unemployment rates: Possible paths.
(Elaborated by the author).

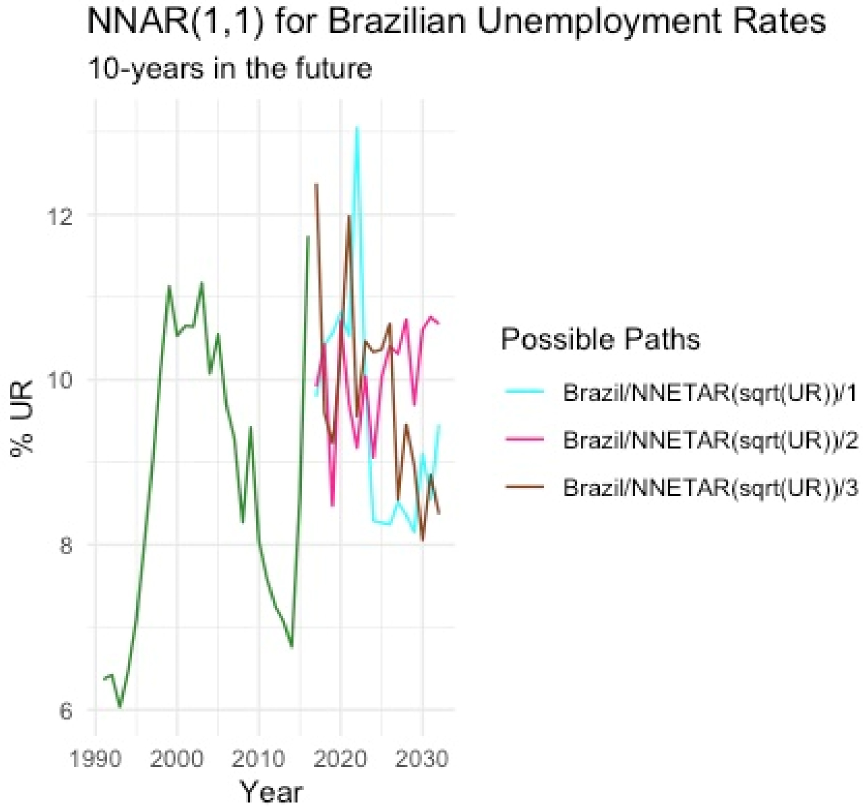


Figure 65 presents an ANN projection for unemployment rates in Brazil considering the remainder of data after the training portion and 10-years in the future, from 2022 (last data observation available) up to 2032. On figure 65, the simulation 1 is the one suggesting a more relatable scenario comparing with the real unemployment that occurred from 2017 to 2022, as the cyan line appears to capture well the behaviour of historical data and its cycles of peaks and lows.

Scenario with the second simulated path, presented on figure 65 by the pink line, seems to be the desirable one, given that does not peak high as the other two and the cycles of unemployment are more contained on the 8% to 11% range. Third possible path, presented by chocolate coloured line ha steeper highs and lows, lowering under when comparing observations from historical data, being a potential indicative that better scenarios could come to fruition if Brazilian entities dedicate some substantial efforts on labour market policies.

However, when using these simulated paths, as this process is iterative and performed using the `generate` function from `fable`, it is not possible to check for accuracy of the ANN method to compare with the other ones applied (STL, ETS, and ARIMA). Therefore, we proceed with the usage of the basic `forecast` proceeding and function, to perform fixed values using the Brazilian NNAR(1,1) model in a 10-year forecast horizon to later proceed with the accuracy checking.

Values comparison, similar as done previously on the ARIMA method, are presented next on table 85. We assess these results and present the accuracy values after the table.

Table 85
Forecasting Brazilian unemployment rates with NNAR(1,1) model: Values comparison. (Elaborated by the author).

Country	*Years*	*Real UR*	*NNETAR projected UR by mean*	*Difference projected to real*
BRAZIL	2017	12.93	10.50	minus 2.43
	2018	12.46	10.20	minus 2.26
	2019	12.05	10.10	minus 1.95
	2020	13.93	10.00	minus 3.93
	2021	13.34	10.00	minus 3.34
	2022	9.46	9.99	plus 0.53

Results projected by Brazilian NNAR(1,1) model when comparing with the real unemployment rates extracted from World Bank and ILO are relatively well approximate. Usually, the forecasted values are underestimating the real data by 2% more or less. Higher deviations are on the COVID-19 pandemic year and the following one (2020 and 2021), despite of it, in 2022 the projected seems to be more accurate, projecting slightly above the real unemployment that happened on 2022.

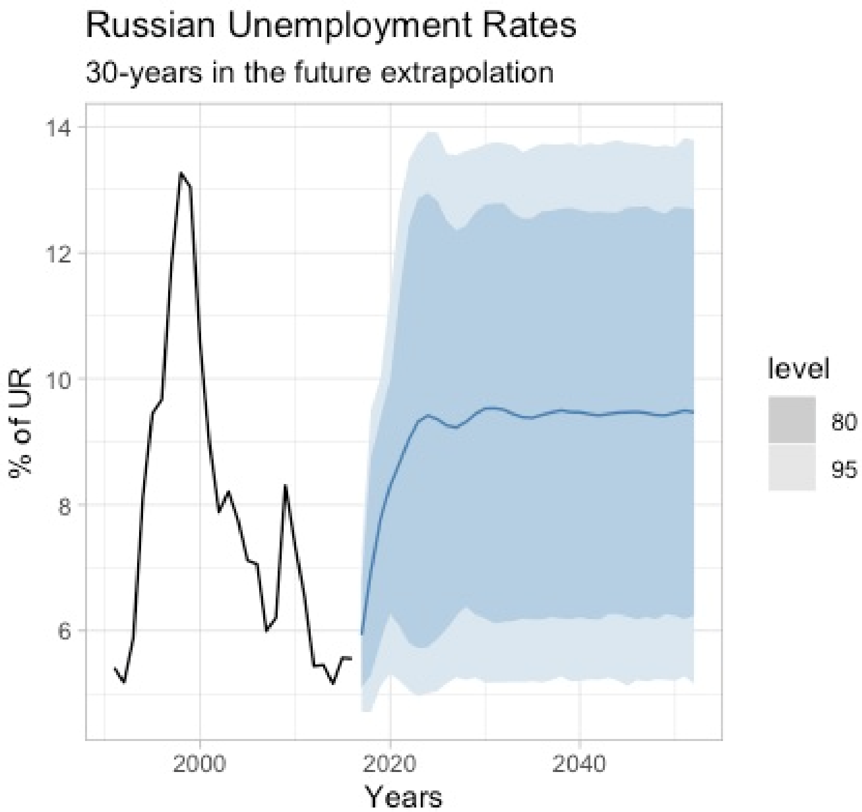
Indeed, neural networks for Brazilian unemployment rates in the future although fitted by a basic model, the NNAR(1,1), apparently captures well the potential labour market in the future for Brazil and the projected simulated paths we conjecture about it on figure 65. Still about figure 65, when having the mean values projected, the Brazil/NNETAR(sqrt(UR))/2 is the one most adjusted considering that unemployment percentages forecasted for after 2022 are all around 9%. As far as accuracy metrics, the following ones were obtained: Winkler score = 28.30, percentile = 1.89 and CRPS = 1.88.

Russian data is the next one to be analysed. We remember that Russia has one-year short of information, ranging on the complete dataset from 1991 to 2021, meaning that the test set of data is of five year observed value, ranging from 2017 to 2021 whereas after this portion tis the training dataset. `NNETAR` function applied on the training portion of data identifies a model NNAR(2,2), indicating 2 lagged inputs and 2 nodes in the hidden layer of Russian network.

NNAR(2,2) suggests for the retrieved data from Russia a neural network having the two last observations as the main input when forecasting and, also, that two neurons are necessary to stabilize the hidden layer on the network. Square root transformation of unemployment rates is used to assure that forecasted values remain positive on future years. For Russia this is particularly important, given that on previous methods (ETS and ARIMA) negative rates were estimated.

Russian NNAR(2,2) model has average of 20 networks, each of which is a 2-2-1 network with 9 weights options were linear output units and the σ^2 is estimated as 0.017. As for the Brazilian data, we perform an extrapolation by 30-year on the forecasts for unemployment in Russia to better understand the NNAR(2,2) model captures and replicates, or not, the Russian historical data. These extrapolated forecasts are presented next, on figure 66.

Figure 66
Forecasts from a Russian NNAR(2,2) model: 30-years extrapolation.
(Elaborated by the author).

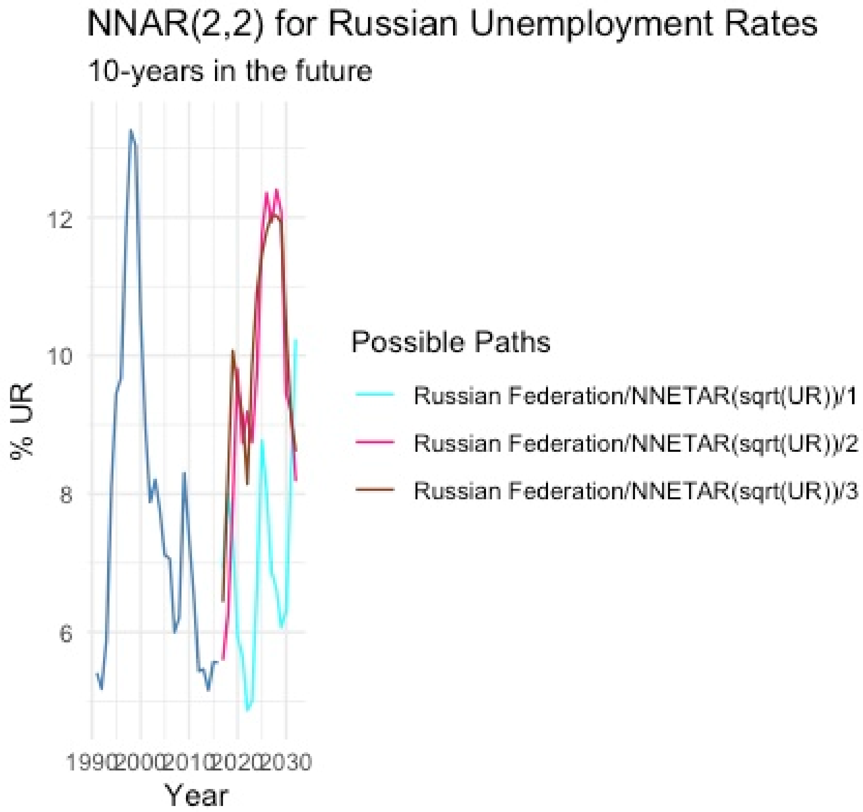


In the extrapolated future for Russian unemployment presented on figure 66, the data appears to replicate what happened in historical values observed of unemployment. Following an accentuate decline, that happened around 2000, some growing occurred on the mid 2010s years, only to precede a new decline that materialized after. Apparently, the visual extrapolation captures well the past Russian data behaviour, anticipating a coming crescendo of unemployment in the 2020s years to decline and a stabilization after this, by around 2040.

Russian unemployment asymmetries appear to be well captured by the NNAR(2,2) model indeed and from this assumption we move on for the potential future path propositions. As done with data from Brazil, we use a simulation process to conceive three future sample via bootstrapping the residuals of available data from Russia. Each of these path's projects the next 10 years after last observation available into an iterative procedure.

We remember that the usage of bootstraps and iterative simulations are not replicable and the `generate` function, to compose the scenarios forecasted, offers different results on each time it is applied. We create and present one of our iterations only for illustrative purposes on the following figure 67.

Figure 67
NNAR(2,2) forecast for Russian unemployment rates: Possible paths.
(Elaborated by the author).



ANN projection for unemployment rates in Russia illustrated on figure 67 presents two very similar scenarios, the ones presented by simulation 2 and 3 (lines on pink and chocolate colours, respectively), where projected unemployment rates seem to go slightly above 12% high-peak on Russian cyclicality, and as low as closing to 8% on minimal values predicted considering these two scenarios. Both pink and chocolate lines on figure 67 appear to be feasible with historical data, somewhat replicating what happened on Russia's past cycles of labour market.

Simulation 1, presented on figure 67 by the cyan coloured line, demonstrates the best-case scenario for Russian unemployment, that do not pass too far from 10% of unemployed people, at the beginning of 2030s years, and go under 5% on the 2020s years. Overall, considering the three simulated forecasts when comparing with the previous one made for Brazil, Russian unemployment is more stable and has capacity to cope with dysfunctions on labour even when unexpected shocks come to reality.

Nonetheless, we have some evidence to not disregard that unemployment could go to the bad side for Russian interested parties as politics, economists, and social scientists, for example. As the projections on simulations 1 and 2 suggested rates that are very like the peak of unemployment the country experienced on the start of 21st century. Meaning that policies of job allocation and labour policies must remain consistent to not repeat the indexes that occurred in the past.

Changing from `generate` to the `forecast` function, similarly as we done for the Brazilian case, we perform fixed forecasted values using the NNAR(2,2) model for Russia considering our 10-year in the future horizon. Comparison of values for the Russian ANN forecast is presented on table 86. Accuracy check results comes in the sequence.

Table 86
Forecasting Russian unemployment rates with NNAR(2,2) model: Values comparison.
(Elaborated by the author).

Country	*Years*	*Real UR*	*NNETAR projected UR by mean*	*Difference projected to real*
RUSSIA	2017	5.21	5.95	plus 0.74
	2018	4.85	6.95	plus 2.10
	2019	4.50	7.76	plus 3.26
	2020	5.59	8.25	plus 2.66
	2021	4.72	8.65	plus 3.93

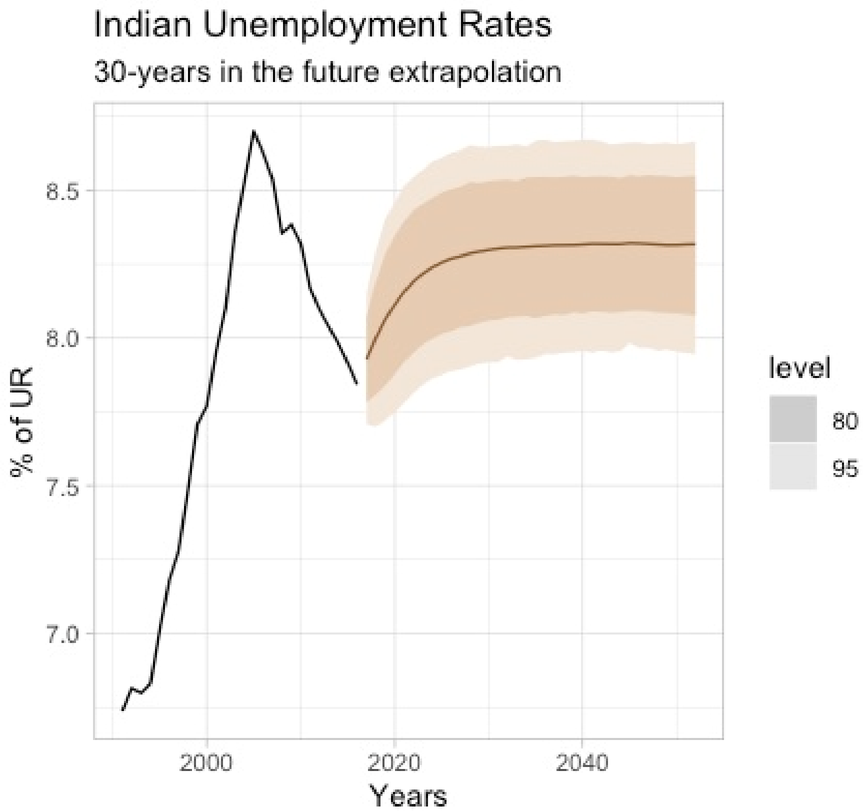
Russian NNAR(2,2) model, according to the results on table 86, consistently overestimates the forecasted values in comparison with the real unemployment rates. For all timespan that covers the test dataset for Russia (2017 to 2021), the projected values are above the real-life data and despite the 2017 years, every estimation is above 2% higher than the real unemployment in the country.

This overestimation may be an indicative that if one of the simulations projected and presented on figure 90 occurs, probably there would be something similar with the cyan projected line respective to simulation 1. Our `forecast` results indeed suggest a peak of 9.50% of unemployed people in 2031, there are no value higher than this on the 10-years projection. Accuracy checks for the NNAR(2,2) model from Russia is the following: Winkler score = 14.00, percentile = 1.82 and CRPS = 1.81.

Third of the five BRICS countries to be assessed is India. Indian dataset returns to 32 observations on the full dataset, ranging from 1991 up to 2022. Complying with the 80/20 proportion on dividing this data, the training portion, the one used for the forecasts, goes from 1991 to 2016 and the test portion, used for accuracy checking, ranges from 2017 to 2022. `NNETAR` suggests a model NNAR(1,1) averaging 20 networks each of them as 1-1-1 with 4 weights and a sigma2 ($\boldsymbol{\sigma^2}$) estimated as 0.0003.

Parameters are the same for Brazilian dataset, putting more weight on the last observation available before to project unemployment rates in the future. Also, there is only one neuron on the hidden layer of the Indian ANN. India's forecasts produced so far, using STL, ETS, and ARIMA does not suggested negative values for future rates, nonetheless, we perform the square root transformation to have a standard for all countries. 30-year in the future projection is performed here as well and these extrapolated predicted unemployment rates are presented on figure 68.

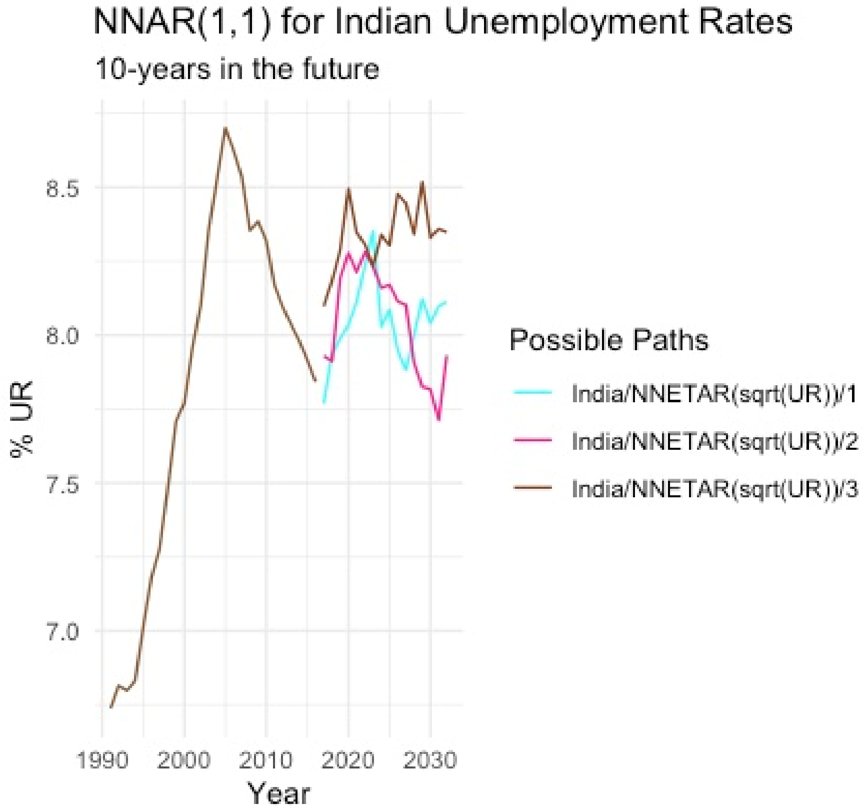
Figure 68
Forecasts from a Indian NNAR(1,1) model: 30-years extrapolation.
(Elaborated by the author).



Projected Indian unemployment presented on figure 68 appears to be more stable, on not oscillating excessively, if comparing historical data for the country's unemployed percentage over the years of data available and assessed. Nonetheless, figure 68 offers a suggestion that unemployment may not go as high as was around mid-2000s years, around 9%, and not as low as the starting of the 1990s, around 6%.

There is a slight suggestion by the of out 30-year extrapolation that Indian unemployment could start to augment, may be and indicative that if we extrapolate on a larger horizon, asymmetries from past values could repeat itself in the future. Despite that happening or not, we presume the NNAR(1,1) Indian model as suited and repeat the proceeding on propose potential future paths for unemployment in the country. This illustration is on figure 69, presented next.

Figure 69
NNAR(1,1) forecast for Indian unemployment rates: Possible paths.
(Elaborated by the author).



Simulation of future unemployment rates in India using the NNAR(1,1) model and presented on figure 69 are well convergent with the extrapolation previously presented on figure 68. For India, different that happened for Brazil and Russia where some simulations are very similar, Indian data generates three well defined good (simulation 2 on the pink line), bad (simulation 3 on the chocolate line) and medium (simulation 1 on the cyan line) scenarios.

Simulation 3, the bad one, does not indicate rates above 8.5%, not as a bad potential path as happened on historical Indian data around 2005. On the opposite, scenario 2, the best path indicated, does not present unemployment levels going lower of 7.5%, which is not so good as earlier percentages of unemployed in earlier timespan around 1990s years.

Considering the Indian economics context and the fact that they have one of the most populated territories in the world, maintain these levels of unemployment, even the bad scenario it is not necessarily a bad path to anticipate. Labour market efforts in the country are not to be underestimated however, given that as we perceived and discussed on about figure 69, it exists an avenue to the future do not be so stable as may appear.

Using the `forecast` function, we move from the simulation to a fixed Indian forecast considering the NNAR(1,1) Indian model and the intended 10-year in the future horizon. Comparison of real and estimated values are presented on the table 87.

Table 87
Forecasting Indian unemployment rates with NNAR(1,1) model: Values comparison. (Elaborated by the author).

Country	*Years*	*Real UR*	*NNETAR projected UR by mean*	*Difference projected to real*
INDIA	2017	7.73	7.93	plus 0.20
	2018	7.65	8.00	plus 0.35
	2019	6.51	8.06	plus 1.55
	2020	10.19	8.11	minus 2.08
	2021	7.71	8.15	plus 0.44
	2022	7.33	8.19	plus 0.86

NNAR(1,1) Indian model, according to table 87 presents an overall accurate estimation. Excepted by 2020, when there was an underestimation by more than 2% comparing forecasted value and real unemployment on that year, results seem well approximate with real life data we retrieved from World Bank and ILO repositories. After the unexpected COVID-19 effects on labour market, estimations return to be well adjusted in 2021 and 2022.

This consistency presented on table 87 gives us a suggestion that, between the observed country's data up to this point, the Indian ANN forecast may be the one that produces better results, suggesting that forecasted values for 2022 and beyond may very well represent a feasible Indian reality. Indeed, forecast considering this timespan are contained on above 8% and under 8.5% of unemployed people in the future, something alike the presented on the cyan coloured line on figure 69. NNAR(1,1) Indian model had the following indexes: Winkler score = 24.10 whereas percentile and CRPS both are equal to 0.82.

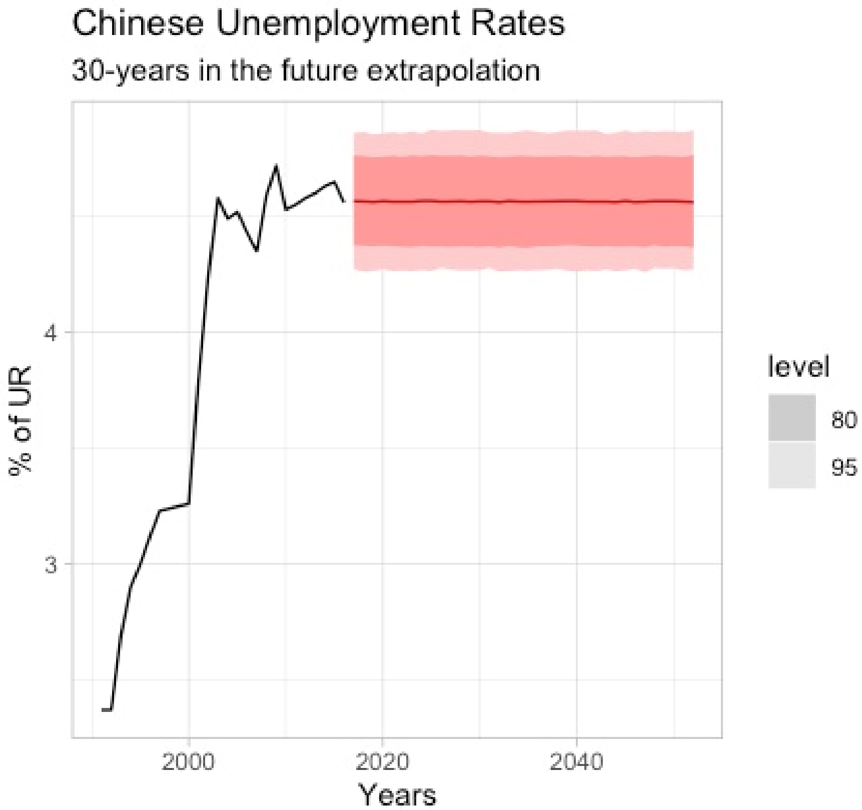
Chinese unemployment data, which comprises numbers from 1991 to 2022 on the complete dataset is the next one to be analysed. Using the range of 26 from 32 total observations we apply the `NNETAR` function to fit the ideal model to proceed with the forecasts. Remaining 6 available information are used later to check accuracy of projections made for unemployment in the Country after 2022 up to 2032.

`NNETAR` on the Chinese training set returns an ideal model with similar parameters with the Brazil and India data, a NNAR(1,1). For China's case the composition is by an average of 20 networks, each of them 1-1-1 types with 4 weights options and linear outputs. Sigma index ($\boldsymbol{\sigma^2}$) estimated is equal to 0.001. These features of the model suggest a neural network where the last observation of the training data is the main driver for forecasts and that Chinese ANN has on it one neuron on its hidden layer.

As done for Brazil, Russia, and India we present a visual extrapolation of 30-years in the future forecasts for unemployment in China as a first step to understand how NNAR(1,1) model captures the patterns on the future when comparing with historical

Chinese data regarding unemployment. This presentation is on the following figure 70 and we discuss about it in the sequence.

Figure 70
Forecasts from a Chinese NNAR(1,1) model: 30-years extrapolation.
(Elaborated by the author).



Cutting in 2016, last year and observation on China's training set, and projecting from thereafter, we may perceive on figure 70 that unemployment in China does not seem to extend itself above 5% at no point on the 30-years extrapolated future. Even when expanding the confidence interval from 80 to 95 percent, people outside the labour force on the country does not peak above the mean historical data neither as low as the beginning of data availability.

We may presume that NNAR(1,1) Chinese model reasonably captures the cyclicality of historical data of unemployment in the country despite extrapolating the forecast horizon, the model does not escape from general historical patterns. Although the proposition of predictive interval is not in the essence of ANN method, we may generate these using a simulation process that creates future sample from using bootstrapping residuals from the available data.

Similar as done for the other three countries analysed up to this point, we simulate three possible future paths for unemployment rates in China. Each path covering the next 10 of our intended forecast horizons plus the years included on the test set. This bootstrapped process being iterative may create infinite possible paths and for illustrative purposes we present just one iteration on the following figure 71.

Figure 71
NNAR(1,1) forecast for Chinese unemployment rates: Possible paths.
(Elaborated by the author).

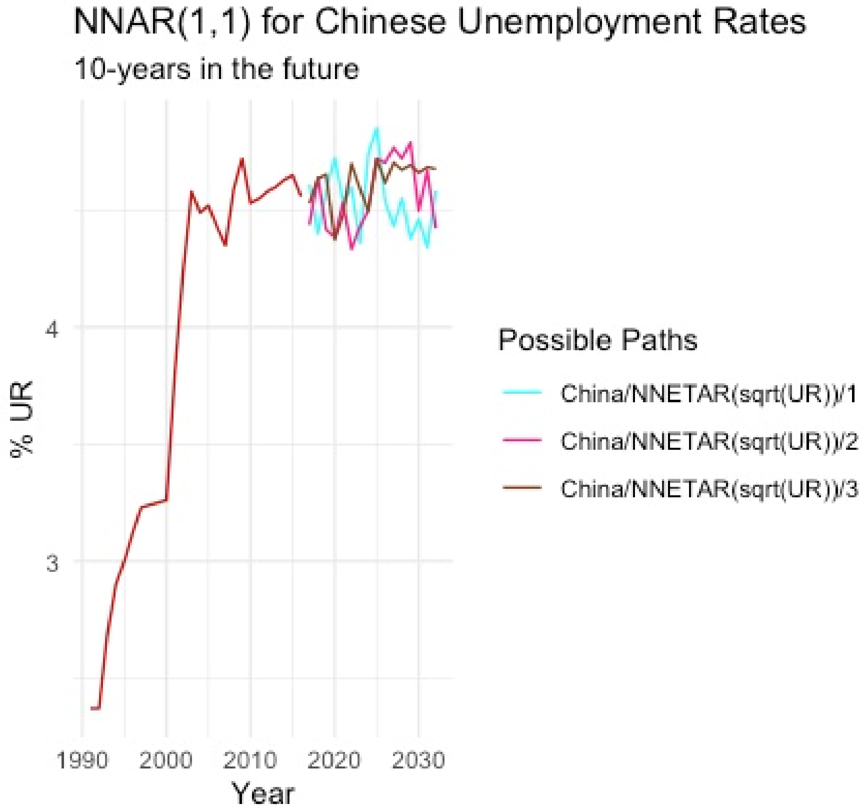


Figure 71 simulated paths resonates solidly with the extrapolation presented on figure 70. When the simulations have a comparatively peak on it a decline comes in the sequence whereas the opposite is true as well, if the projections are lowering in the beginning, they peak after. This behaviour of peaks and lows are well represented on the simulation 1 (coloured as cyan) and 2 (coloured as pink). Both simulations are alike historical data, especially on the mid 2000s up to 2016.

The simulation 3, the line coloured as chocolate, is more consistent with historical data that happened from the end of 2000s and in the early 2010s years, oscillating on peak and lows more rapidly. Important to notice that these referred crescendo and declines on Chinese unemployment rates, when comparing with other BRICS countries are very low. Percentage of unemployed people in the country does not go to 5 % and above and neither to 4% nor under.

Considering the BRICS context, no country appears to cope better with labour market policies than China. Even considering unexpected economic shocks, Chinese decision-makers dealt in a manner that in 2020, peak COVID-19 pandemic year, unemployment rates were by 5%, half-point above the previous year and in 2021 the rates fell this 0.5 point as well, being from 4.55%. Chinese politics and decision regarding labour are the ones to be referential for the other four countries within the group and outside as well.

We proceed with the usage of `forecast` function, to perform a new projection for unemployment using the Chinese NNAR(1,1) model. Prediction is to a 10-year horizon and later we apply some accuracy statistics. Before that, table 88 presents the values comparison between real unemployment and the forecasted data for this model.

Table 88
Forecasting Chinese unemployment rates with NNAR(1,1) model: Values comparison.
(Elaborated by the author).

Country	*Years*	*Real UR*	*NNETAR projected UR by mean*	*Difference projected to real*
CHINA	2017	4.47	4.56	plus 0.09
	2018	4.31	4.56	plus 0.25
	2019	4.56	4.56	equal
	2020	5.00	4.56	minus 0.44
	2021	4.55	4.56	plus 0.01
	2022	4.88	4.56	plus 0.32

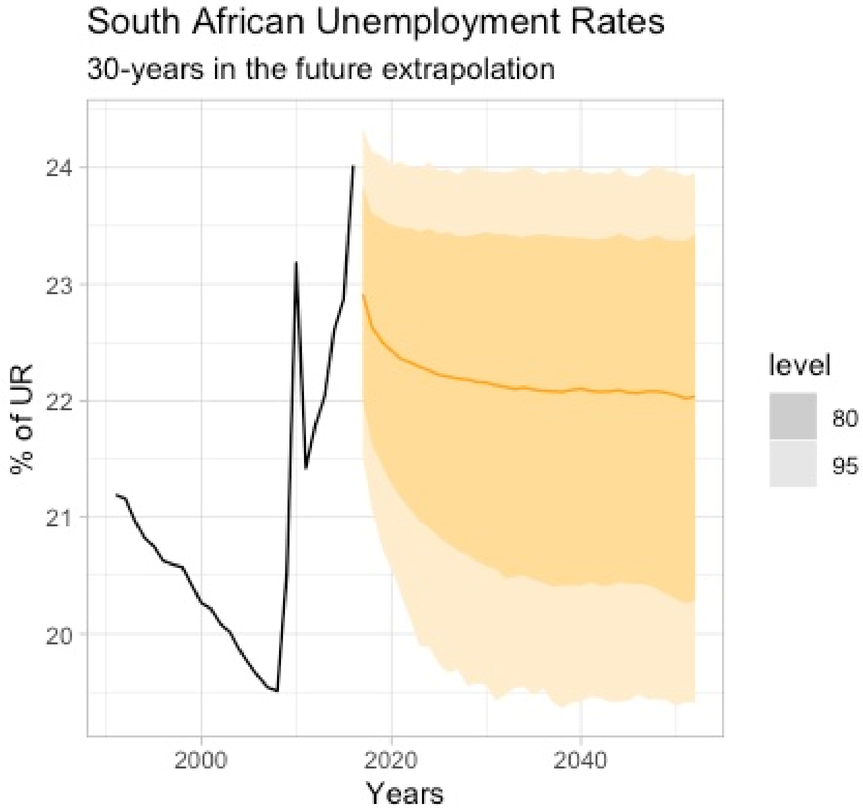
Chinese NNAR(1,1) model considering the comparison presented on table 88 is very well accurate. There is no index projected that misses the real unemployment by 0.50. Largest difference is in 2020 and is an underestimation only by 0.44 whereas the overestimation is in 2022 by 0.32. Chinese unemployment in recent historical data seems to be indeed very consistent in almost a linear setup that may very well replicate the results earlier presented on figure 71 into real-life. NNAR(1,1) model for China could not be complex in neurons or layers but captures very well what happened in Chinese past unemployment and could be a solid prospection about what may occur in the future. Accuracy metrics for the model: Winkler score = 1.61, percentile and CRPS are equal to 0.15.

Fifth and final country to be checked individually is South Africa. Same from Brazil, India, and China, South African full dataset comprises 32 observations from 1991 to 2022. Considering the 80/20 proportion, 26 observations (1991 to 2016) are on the training set and 6 remaining values are on test set (2017 to 2022). `NNETAR` function indicates NNAR(1,1) model, averaging 20 networks each of them as 1-1-1, having 4 weights and a sigma ($\boldsymbol{\sigma^2}$) values estimated as 0.005.

Model parameters are the equal for South Africa, Brazil, India, and China datasets, where (1,1) term indicates that last observation available is the main feature of the network on the projections of unemployment rates in the future. As was done for the other BRICS countries, the South African forecasts using ANN method is based-on a square root transformation to ensure that estimated values on the forecast will be only positive.

A 30-year in the future extrapolation is performed and presented on the following figure 72 considering the defined model according to the `NNETAR` function. We do not go deep on the values projected; our intention is to have a visual suggestion about how data can behave in a larger future than the one we intend to forecast for this thesis purposes.

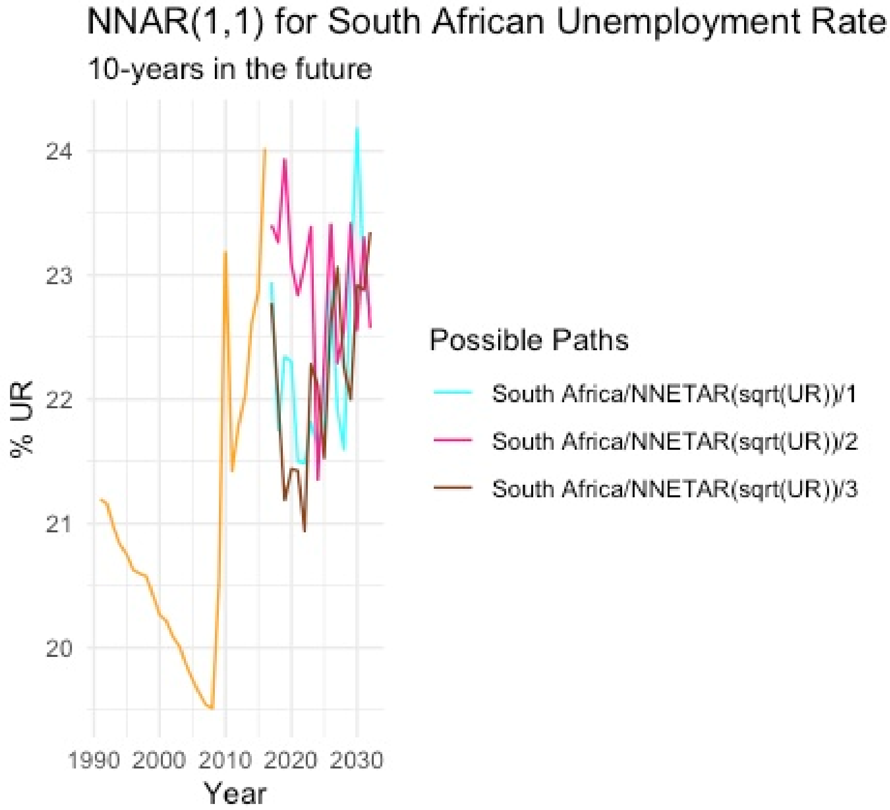
Figure 72
Forecasts from a South African NNAR(1,1) model: 30-years extrapolation.
(Elaborated by the author).



On South Africa we have the highest unemployment rates in all metrics and methods within the BRICS countries group and using ANN results are not differing. Figure 72 suggests that when observing the 95% confidence interval on projections the unemployment could go above 24% and under 19% by 2030 year and after. There it seems to appear a declining trend on the number of unemployed people in the country what could be a repeat on historical data for South Africa, where highs and lows are steeper on both ways.

Visual extrapolated forecasts presented on figure 72 seems to identify and replicate well asymmetries and the cyclicality of past values on South African unemployment, suggesting that the NNAR(1,1) model for the South Africa dataset maty be fitted. From this assumption, we proceed for the proposition of potential future path for unemployment in the country, same procedure applied on the previous nations analysed. This illustration is on figure 73, presented next.

Figure 73
NNAR(1,1) forecast for South African unemployment rates: Possible paths.
(Elaborated by the author).



All three simulated paths of unemployment rates in South Africa using the NNAR(1,1) model presented on figure 73 seems to replicate in an effective manner what happened with historical data. In fact, three simulation suggests very similar patterns in a broader sense. If we aim to define a good, a bad and an intermediate simulation, it is possible to define the South Africa/NNETAR(sqrt(UR))/1 as the worst, the one peaking above 24% of unemployed in 2030 while the South Africa/NNETAR(sqrt(UR))/3 is the among three simulation that reaches the minimum by under 21%.

On the three iterations we simulate, there is no scenario where unemployment reduces significantly in South Africa. Even the best of the three simulations, the chocolate-coloured line on figure 73 does not come close to lowest percentage of unemployed in the country that were by 19.5% around 2007 and 2008. On opposition, there is a path to the current rates may be even worsened, as projected on the cyan-coloured line.

We remember that these simulations are residuals and bootstrapped based considering the available data through an iterative process by `generate` function. Many other paths may exist, for better or worse, our intention is only to offer a glimpse on potential but not necessarily factual scenarios. Nonetheless, South African unemployment indeed is a historical problem and with ANN, as happened on the other applied methods as well, future does not give many indicatives of a better scenario.

South African government, private companies and even outside the country interested parties, like other BRICS countries for example, may be needed to an effective and collaborative effort to mitigate the nation's labour market dysfunctionalities. Incentives for self-employment, better allocation of resources, scholarship programs to cope with early unemployment and many other initiatives must be put in practice on anticipating what may come in a relatively nearby damaging future.

As the simulated process does not give a definitive result to accuracy measurements, we move for `forecast` function considering the NNAR(1,1) South African model and our intended 10-year in the future horizon. Comparison of real and estimated values are presented on table 89.

Table 89
Forecasting South African unemployment rates with NNAR(1,1) model: Values comparison. (Elaborated by the author).

Country	*Years*	*Real UR*	*NNETAR projected UR by mean*	*Difference projected to real*
SOUTH AFRICA	2017	23.99	22.90	minus 1.09
	2018	24.22	22.60	minus 1.62
	2019	25.54	22.50	minus 3.04
	2020	24.34	22.40	minus 1.94
	2021	28.77	22.40	minus 6.37
	2022	29.80	22.30	minus 7.50

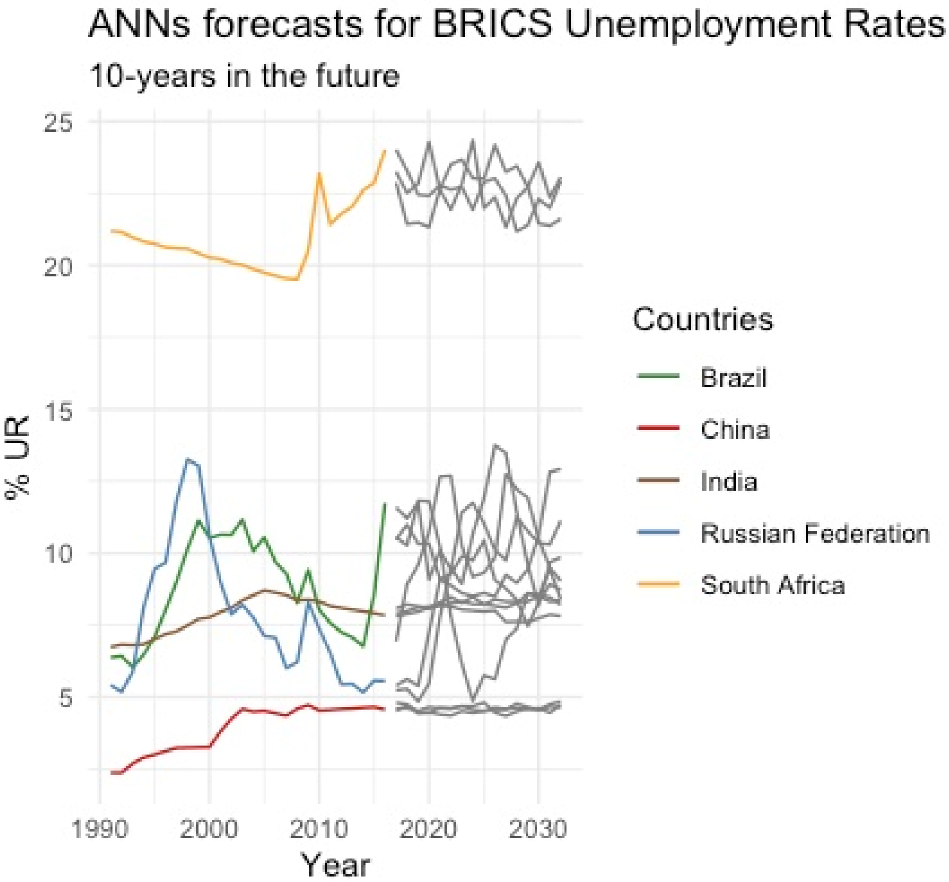
NNAR(1,1) South African model, according to table 89 presents potentially the worst accuracy in comparison with already analysed Brazil, Russia, India, and China data. All estimated values on the years that is covered by the test portion of data are underestimating real-life unemployment rates. Particularly in 2021 and 2022 the projected unemployment misses the real values by more than 6% and 7% respectively.

Results presented on table 89 gives the impression that unemployment in South Africa have been growing rapidly and at alarming rates on the following of COVID-19 pandemic years. 2021 and 2022 are largely affected by what happened in 2020 regarding labour market and we do not have many evidence suggesting that the scenario on the future could be a better one. If projection by ANN method continues to miss real unemployment by the values in table 89, unemployment could go higher than 30%. NNAR(1,1) South African model presents the following indexes: Winkler score = 85.80, percentile = 3.12 and CRPS = 3.11.

At last, as performed on STL, ETS, and ARIMA methods, we produce some assessments of ANN for the BRICS as a group. Bearing in mind that the complete timespan of observed values of unemployment rates covers 1991 to 2022 years for Brazil, India, China, and South Africa whereas Russia has one-year shortage of data. For all five countries when analysed individually, and considering them as a BRICS group as well, the proportionality of the complete dataset is divided into an 80/20 for training and test sets.

`NNETAR` function suggests NNAR(1,1) models for Brazil, India, China, and South Africa data. Only Russia differs the model indicated, being an NNAR(2,2) model. Projected simulated paths for future of unemployment for these nations, replicating the country-by-country analyses, is presented on the following figure 74.

Figure 74
ANNs forecasts for BRICS countries' unemployment rates: Possible paths.
(Elaborated by the author).



We remember that these paths are simulate three times for each country considering their own availability of historical data. Having countries' information in hand we used the `generate` function for simulate these paths presented on figure 74 by an iterative proceeding that uses bootstrapping and residual values for every one of the five BRICS nations. As a bootstrapped is performed, every iteration generates different results and, therefore, different possible paths. Probably, the simulations presented on figure 74 are not necessarily the same we present when observing each country; however, we use the figure for illustrative purposes of our forecasts.

These simulations are useful to perceive the extremes on the BRICS: South Africa at the top and China in the bottom of the plot, having the highest and the lowest rates of unemployment. As well the behaviour of Brazil, Russia, and India, historically are relatively apart from each other in some manner but may come to a potential future with more similarities between themselves.

Descriptive statistics for the forecast using ANN method are presented on the following table 90. We present data for each country that were built upon their historical data, test sets, and on 10-year in the future forecast horizon (2023 up to 2032). After that, accuracy of the method was performed through Winkler score, percentile, and Continuous

Ranked Probability Score. We remind you to beyond to compare the presented next on table 91 regarding accuracy measurements table with the previously presented in the table 62, where are the accuracy on these same statistics are depicted considering the STL and ETS methods.

Table 90
ANN forecasts: BRICS unemployment rates descriptive statistics.
(Elaborated by the author).

Country	*Min.*	*1st Qu.*	*Median*	*Mean*	*3rd Qu.*	*Max.*
BRAZIL	9.88	9.89	9.89	9.99	9.99	10.62
RUSSIA	5.99	8.52	9.15	8.70	9.25	9.45
INDIA	7.93	8.14	8.24	8.20	8.28	8.30
CHINA	4.56	4.56	4.56	4.56	4.56	4.57
SOUTH AFRICA	22.14	22.18	22.26	22.33	22.39	22.90
BRICS	**4.56**	**7.98**	**9.15**	**10.75**	**10.00**	**22.90**

Results on table 90 complement well all the individual analyses and the illustration on previous figure 74. China is stable on the less damaged by unemployment and whereas South Africa does not go under 22% not even in the minimum value projected. Observing BRICS countries collectively the median and mean values are around 10%, reasonable that Brazil, Russia, and India in the middle compensates the highs and lows by other two countries.

Overall, the Artificial Neural Network method produced forecasted values, consistent with other techniques we applied (STL, ETS, and ARIMA). Illustrations presented on all figures that uses simulation, even considering the iteratively nature of the ANN method produced similar results in different iterations. To statistically assess the accuracy of ANN, we check and present results on the following table 91.

Table 91
Accuracy checking for the ARIMA and ANN forecasts: BRICS unemployment rates.
(Elaborated by the author).

Country	*Forecast Model*	*winkler*	*Percentile*	*CRPS*	*Forecast Model*	*winkler*	*Percentile*	*CRPS*
BRAZIL	**ARIMA (0,1,2)**	9.67	1.03	1.03	**NNAR (1,1)**	28.30	1.89	1.88
RUSSIA	**ARIMA (0,1,2)**	9.55	0.63	0.63	**NNAR (2,2)**	14.00	1.82	1.81
INDIA	**ARIMA (2,1,0)**	17.90	0.63	0.63	**NNAR (1,1)**	24.10	0.82	0.82
CHINA	**ARIMA (1,1,0)**	1.79	0.16	0.16	**NNAR (1,1)**	1.61	0.15	0.15
SOUTH AFRICA	**ARIMA (1,0,0)**	61.60	2.59	2.59	**NNAR (1,1)**	85.80	3.12	3.11

On table 91 we have accuracy checking for ARIMA and for the ANN method. We compare these two methods presented with was previously depicted on table 62, where was the accuracies for STL and ETS forecasts. Summarized on these two tables, 62 and 91, we may compare which of the four applied methods are performing better for each country, meaning the one having lower values on Winkler score, percentile and

CRPS indexes. The lower it is, minimised are the errors on the forecasts using each of these methods.

On specific to ANN usage, as this is the method we are applying on this topic, the technique does not perform better in comparison with any of the other three methods. Not even for the Chinese data, the country having lower values than the other four of the BRICS group when using ANN, the results are not good. STL forecast and ETS, for China's case, presents lower percentile and CRPS indexes (0.14 on both while on ANN is 0.16) and the Winkler score using ANN is by 1.61, which is better than the ARIMA application but worse than both STL and ETS.

We clarify that we do not believe that the Artificial Neural Network is a non-reliable method for forecasts. In fact, as we present earlier on this chapter, machine learning techniques like ANN are improving future studies in some areas and may be the next advance on the field. Maybe, as we are dealing with a relatively small sample of data, the method does not have many inputs to generate networks more complex, and, therefore, more robust forecasts for our case.

ANN is also in its essence an iterative method, if we perform the technique other times, we will have different results that may improve or even worsening the results we obtained and presented on the presented results here on this topic. Although we have run the procedure many times to perceive the consistency on what we have shown so far, if anyone reading this try to run the technique by itself may have distinct outputs. This is a caveat and limitation of the method important to be acknowledged but we do not believe that turns our results as unreliable or insignificant.

If we are analysing other macroeconomic variables, that has an extensive amount of data as GDP or inflation, for example, results using ANN to forecast could portrait a more effective scenario than the one we found when analysing unemployment rates in five countries with, at maximum, 32 observations of data available on the most consolidated repositories about the theme.

Nonetheless, we believe that the Artificial Neural Network applications still offer some reliable outputs and relevant insights for possible scenarios that unemployment rates could evolve in Brazil, Russia, India, China, and South Africa realities. Also, as we mentioned on the ARIMA application, we do not aim to unveil a best overall method to forecast. On opposition, we decide to use different approaches to identify positives and negatives on all of them.

One methodology could be a better fit for one country while at the same time it could perform poorly for another one. Given that we are analysing a group of countries, the BRICS, this mixed scenario is only natural to occur. It could even be the case that combining methods a more accurate forecast could be generate, offering a better adjust potential to foresight future values than the usage of just one technique in isolation. About this combination of proceedings, specifically the ones we present up to this point, the STL, ETS, ARIMA and ANN we present our findings on the next topic.

4.4.5. An application of combined forecasts.

We proceed this topic under the premises discovered on the previous application. There is no method, at least among the four performed here, that is well adjusted and a better fit for all countries. Having no method above all, the idea is to combine them to have an idea if together they present a minimisation of errors on the forecasting process and, as consequence, it improves our estimations of unemployment rates in 10-years into the future forecasts.

This topic is based on the chapter 13.4 on Hyndman & Athanasopoulos (2021) Forecasting: Principles and Practice book, where they suggest that the usage of different methods in a same timeseries to average resulting forecasts as a manner to improve accuracy from the isolated applications. Literature (e.g., Bates & Granger, 1969; Clemen, 1989) offers some reason to believe that indeed combined forecasts may lead to a better performance on accuracy measurements.

More recently, Wang et al. (2023) developed a 50-year literature review on studies and areas that uses forecast combinations. Authors point to potential explanations to why the combining process not rarely leads to stronger performances, such as the incorporation of partial information coming from different methods and the possibility to mitigate structural breaks or other instabilities emerged from one specific technique or any data observer specificity.

Future values to be forecasted are unknown or, by the defined-on Wang et al. (2023), a "meta-parameter". If all individual forecasts and methods, as we used here doing the Seasonal and Trend Decomposing using Loess, Exponential Smoothing Technique, Autoregressive Integrated Moving Average, and Artificial Neural Network techniques, are providing estimates, it seems reasonable that averaging these estimations may offer the indicative to provide a new and potentially improved estimates (Wang et al, 2023).

There has been an extensive range of research where combined forecasts uses weighted averages, or other more complexes approach (i.e., time-varying weights, nonlinear combinations, correlations among components, and cross-learning), there are not many combinations that are significantly outperformed by simple average of forecasts performed individually (Hyndman & Athanasopoulos, 2021; Wang et al., 2023). From this premise we may start to build our combinations going forward.

Methods we selected to use (STL, ETS, ARIMA, and ANN) were already discussed on the methodological segment and at the beginning of each applications rounds we proceed on the previous four subtopics of this results and discussion section; therefore, we do not replicate to overextend on explanations about them and the idea of combination of forecasts is indeed simple as it seems.

We follow Hyndman & Athanasopoulos (2021) scripts taking in consideration our data characteristics and the results comparing the ensembled forecasts with the other four methods are compared at the end of it. Considering that our study is not focusing on a country in specific but considering a cross-country analysis within the scope of BRICS countries, we do not perform a country-by-country analysis as it was done on the methods used in isolation.

Main premise on this topic proposition is to simultaneously automatize forecasting techniques that were applied individually. Using each of STL, ETS, ARIMA, and ANN outputs intention is to conceive a combination of these four methods and assess if a combining of these performs better than those individual applications. If this happens to be the case, as occurred on other studies in line with this premise (e.g., Bates & Granger, 1969; Clemen, 1989; Hyndman & Athanasopoulos, 2021; Wang et al., 2023), the combining output will be the referential for the later presented scenario forecasting.

If our data analyses and results when combining forecasts does not suggest a better performance in comparison with methods used on isolation, we intend to use as basis the methodological proceeding that presents better results among the four methods. Before to proceed with this assessment, some features about this combining process are necessary to be cleared up before to present our results.

First, data by the five countries on BRICS group are combined into a single "Aggregate unemployment" variable, that sums all countries unemployment rates individually. We divide this new variable on training and test portions of data. Similar process as the other applications performed earlier. Results of this combined variable are presented next, on the following table 92.

Table 92
Unemployment rates on BRICS: A combining aggregation.
(Elaborated by the author).

Variable	*Year*	*% value*	*Year*	*% value*
AGGREGATE UNEMPLOYMENT	1991	42.08	2007	47.70
	1992	41.94	2008	46.93
	1993	42.37	2009	51.33
	1994	45.16	2010	51.43
	1995	47.30	2011	48.26
	1996	48.63	2012	47.15
	1997	51.91	2013	47.20
	1998	54.71	2014	47.14
	1999	55.55	2015	49.56
	2000	52.40	2016	53.72
	2001	51.60	2017	54.33
	2002	50.95	2018	53.49
	2003	52.33	2019	53.16
	2004	50.71	2020	59.05
	2005	50.63	2021	59.09
	2006	49.44	2022	51.48

Having as referential these combined values presented on table 92, that aggregates unemployment rates considering Brazil, Russia, India, China, and South Africa, we may move on to forecast future values having these ones as basis. We may see observing the depicted on the table 92 how well aggregate unemployment are capturing economic shocks that happened on the world during the timespan of data covered.

To illustrative examples of what may be captured by the aggregate unemployment presented on table 92, we see that after 2008 crisis (Lee, Schmidt-Klau & Verick, 2020), unemployment spiked on 2009 and 2010, the same happening after the COVID-19 pandemic outbreak, where percentage of unemployed peaked high in 2020 and 2021. We present the forecasts using each of the four methods plus the combination on the following figure 75.

Figure 75
Unemployment rates on BRICS: A illustrative presentation.
(Elaborated by the author).

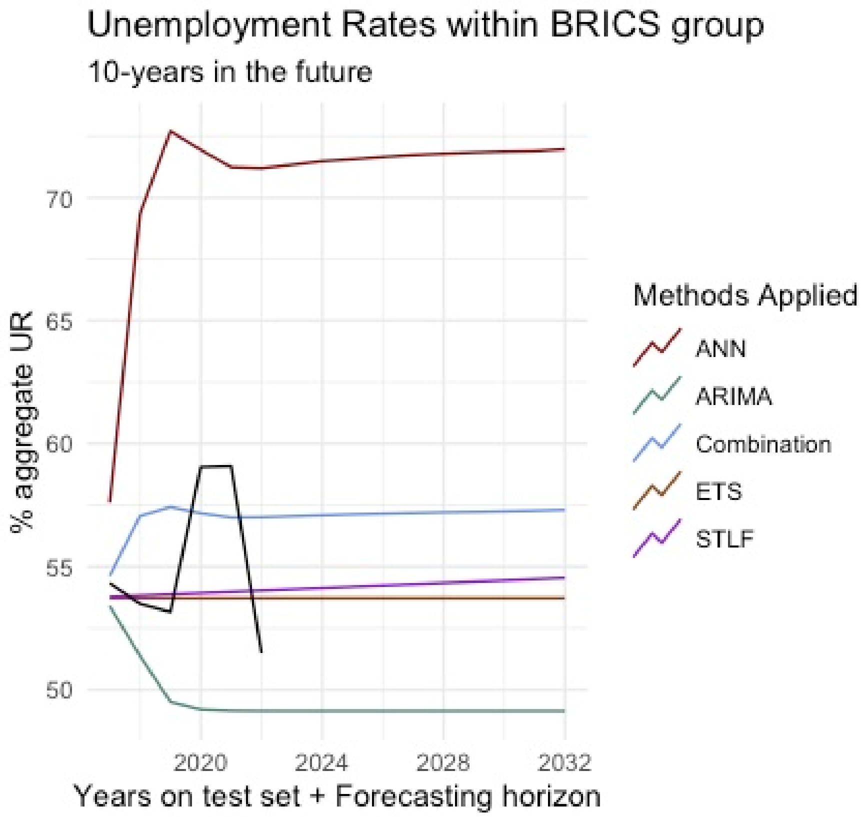


Figure 75 presents five lines on the plot. Coloured ones are respective to each method (STL, ETS, ARIMA, and ANN forecasts) whereas on the black line is the real observed unemployment rates that were extracted from World Bank and ILO repositories and presented as well on table 92. Worth of remembering however that ANN here, as happened for the method applied in isolation on topic 4.4.4 present an iterative forecast, different application of the method would generate distinct results, we are portraying one iteration if others run the same technique, with the same data, results would differ.

ETS suggests almost a straight line with mean values, STL forecast indicates a slight increasing on unemployment starting on the second half of 2020s years. ANN line on figure 98 is positioned at the top of the plot, suggesting unemployment that may close to 75%, considering the five countries, while the ARIMA line is the one having the lower values, under 50% at the bottom of figure 75.

Accuracy measurements of the forecasts visually demonstrated on figure 75 are presented on table 93. On accuracy measurements and proceedings, we perform a several of them and we have been using as a rule of thumb a premise of the less, or minimal, is better (Hyndman & Athanasopoulos, 2021). Meaning that the lower values on the metrics are the ones that may provide more assertive forecasted values. We strongly recommend the reading of Makridakis (1993), where is discussed concisely the importance of these metrics on forecasting purposes.

Table 93
Accuracy from BRICS models: STL, ETS, ARIMA, ANN and Combination.
(Elaborated by the author).

Method	***RMSE***	***MAE***	***MAPE***
ANN	15.00	13.90	25.50
ARIMA	6.06	4.81	8.44
ETS	3.24	2.39	4.18
STL	3.16	2.40	4.21
Combination	3.40	2.94	5.45

Results on table 93 presents that ANN method performs particularly bad on our sample and purposes. This may be due we have been dealing with a relatively small number of observations for a robust technique that is anchored on iterations and probably would perform more accurately with a larger sample. We remember however that ANN is a machine learning iterative process, other applications may have slightly different results for better or worse. ARIMA as well does not perform well in comparison with the other options, being the second worst of the applications used.

On opposition, ETS and STL are very alike overall. Difference on the indexes of these two are not too apart from each other and the better results suggesting by MAE and MAPE are slightly better for Exponential Smoothing Technique. RMSE number is lower on Seasonal and Trend Decomposing Using Loess, 3.16 whereas on ETS is 3.24. Differences among STL and ETS are not accentuated indeed, as table 93 indicates and we may also perceive on the approximate values on MAE and MAPE indicators on both methods, being the ones on ETS slightly better.

Combination of methods does not perform bad as ANN or ARIMA but as well does not return better numbers than ETS and STL. Combining methods presents itself as a midterm, where its results, at least on this first analysis, are not to be disregard considering that it already outperforms one of the recurrently used forecasting techniques, the ARIMA, and the expanding on usage method, the ANN.

Still following the Hyndman & Athanasopoulos (2021) proceedings we replicate the simulation process that was used on the STL and ANN forecasts, where we simulate some potential scenarios for the future of unemployment on BRICS countries. The idea is to create intervals where aggregate unemployment on the group may reside in the future considering the 80 and 95 percentages of confidence.

Simulating process as we mentioned when applying the ANN technique is iterative. Potential paths are simulated iteratively using residuals based on available data and bootstrapping procedures. Therefore, anytime the running of these simulations is performed different results may appear. We proceed and present one simulation just to perceive if the forecasts illustrated on figure 75 are within the range of potential paths.

Replicating a similar process as on Hyndman & Athanasopoulos (2021) textbook, we generate 1000 future sample paths and compute distributional forecasts based on these one-thousand possibilities. We enlarged the potential paths for unemployment to assure that we have a good as possible outcome considering the iteratively nature on this proceeding. Following figure 76 presents our simulated paths, considering an interval confidence by 80 and 95 percentage. After this visual illustration we present on table 94 the accuracy checking of our simulations.

Figure 76
Simulated paths for aggregate unemployment rates on BRICS.
(Elaborated by the author).

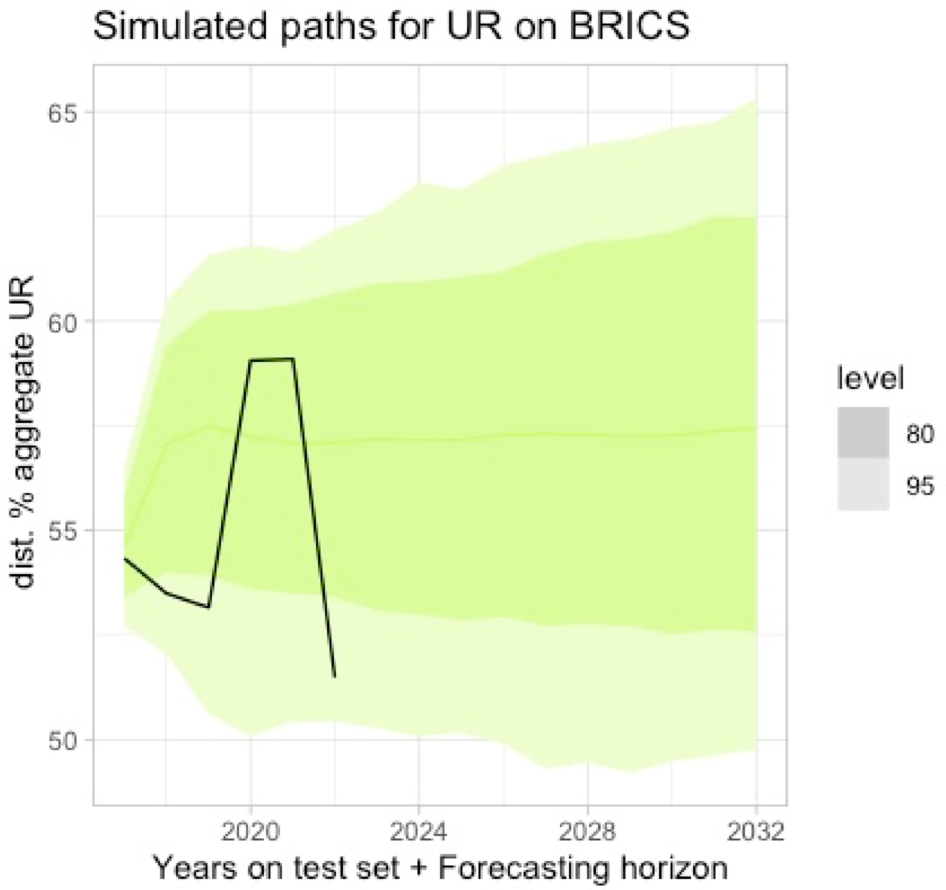


Table 94
Accuracy from BRICS models: STL, ETS, ARIMA, ANN and Combination simulations.
(Elaborated by the author).

Method	*winkler 80%*	*winkler 95%*
ANN	25.10	26.40
ARIMA	25.10	49.00
ETS	10.50	15.80
STL	11.00	16.80
Combination	11.10	9.67

On figure 76 we may see that the intervals described on the simulation covers relatively well the presented on previously presented figure 98. At the top of figure 76 we see a potential growing on the end of the 10-year forecast horizon that may implicate that eventually could reach what predicts the ANN forecast from the figure 75. On the opposite, bottom of the figure is essentially replicating the ARIMA projected values around 50% on aggregate unemployment rates.

Accuracy checking results presented on table 94 is in line with table 93. ETS that on table 93 performed better on MAE and MAPE metrics on the 80% interval of confidence is slightly more accurate than STL and the combining method. However, as presented in the table 94, expanding up to the 95% confidence interval, the combining methods significantly outperforms the others when individually applied.

All things considered we have evidence to believe that the combining method procedure that aggregates unemployment and forecasts for the ANN, ARIMA, ETS, and STL methods, bearing the intentions to suggest a BRICS collective into the future scenario is feasible to work as main referential for propositions about behaviour of unemployment rates in the group. Therefore, when the five countries are to be analysed as a collective (the BRICS), the combined method will potentially offer solid results.

On the other hand, for countries individual assessments, each of them is more suited with one of the four methods applied solely. Russia and China both have better forecasting performance on the Exponential Smoothing Technique (ETS) and Seasonal and Trend Decomposing Using Loess (STL) methods respectively, considering all accuracy checking we apply: Winkler score, the percentile, and the Continuous Ranked Probability Score (CRPS).

Brazil and South Africa have similar results, which may suggest a strong relationship between these two within the group. Their performance is better using ETS considering percentile and CRPS indexes whereas the Winkler score is lower, minimising errors, in STL. We assume ETS as better method for both countries having in mind that exists two better indexes in this method when comparing with STL.

India is the country with more mixed results. Autoregressive Integrated Moving Average (ARIMA) and ETS have same results for percentile and CRPS indexes while the Winkler score, similar as happened for Brazil and South Africa is better on STL technique. As apparently general STL in general tends to suggest lower values for Winkler metric for most of the countries analysed, we use decimal cases to decide in favour of ARIMA as the best method for Indian data.

Artificial Neural Networks (ANN) does not appear as the most suited method for none of the five BRICS countries and performs poorly on the aggregation of methods also, as we could perceive in this topic when the comparison is made with the other techniques on figure 75 and on both tables 93 and 94. Nonetheless, we do not disregard the method inputs at all, the simulating process we perform using ANN were particularly insightful due the range of potential paths we discover on the simulated scenarios.

In the end, every method gives their own contribution as we presume by the option to perform a multimethod approach. Specificities by each of the methods are useful on building up knowledge that we need to proceed with the scenario forecasting proposition. As we delimited early on this chapter, the idea is to perform a cross-country analysis, particularly within BRICS context, therefore, we intend to rely significantly on the combined methods results to our scenario forecasting proceeding that will be performed on the next topic.

Country's level analysis and ANN, ARIMA, ETS, and STL forecasts applied on isolation will not be completely put aside, when some specific country's suggestion on the scenario forecasting is to be made, we rely on the technique this nation have a better performance on errors minimization. ARIMA to India, ETS for Brazil, Russia, and South Africa and STL forecast for China. Again, we recommend the checking of tables 62 and 91 where the accuracy measurement for each method is presented with more detailing.

Before the scenario forecasting proceedings, for descriptive purposes, table 95 presents the forecasted values obtained by the usage of combination method. These results are complementary on table 92, where aggregate results are presented considering the data retrieved for World Bank and ILO repositories. Table 95 and its results are the foundation for our scenario forecasting that will be presented on the sequence.

Table 95
Unemployment rates on BRICS: Forecasting values by combining methods.
(Elaborated by the author).

Variable	*Year*	*ARIMA forecasts*	*ANN forecasts***	*ETS forecasts*	*STL forecasts*	*Combination forecasts***
AGGREGATE UNEMPLOYMENT*	2017	53.40	57.60	53.72	53.77	54.62
	2018	51.37	69.35	53.72	53.82	57.06
	2019	49.50	72.70	53.72	53.87	57.42
	2020	49.19	71.95	53.72	53.92	57.17
	2021	49.14	71.23	53.72	53.98	57.00
	2022	49.13	71.19	53.72	54.02	57.01
	2023	49.13	71.33	53.72	54.08	57.05
	2024	49.13	71.48	53.72	54.13	57.09
	2025	49.13	71.56	53.72	54.18	57.13
	2026	49.13	71.65	53.72	54.23	57.15
	2027	49.13	71.73	53.72	54.28	57.18
	2028	49.13	71.78	53.72	54.33	57.21
	2029	49.13	71.83	53.72	54.38	57.22
	2030	49.13	71.86	53.72	54.44	57.25
	2031	49.13	71.89	53.72	54.49	57.27
	2032	49.13	71.97	53.72	54.54	57.31

* Aggregate unemployment variable comprises a timespan from 2017-2022. This indicates the 6 years that composes the test portion of the dataset plus 10-years into the future forecast horizon from 2023 up to 2032.
** ANN forecasted values presented on table 95 are respective to one iteration we perform to compose this table. As the method is iterative by machine learning, different iterations may present distinct results. Considering that ANN forecasts may change the combining of the four methods would change as well.

ANN forecasts have the most elevated numbers, closing to almost 72% of aggregate unemployment on the five BRICS countries whereas ARIMA goes as high as 53%. ETS project a straight mean value by 53.72%. STL and combination of the four methods are the ones that present some variation lower values being around 54% and higher by 57%. Results on table 95 are in line and replicating what we illustrated on figures 75 and 76 the caveat imposed by the ANN usage where results may differ on different rounds of applications, and we are basing our analysis at one iteration applied.

From table 95 results, the discussed so far on regard the combining procedures as well on the previously presented from individual methods applications, we move forward to what may be understood as one of main contribution proposed by this thesis, the scenario forecasting for unemployment in 10-years into the future of BRICS countries presented on the next topic.

We remember that these scenarios condense all the before developed on this thesis, the literature review about unemployment, the exploring of potential determinants that composes unemployment rates and the multimethod proceeding we developed on this chapter. All proceedings established are basing our scenario building as we hope to propose results-based scenarios that goes beyond speculative thinking or mere conjectures, believing that all things we put effort on it are solidifying everything that comes on the sequence.

4.4.6. Scenario forecasting.

Scenarios are to be plausible and coherent descriptions of possible future situations that may be used for strategic planning, decision-making, and exploring alternative managerial procedures (Schaars, 1987). Scenarios are valuable tools for navigating on a volatile context, where uncertainties are high or there exists a plenty of factors that affects a main subject of interest, as is the case for unemployment phenomenon identifying throughout our study.

Still according to Schaars (1987), scenarios approach on despite the technique they are developed may provide a well-structured method to conjecture-thinking about alternative futures. Scenario's proposition involves several key steps that goes from scope identification, factor analysis, storyline creation, and scenario refinement. The method can be applied into various fields, including the ones we insert our ongoing study, the business, government, and policy planning (Schaars, 1987) analyses.

We acknowledge that in its essence the scenario building is a purely qualitative approach although some quantitative efforts, as the one we intend to propose here, have being incorporated on the field as the likes of trend-impact and cross-impact analyses (Bunn & Salo, 2003). Trend-impact and the forecast-adjustments may be understood as a type of bridge between the quantitative proceedings that we applied so far on this thesis and the qualitative approach we about to perform from hereafter.

Before that, we present some arguments on Bunn & Salo (2003) that correlates with our premises on the usage of scenarios assuming an econometric and quantitative orientation. Trend-impact analysis, in our case the potential tendencies of unemployment 10-years into the future for BRICS countries, commonly relies on many inputs to unveil tendencies (Bunn & Salo, 2003). We are relying on our multimethod approach that combines outputs from ANN, ARIMA, ETS and STL forecasts to have a glimpse on how rates of unemployed may grow or decline on our proposed scenarios.

There is however a caveat on this effort on using a subjective approach grounded on quantitative procedures. Bunn & Salo (2003) alerts to the possibility that previous mathematic results may deceive researchers to believe that they fully anticipate or mastered uncertainties within the environment analysed. This is the reasoning about we intended to expand our data inputs as much as possible. Doing that, even in a relatively short timespan (1991-2022) we covered two major economic crises, in 2008 and 2020.

The insertion of these unexpected shocks offers to our working data some level of robustness to peaks on unemployment and to perceive the time countries observed, Brazil, Russia, India, China, and South Africa, needed to recover some stability on their respective labour markets. We do not believe nor presume to design what will effectively occur 10-years from now; however, we credit our efforts on avoid biases and minimise errors through all four individual forecasting techniques we performed before the proposal of our scenarios.

Forecasting by the usage of scenarios may help decision-makers to expand their prospective thinking more broadly about potential futures, allowing them to navigate more assertively through uncertainties and to make more informed decisions and resilient plans to a diversified possible future. Our simulated paths presented on STL, ANN and combined methods are particularly helpful on this regard, given that we develop at least three per country (and method) possible routes for unemployment on BRICS up to 2032 year.

We also made use of both point and interval forecasts. The former one predicts a specific value – a point –, in our case of unemployment rates, to happen in the future whereas the latter may complement point forecasts due the offering of a range of potential values rather than one in specific (Li et al., 2019). Naturally, if we are expanding possible outcomes of unemployment, this would correlate with an increased variability and uncertainty about the forecasts that does not would describe likelihood of possibilities, which considering managerial interests may be desirable (Li et al., 2019; Wu et al., 2021).

From all the outlined we move for the scenario forecast writing. We share Schaars (1987) perception that three scenarios are consensual on the related literature. Two scenarios would be reductive to a “good-and-bad” outcome whereas more than three could lead to an elevation on uncertainties and variability that happens as well on point and interval forecasts, and we intend to mitigate through our quasi-qualitative approach of scenario forecasting.

We design our scenarios following a summarized version of the seven steps to scenario building according to Wright & Spers (2006) and Spers, Wright & Amedomar (2013). Brief descriptions about how we adapt, considering our research specificities, to follow the general framework presented by these authors’ work is presented on the following:

- **Step 1**: Starting point on the building process is to have a clear definition about scope and goals intended on the scenario. We believe to have this covered due the restrict sample been discussed on our research, BRICS countries reality. Also, we are projecting these countries potential unemployment rates in a short-medium horizon (10 years).

- **Steps 2 to 5**: We recommend to those wanting a deepening on these steps the reading of Spers, Wright & Amedomar (2013) study but in a succinct definition these four steps are condensing the operational part preceding scenario's writing.

On steps two up to the five are features like the identification of a target variable, unemployment in our case, driving forces and themes that could influence it (see chapter 2 of this thesis about emerging topics on unemployment-related literature and chapter 3 that discuss unemployment determinants) and the projection of future state of the target variable, something we have been performing throughout this chapter forecasting proceedings.

- **Step 6**: Is respective to what authors named the assembling of a morphological matrix of the scenarios (Spers, Wright & Amedomar, 2013). At this point are present the combining of all identified variables, relationships, projections, and trends that composes the overarching scenario that will be later depicted on three possible futures for unemployment within BRICS countries.

Our scenarios are developed considering the emergent themes on unemployment-related literature that we unveil by bibliometric analysis applied on chapter 2 of this thesis. Recently recurrent topics discussed that we extracted includes the unemployment insurance programs, youth unemployment rates, monetary and labour policies, COVID-19, entrepreneurship, and self-employment (see chapter 2 and the bibliometric analysis we early applied).

Determinants that compose and explain unemployment rates were assessed and measured on chapter 3 by Vector Error Correction Modelling (VECM) are as well considered. Gross domestic product, self-employment rates, and labour productivity, has a negative association, whereas inflation, youth unemployment, and savings-rates, have a positive association with unemployment rates. VECM that we applied indicates that the leading influential variable on unemployment is inflation, followed by youth unemployment and GDP.

Forecasting procedures: ANN, ARIMA, ETS, STL, and the combining of these four methods, each of them that suggests potential unemployment rates in the future are as well orienting the build-up of our scenarios. As mentioned earlier, when BRICS is observed as a collective of nations, we emphasize the combined method and aggregate unemployment levels that we present on table 95. Complementary to that, when describing country's level analysis, we respect the method in which that country performed better.

- **Step 7**: At last, but not least important, the final proceeding on scenario building according to Wright & Spers (2006) and Spers, Wright & Amedomar (2013) is the textual development of the scenarios, detailing each of their own variables, relationships, causes and effects that lead to that sketch of projected future.

Considering our study purposes this is the point where subjectivity of scenarios method comes into the mix. Nonetheless, all conjecture is results-based and respecting premises that guides our research. We undertook some effort to mitigate biases and subjectivism that could influence the scenario forecasting that we are presenting to close our results and discussions topic. Following table 96 presents the morphological matrix of our scenario forecasting.

Three scenarios from everything we developed and presented up to this moment are to be presented. Two of these scenarios are representing best- and worst-case that unemployment rates could evolve to 10-years into the future. There is a third setting embodying, maybe, the most likely outcome that may occur. On this alternative scenario we are considering historical data, forces, labour market interactions and economic events would somewhat replicate themselves from past into the future.

Table 96 (continues next page)
Morphological matrix of the scenarios.
(Elaborated by the author).

	Scenario 1: "Inflationary headwinds"	**Scenario 2: "Youth unemployment resilience"**	**Scenario 3: "Tech-led economic labour diversification"**
Variables on the scenario	Envisions a period of significant economic transitions, inflation-led, and political restructuring leading to increased unemployment rates.	Despite facing certain challenges and uncertainties, BRICS countries maintain a relatively stable employment environment. Many factors contribute for this stabilization and resilience to potential disruptions, especially due the youth labour force insertion.	Represents a positive outlook where economic growth, tech-led advancements, global and within BRICS countries impulse the employment situation for these nations.
Aggregate unemployment rates	Around 70% - 75%	Around 55% - 56%	Under 50%
Unemployment and labour market characteristics	Heightened uncertainty and volatility on job-opportunities. As inflation erodes the purchasing power reduces, producing costs increases leading to a slow down on hiring and workforce reductions.	Public policies are geared towards providing young workers with better-oriented education and early school leavers for training and entrepreneurial opportunities. Employers may benefit from a pool of skilled and innovative young talents what could fostering a competitivity on labour market.	Presumes a shift away from traditional industries, potentially reducing unemployment rates due the creation of new, high-tech sectors creates demand for reskilled and upskilled workers. Emerging sectors as artificial intelligence, renewable energy and digital services leads the way.
Maximum unemployment per country within BRICS	Brazil: 27% Russia: 9% India: 16% China: 3% South Africa: 45%	Brazil: 25% Russia: 10% India: 16% China: 4% South Africa: 45%	Brazil: 24% Russia: 13% India: 14% China: 6% South Africa: 43%
Relationship between employers and employees	Employers facing increased productive costs may implement cost-cutting, reduction on benefits, wage freezes and layoffs. Employees seeking to	Companies, aiming to retain and attract skilled and young talented workers may invest on mentorship programs, training opportunities a flexible work-	Relationship between employers and employees in a tech-led economic scenario is characterized by adaptiveness and collaborative efforts. Employers valuing

	Scenario 1: “Inflationary headwinds”	Scenario 2: “Youth unemployment resilience”	Scenario 3: “Tech-led economic labour diversification”
	cope with higher con living costs due inflation may become to be contentious within their job relations.	arrangements. Creating a positive work environment based-on sharing commitment and a mutually beneficial relationship.	expertise and capabilities tends to shift toward a partnership-based relationship with their employees.
Job channels	Traditional full-time options for work may decline due the inflationary pressures on cutting costs. Temporary or part-time employment may become the normal. “Gig-economies” jobs and freelancers may spike to cope the crescendo on unemployment rates.	Internships, apprenticeships, and mentorships programs may lead the way to fostering job opportunities to young workers. Gig economy could continue thriving offering flexible opportunities. Job channels may be a blend of traditional, innovative, and youth-focused initiatives.	Technology-driven fields such as artificial intelligence and digital services could face a growth. Skills-based jobs focusing on specific competencies may emerge as primary routes for employment. Online platforms and talent marketplaces might gain prominence.
Health, environmental, and economic labour laws	Economic pressures may tempt governments to relax certain labour and environmental regulations to ease the inflationary burden on business. On the other hand, this could lead to potential health and safety hazards for workers.	To attract and retain young talent, laws supporting work-life balance, mental health and equal opportunities might be strengthened.	New regulations may emerge to address ethical considerations of the emerging technologies, ensuring the responsible use of artificial intelligence and other innovations. Environmental laws might emphasize sustainable practices in technology industries whereas gig economy may be law regulated.

Table 96 presents our morphological matrix and the three scenarios we propose based on our forecasting quantitative techniques developed in this and on previous chapters of this thesis. We are referring our matrix based-on Spers, Wright & Amedomar (2013) study but in a more succinct version that is developed by our own research results. Three broader scenarios and 7 driving themes common to them will orient the description of scenarios that will come in the next topic.

Some paths to BRICS countries cope with unemployment phenomenon may be already inferred by the presented-on table 96. For example, youth-oriented efforts for job allocation, particularly due the large population in countries like China and India that have a high populated territory. Tech-led initiatives also appears as a potential driving force to create and evolve job opportunities if workers are prepared to be better skilled for future demands.

A bad scenario is also presented in table 96 by the entitled “Inflationary headwinds”, showing a potential unemployment growing driven by economic negative effects, as elevated inflation, that may influence private and public efforts on measures that led to reducing employment due increased productive costs on goods and services. The scenarios we propose are well in line with the potential aggregate unemployment that we present on figures 75 and 76, suggesting three paths that we intend to further explore from hereafter.

4.4.6.1. Scenario forecasting – Description of scenarios.

The scenario writing phase are respecting the seventh of seven steps that we described earlier according to Wright & Spers (2006) and Spers, Wright & Amedomar (2013). The topics before this one we covered during the entire development of this thesis up to the morphological matrix presented on table 96. We present three scenarios for the unemployment in BRICS countries into a 10-year in the future timespan; each of these scenarios are described on the following paragraphs.

Scenario 1, named "**Inflationary headwinds**", unfolds a narrative of ongoing substantial economic transitions and political restructuring within BRICS nations. Influenced by the relatively recent, but still damaging, effects from COVID-19 pandemic and the non-full recovery by countries an era of elevated unemployment rates across Brazil, Russia, India, China, and South Africa rises. Spanning the next decade that goes from 2023 to 2032, we envision a period marked by political instabilities leading to inflationary pressures and consequential negative implications on labour markets.

Envisioned period is characterized by a pressing economic recalibration that is mainly oriented by inflationary forces that permeate each facet of the society business and relationships landscape. As the presence of growing inflation numbers looms large, purchasing power across all five countries decreases whereas production costs escalate. In this scenario each country reacts accordingly their historical realities but with their own needs and urgency a discernible reduction in hiring activities implicating on accentuate labour force reductions. This projected disequilibrium between supply and demand for labour arrays a lingering period of uncertainty on jobs opportunities.

Reflecting the disparate impact of economic pressures, projected unemployment rates unveil a nuanced narrative across the BRICS countries. Brazil, grappling with enduring economic complexities and a historical inflationary problem, would represent 27% of total unemployment in the BRICS. Brazilian unemployment rates would peak historically, surpassing the pandemic years (around 13% on 2020 and 2021) and closing to 20% or even more Brazilians outside the labour force. Brazil's public policies regarding to labour would face an unanticipated shock with the poverty in the country also growing exponentially.

Russia does not escape the negative effects of the scenario, although it navigates in a comparatively moderate situation being responsible by 9% of aggregate unemployment within BRICS. Effects from military conflicts with Ukraine as well the potential to new conflicts in the region emerge do not ease the consequences on Russian labour market. Unemployment in Russia considering the 10-years projection could repeat historical indexes that happened in the mid 1990s years, when the dissolution of USRR where still looming. Unemployment rates could be going around 13% to 15% of people exonerated from their jobs or not being reallocate due to economic cuts.

India contends with a substantial unemployment rate among BRICS countries, representing 16% of total unemployment in the group. Indian historical and enduring uncontrolled population growth would lead to an equally large non-allocation of demand to fulfil the large supply of Indians desiring for a job opportunity. Peak of unemployment in India happened in 2020, above 10%, in this scenario this rate would be surpassed and one of the main drivers for it would be a growing migration of better-skilled workers that would prefer to seek for employment in other countries due the increasing life-costing expenses due the inflation being large.

China's, even in this worst-case scenario, would remain the most stable country among their BRICS colleagues, representing 3% of aggregate unemployment in the group. Chinese reality would be of a non-alarming unemployment rates growth due the economic power and relevance the country has in worldwide context. Unemployment could remain steady with historical data or even decrease to 3% that happened in early 1990s, considering that China has a better environment to navigate and even profit on inflationary pressures that could lead to damage in other nations.

South African longstanding unemployment problem in this scenario would only be increased. Entangled by the forced economic restructuring the country would be responsible by 45% of total unemployment in the BRICS group. Historical peak of people without a job, 29.8%, that happened in South Africa by 2022 would be worsened by the nation in this scenario's context most certainly passing the 30%. One third of Africans would not find any opportunity to be inserted into the country's labour force and public policies would reflect on it poorly not establishing any insurance programs for a such large number of people looking for employment.

These country-specific variations highlight the heterogeneous nature of challenges posed by inflationary headwinds. On the employer-employee dynamics the situation would not be different. Faced with escalating production costs, employers resort to several cost-cutting measures, ranging from reductions in benefits to wage freezes and, inevitably, layoffs. Traditional full-time employment avenues witness a paradigmatic shift in response to the inflationary pressures gripping businesses.

Part-time or temporary employment becomes prevalent as businesses seek adaptive measures to weather the economic storm. Simultaneously, the gig economy experiences expand largely, with freelancers and gig workers assuming pivotal roles in mitigating the escalating unemployment rates especially in the most-affected countries. Public policies and political measures are confronted with the dilemma to ease the burden on businesses by relaxing certain labour and environmental regulations whereas they remain able to extract profit to sustain insurance policies for the jobless.

Stakeholders, public or private, are urged by the necessity of striking a balance between economic imperatives and the protection of workers, becoming subject to an intricate dance that demanding nuanced policy decisions to navigate the complexities of the inflationary epoch. In summary, scenario 1 unveils a narrative of economic turbulence that reverberates through the very fabric of employment structures, prompting adaptive responses from businesses, employees, and policymakers alike. BRICS countries would need to be resilient and adaptable as nations tries to cope with formidable challenges imposed by inflation-led uncertainties.

"**Youth unemployment resilience**", the scenario 2, envisions a period where BRICS countries navigate labour market challenges having as focus the stabilizing of employment landscape, particularly for the youth, to consequently mitigate their overall unemployment levels. This scenario, covers years from 2023 to 2032, emphasizing factors that may contribute to stability and resilience in the face of possible disruptions, with a spotlight on the potentialities of youth labour force.

In this scenario, public policies take the main role on directing efforts towards providing young workers with better-oriented education and training opportunities. Early school leavers are targeted for specialized training and entrepreneurial ventures. This emphasis on skill development on the early days of working life creates a competitive advantage in the labour market. Companies may benefit from these policies having a pool of skilled and inventive young talents which fosters a dynamic and competitive labour environment that could in medium and long term mitigate unemployment.

Reflecting the concerted efforts to bolster youth employment, projected maximum unemployment rates within BRICS showcase a relatively stable landscape. Brazil, with a focus on skill development, youth-oriented policies and entrepreneurial foments would be responsible by 25% of the unemployment on the group. Brazil, however, faces a population-ageing problem what is a twofold phenomenon. May implicate on more opportunities to young people to be inserted on labour market whereas would demand insurances for those exiting from it, demanding a pivotal public policy interference to strike an economic balance.

Russian public policies as well are oriented on youth availability on regard to create for these more opportunities and to retain their potential labour force on the country. Within the group the estimative is that Russia would represent a maximum of 10% of unemployment among the five nations on the BRICS. Russia alike the Brazilian case have an ageing number of people the caveat that could present a higher unemployment in this scenario when comparing with scenario 1 is that youth population may prefer to leave the country due their restrictive politics and ongoing problems with military conflicts.

It is a known fact that India must deal with a large population for better or worse. Most populated country on the world have at their disposal a large youth population as well, what could present as opportunity to, if better prepared and oriented in early stages, those entering the labour market may occupy or even create job opportunities for themselves. Within the BRICS we project that India would maintain a steady 16% of the unemployment on the group but these number could go lower if public policies are well-designed to foment entrepreneurship and self-employment initiatives, that could tackle unemployment problem at the source to portray a better future.

China remains with the leveraging of its economic power being the most resilient country within BRICS countries. On the group perspective the Chinese labour market would be responsible by 4% of BRICS' unemployment, a rate that is similar as historical mean values that happened analysing country's available information going from 1991 to 2022. China not being directly involve in military conflicts, like Russia, have as well a restrictive governmental administration that could led to people intending to escape from the country although this could be offset by a larger population and potentially more people interested to migrate for the country seeking for better labour opportunities.

South Africa would remain stable with higher unemployment levels within the BRICS being the country that rates are representing 45% of jobless individuals on the group. However, on this scenario we do not anticipate an above the historical unemployment rates surpassing 30%. In fact, this may be the scenario where South African public policies and private efforts may envision a path to cope better with their enduring unemployment disfunctions. Aiming efforts on youth initiatives as soon as they enter in the labour market could restrain their unemployment at least on some stability.

Private companies, influenced by public policies making, recognizes the potential benefits and importance of retaining and attracting skilled young talent. Higher investments in mentorship programs, training opportunities, and flexible work arrangements passes to be key elements on managerial efforts. Employer-employee relationship changes for the imminent fear of firing into a positive, mutually beneficial partnership. Centred around shared commitment and the cultivation of a conducive work environment, job satisfaction, increasing productivity and innovation are enhanced.

Internships, apprenticeships, and mentorship programs are the main pathways for young workers to immerse on labour market. Gig economy continues to thrive, offering flexible opportunities tailored to the preferences of the younger workforce and their own entrepreneurial initiatives. Job channels become a blend of traditional structures, innovative initiatives, and youth-focused programs. This diversification of channels ensures that the evolving needs of the youth are met while maintaining a dynamic and adaptable employment landscape for companies' interests.

Recognizing the importance of work-life balance and mental health in attracting and retaining young talent, labour laws with this focus are strengthened. Emphasis is placed on fostering equal opportunities, creating an inclusive environment that aligns with the values and expectations of the younger labour force. In summary, scenario 2 unfolds a narrative of resilience, stability, and adaptability, driven by strategic public policies and a concerted effort to empower youth people on the labour market. This scenario emphasizes the pivotal role of a younger workforce in shaping a competitive and innovative labour market across the BRICS nations.

Named as "**Tech-led economic labour diversification**" the scenario 3 portrays a most positive outlook for BRICS nations. Foreseeing a future where economic growth, technological advancements, and global initiatives within the BRICS countries drive a transformative shift in employment dynamics, this scenario, spanning the years from 2023 to 2032, represents a departure from the traditionalist industries and labour relationships accompanying a tech-driven era that significantly impacts the labour market.

This scenario anticipates that this departure from traditional labour structures would lead to an accentuate overall reduction in unemployment rates. The emerging of a new high-tech sectors for labour opportunities, particularly in artificial intelligence (AI), renewable energy, and digital services, becomes the driving force behind this potential impulse for the decreasing of unemployment levels.

Considering that this scenario becoming reality, to fully accompany this changing of landscape a demand for reskilled and upskilled workers is created demanding as well an effort on the supply side of the labour force. Under the umbrella of tech-led economic diversification, the BRICS nations would exhibit a positive trend with the aggregate unemployment rates in the group anticipated to be under 50%, the lowest of the three scenarios we have been proposing.

Brazilian industries are focused on a shift towards technology-based initiatives. Agriculture sector in the country leads the way on advanced methods to produce more to export and intern consumption effecting the economic landscape and, consequently, creating more job opportunities in different sectors not only on the production of these goods, but in service offered as well. Within the BRICS countries aggregate unemployment Brazil would be responsible by 24% of unemployed people in the group. Brazil would be dependent on partnerships in- and outside the BRICS to foster tech-led initiatives beyond agribusiness sector, probably the most developed sector in the country.

Russia, within the BRICS colleagues, navigates in a comparatively stable scenario, only China performing better. Military conflicts that in the other two scenarios have been a negative influence spreading to the labour market, here would turn to be a positive influence. Tech-led innovations early directed for military intentions would pivot to other areas playing a key-role on technological advancements that may foster better situation into labour market. Russian unemployment would not repeat historical peaks that happened in the 1990s and potentially lowering the rates occurred in early 2000s.

India, with a focus on digital services and skill-oriented jobs would have in scenario 3 the better positioning among BRICS countries, being responsible by 14% of total unemployment in the group. Indian efforts would be directly oriented to a better and oriented development on their labour force, through training programs, skill developments on new job opportunities that emerged and so on. India would profit on partnerships of more developed economies, as Russia and China for example, that in the other hand would profit from the larger Indian labour force.

China would have on this scenario the worst contribution on the aggregate unemployment within BRICS countries, peaking at 6%. Leveraging on its technological prowess, Chinese labour market still maintains an extremely favourable labour market context on regard to unemployment, not going above the single digit of Chinese people being outside the labour force. This scenario would be the bad one for China’s specific because their technological advancements could be so efficient that could lead to some jobs starting to be performed by machine and not by individuals.

South African, opposing to the Chinese case, would have the better scenario here in this scenario, representing 43% of unemployment within BRICS countries as in scenarios 1 and 2 this rate was by 45%. Tech-driven initiatives would mitigate the South Africa unemployment disfunctions, considering that this new global-tech context would enable people from South Africa to benefit from job opportunities that could emerge from other parts of the world without the necessity to leave their own country. As we describe for Brazil and India the tech innovations on South Africa would also be dependent on partnerships for the democratization of employment opportunities.

In this tech-led economic scenario, the relationship between employers and employees undergoes a paradigm shift characterized by adaptiveness and collaborative efforts that would be mandatory for both sides. Employers, valuing expertise and capabilities, transition towards a partnership-based relationship with their employees. This shift into a collaborative association fosters an environment where continuous learning and skill development are not only encouraged but become integral to sustaining competitiveness in a tech-driven landscape.

Technology-driven fields, including artificial intelligence and digital services, witness significant growth, becoming primary avenues for employment as well the place where skilled individuals could look for better job opportunities. Skills-based vacancies would emerge on the labour market pool, offering prominent routes for jobseekers. Online platforms and talent marketplaces grow on relevance, facilitating a dynamic and flexible market that aligns with the demands of the emerging tech industries.

This rising of technology brings to discussion a brand-new landscape of regulatory considerations to address ethical implications of emerging technologies. A special focus on ensuring responsible use of artificial intelligence would be central on all labour relationships, where the lines regulating AI use must be clear and concise. Environmental laws are expected to emphasize sustainable practices, aligning with global efforts for a green and eco-friendly tech sector.

Backing to the employers-employees association and still discussing about this law regulations, in this scenario the gig economy would potentially faces a decrease as we may presume that better prepared and skilled workers would demand for more solid and protective labour contracts. Within this context, part-time jobs and gig economy if closely tied to the tech landscape, must adjust to cope with an increased regulatory scrutiny to ensure fair and just practices on labour relations.

Scenario 3 envisions technological diversification on creating opportunities and becoming a driving force for positive change within BRICS nations and globally. This scenario capitalizes on scientific innovations fostering a dynamic, adaptable, and collaborative labour market that may position the BRICS countries at the forefront of global technological landscape over the next decade. Chinese economic development could be on the driver's seat and with them Brazil, Russia, India, and South Africa may have themselves a better scenario.

4.5. Final considerations and conclusions.

On finishing this chapter, we return to the start where objectives and research question were delimited to all the writing and proceedings we developed up to this point. We have as targeted objective the proposition of possible future scenarios for the unemployment rates in the BRICS countries. We indeed comply with this, designing three scenarios for Brazil, Russia, India, China, and South Africa in a 10-year into the future scenario forecasting. On doing that, we as well answered the questioning about what tendencies could happen within these nations' labour market regarding unemployment.

The scenario forecasting that we use and present it is per se an own method, a purely qualitative type, but before to proceed for the usage of it we run four quantitative techniques to substantiate our scenarios. We alert and reinforce again; we do not believe nor presume that one method is better than any other. This is the reason we apply different methods (ARIMA, ANN, ETS, STL) combine them and after these applications we proceed to scenarios' writing. Premise we believed and used is to have the most inputs by this multimethod approach, maintaining a manageable level of information, to consequently write informed and assertive scenarios.

Before to deepen on the methods and results we obtained applying them we make an apart to justify our choice to have as scope Brazil, Russia, India, China, and South Africa, the BRICS countries. First, we scanned the unemployment-related and forecast/foresight literature and studies that deal with the particularities of these countries are very incipient even at individual country-level. These five nations analysed in a group perspective and on foresight purposes, as far we are developing this thesis, we do not find any similar as the one presented here.

More than this academic opportunity, we envision BRICS as a potential group of countries that could play a key role on world's economy. When these nations decided to collaborate at some level, all of them had in common the "emergent economies", years passed, and China is nowadays a global leading economy and in some metrics is the most powerful country of the world. Chinese example could lead other countries within the group to follow the path if they are, as China's, able to manage well their own resources including each of them labour markets.

We believe that the outputs we presented, and discus may contribute on labour market assessments, especially on the unemployment side, offering reliable about how unemployment rates are in the present and how it could evolve in a short and medium term, considering our 10-year forecast horizon. We do not intend, neither we could, unveil what will exactly happen in the future even describing three potential scenarios about it. More likely outcome is that reality would combine things we present on the proposed scenarios even that one of them ends up being more assertive.

After some focusing on each technique of our multimethod approach, we move for scenarios' consideration. We apply all the four quantitative techniques using R-Studio software and mainly the package `fable`, presented on Hyndman & Athanasopoulos (2021) textbook. Other R-functions and features complement `fable` for other demands especially regarding the visual graphics that we show during the results presentation.

ARIMA, ANN, ETS, STL were applied for the five BRICS countries individually and, as much as possible, we assess nations data simultaneously, as a group.

Diagnostic checks were performed at the end of every application to measure their potential error on forecasted values. Concluded all methods usage in isolation we proceed to a fifth-round of statistical proceeding that is a combining of the four into one combination that is the main referential to our scenario forecasting writing process. As we applied these methods it was possible to perceive that even using the same countries' data on all technique, some performed better using one than other.

Seasonal and Trend Decomposing using Loess (STL) method is the one presenting better performance for China's data. Chinese forecasted values anticipates a stable unemployment in the future, something around 4.5% of Chinese people would be outside the country's labour force. Results for the 10-year forecast horizon are in line with historical data and indeed suggests stability, median values of unemployment in the country where by 4.48% from 1991 to 2022 and there is no peak above 5%. For the future, Chinese labour policies being used could be extended as it appears they are capable of maintain unemployment in a solid proportion even facing unexpected economic shocks.

Exponential Smoothing Technique (ETS) is the method that appears to be offering more assertive results to our envisions of the future purposes. ETS is the method minimising errors, tend to present better forecasts, for three of the five BRICS countries, Brazil, Russia, and South Africa. That is not a surprising result considering the potential ETS presents to perform well in scenarios where timeseries data exhibits a certain degree of regularity, seasonality, or trend as is the case for unemployment rates. ETS may be particularly effective for short-term forecasting and in situations where historical patterns are a presumed indicative of future behaviour about a variable of interest.

Brazilian forecasted values by ETS projects unemployment rates from 2023 up to 2032 having around 11% with some margin to be better or worse than this baseline. 11% for Brazil would implicate on a few more Brazilians citizens not allocated with a job on the labour market, surpassing historical data available where mean unemployment is on 9.44%. Nonetheless, this projected increase on unemployment does not anticipates a historical peak that for the country happened on the COVID-19 pandemic year reaching 13.93% in 2020.

On Russia's case we have a better future possible, comparing with the one we just described for Brazil. Unemployment projections for Russia, also using the ETS method, anticipates a mean value around of 5.56% of Russians being potentially jobless in our projected 10-year into the future horizon. A 5.56% rate would implicate a value under the mean Russian unemployment from 1991 to 2021 (one-year shortage of information relating with other BRICS countries) that is 7.29%. This potential better future however would not reach the minimum unemployment in the country that was under 5% in 2019, before the COVID-19 consequences.

As we have found China as the most stable country among the BRICS on regard to labour market, it was noticeable that the opposite is the case for South Africa, which by any method or analysis appears as the one most damage by elevated unemployment rates. South African historical data ranging from 1991 to 2022 presents a mean percentage of unemployed people by 21.91%. The minimum rate on the country 19.51% in 2008 it is not even close at the maximum reached by any of Brazil, Russia, India, or China. South African projected values by ETS anticipates a mean unemployment by 23.98% surpassing their mean historical numbers. An alarming scenario could unveil in the future for South Africa, labour policies are indeed urgent to be applied or unemployment situation could be worsened in a not-so-distant future.

Autoregressive Integrated Moving Average (ARIMA) technique is the method presenting more assertive results for Indian data. ARIMA method not being the one that is offering general better results came as a surprising result considering that this is probably the recurrently used tool by studies that have some similarities with ours. On specific to India's forecasts, ARIMA projects that unemployment in the country could reach some level around 7% with margin to plus or minus. If our forecast comes to be reality would not be apart from Indian historical data where from 1991 to 2022 mean unemployment is on 7.8%. Forecasts do not anticipate above nor under values from historical unemployment in India, where the maximum was by 10.19% in 2020 and the minimum 6.51% in 2019, suggesting some high volatility in Indian data.

Artificial Neural Network (ANN) method does not appear as the best suited technique for any of the five BRICS countries as we may notice. In fact, comparatively, ANN is the technique offering more deviancy on results for better and worse. We believe that machine learning with such robustness as ANN does not fit well with a relatively small sample and data we are dealing in our study. Also, considering that the method is essentially iterative results from ANN could be deceitful as each time the technique is applied, different results may emerge.

Nonetheless, we do not disregard the contributions that ANN applications brought to our research. Using as advantage the iterative and bootstrapping proceeding that is inherent on the method, we project within the ANN some scenarios unemployment could develop in each of Brazil, Russia, India, China, and South Africa realities. Figures presented throughout the ANN usage offered a glimpse on how these rates could repeat, deviate or be abnormal from countries historical data, something that when dealing with real-life information, as unemployment, may not surprisingly come to fruition.

In the end, every method contributed to our general landscape for unemployment in BRICS countries. Inputs from Artificial Neural Network, Autoregressive Integrated Moving Average, Exponential Smoothing Technique, and Seasonal and Trend Decomposing using Loess are valuable to better inform us about current state and how the unemployment levels for each country may evolve in the delimited 10-years horizon we decided to project.

In line with this assumption of all methods being valuable suppliers of information for a such complex phenomenon as unemployment, we performed a combination proceeding, again, following the presented-on Hyndman & Athanasopoulos (2021) textbook and other research (e.g., Bates & Granger, 1969; Clemen, 1989; Wang et al., 2023) that suggests combining is better. We aggregate unemployment rates, summating all countries individual data, and use this new variable for forecasts. Visual projections and results of this application on topic 4.4.5. and were the driving input to our final piece, the scenario forecasting.

Basing our scenarios on all the above-described quantitative proceedings we indeed believe to have enough inputs to present plausible and potentially likely tendencies for the unemployment rates in the BRICS countries. Research question and chapter's objective, to our intentions are met. Our scenario forecasting does not disregard the speculative nature of this method whereas simultaneously we speculate with solid foundations and rigour to project what may come to reality 10-years from now.

Scenario writing process that we follow is complying with Wright & Spers (2006) and Spers, Wright & Amedomar (2013), which by their own follow seminal premises for the method according to Michael Porter and Michel Godet. We designed three scenarios envisioning potential paths for unemployment rates in the future of BRICS countries, namely: "Inflationary headwinds", "Youth unemployment resilience" and "Tech-led economic labour diversification". On presenting these three we are foreseeing the bad, the good and, maybe, the likely outcomes that could come for our nations scope in a near future.

"Inflationary headwinds" presents a scenario where BRICS nations face a challenging decade marked by economic turmoil, political instability, and high inflationary pressures. Unemployment rates in this context rises and impact every country especially those that already have historical inflation problems and enduring unemployment as is the case for Brazil and South Africa particularly. Traditional job structures shifts whereas new labour relationships as gig economy gains space with employers and employees having to cope with elevated inflation impacting their lives.

In scenario 2, "Youth unemployment resilience," BRICS countries resort to a slightly better perspective. Brazil, Russia, India, China, and South Africa direct their focusing on stabilizing the employment landscape dedicating public and private efforts onto youth population. Public policies pass to be designed to prioritize better education, training, capacitation, and entrepreneurial opportunities, resulting in a relatively stable unemployment scenario. The employer-employee relationship transforms into a positive partnership, fostering innovation and adaptability.

Scenario 3, "Tech-led economic labour diversification," is the one that envisions a positive future for BRICS nations. Driven by some level economic growth especially due tech advancements, unemployment rates may start to follow a decreasing behaviour. Emergent high-tech sectors pressure the relationship between employers and employees to become collaborative, emphasizing continuous skill development for both parts. Technology-driven fields, including artificial intelligence and digital services, become primary channels for employment, with a focus as regulations to deal with this environment tends to be more restrictive.

From a general point of view, we believe that throughout our scenarios proposed we covered well what are the tendencies that may come to fruition on BRICS countries. All scenarios, even scenario 1 ("Inflationary headwinds") portrays situations that does not are too apart what happened on Brazil, Russia, India, China, and South Africa historical data. This is in line with the values we forecast using ARIMA, ETS, STL, and even the broader iterations from ANN method.

Beyond the scenarios proposal we point to some routes that could, or should, be pursued to indeed bring these paths to reality or in the case of the worst-case scenario to better cope with the problems it may accompany it. Overall, we believe that this chapter, our scenarios, and all the proceedings that based their writing process are a useful tool for better planning to different interested players within BRICS nations. For governments, professional associations, and even for individuals what we present and developed could be a reliable source of guidance from what might come in their futures.

We have found an underpinning demand for all scenarios: Cooperation. BRICS countries might have an envisioned better future, but the grouping approach must be real and not only an acronym that connects these five countries. If, as we project, tech-led innovations would spread developments that reaches the labour market relationships, those nations that already have a better developed technology, Russia and most particularly China, should contribute with the lesser evolved. If inflation peaks and economic pressures are exponential the richer should not solely intend to get richer but offer some aid for the most pressured. Empowerment should be across the BRICS board; we believe that a collective effort is a must to navigate the unemployment uncertainties.

Although considering we are offering solid results and foreseeing, some limitations must be declared. Potentially, our multimethod approach may be excessively extensive and at some topics confusing. This is a twofold problem indeed that although results that we retrieve, and present may be excessive they are foundation to the inputs we use to write our proposed scenarios. We delimit a reasonable scope of countries and information, focusing on BRICS enables the opportunity to have a glimpse of distinct economies and development levels, although we could be more assertive restricting even further the scope of country's information analysed.

We faced a particular problem using the Artificial Neural Network method, being this technique iterative results that we present, discuss, and used for writing our scenarios could be different in any other application even using the same data we have worked. It could be the case to remove ANN from our multimethod proceeding something, a path we choose to not follow due the still valuable inputs the method offered. We recommend for future studies derived from this to be alert by this caveat and proceed with caution with the ANN usage.

Limitations and biases considered we believe that this chapter offers a reliable presentation of possible future scenarios that might come to reality for BRICS nations. Having what we developed in hand decision-makers and unemployment-related stakeholders could be better informed to design assertive strategies to cope with perennial uncertainties that permeates labour market and unemployment rates, being able to, potentially, safeguard positive results for themselves. To ensure our results and proposed scenarios we encourage further studies that follows our lead and fill the gaps we left. Unemployment is a multifaceted and complex phenomenon that we do not solve or exhaust with our writing.

5. CONCLUSIONS.

This thesis represents an exhaustive, although not conclusive, exploration into the intricacies landscape of unemployment phenomenon. Comprising three interconnected chapters that assesses past on this theme, by bibliometric analyses on unemployment-related literature; present, observing some influential determinants that composes unemployment rates; and future, by forecasting techniques and scenario-based foresights; we believe it was possible to shed some light into theoretical, empirical, and methodological insights for a better understanding of unemployment complexities.

As we aimed on the initial proposition and research delimitations, each chapter adds a unique layer onto the comprehension of such a complex and multifaceted phenomenon as unemployment is. Besides the independent results and assessments by our chapters 2 to 4, together, they paint a widespread picture of the challenges and potential future that unemployment may depict especially for our final targeted nations, Brazil, Russia, India, China, and South Africa.

We do not resolve all unemployment and labour market overall dysfunctions. That was not the goal we pursued when developing this thesis. Nonetheless, it is our believe that we were able to present an extensive and reliable portrait about unemployment that may be further developed upon the basis we are proposing on this study. Resuming to the main research question that guided our study (How could unemployment rates evolve in a 10-year forecasting horizon for BRICS countries?), it is possible to confirm that a satisfactory answer has been accomplished when all three chapters that precedes this one are developed and presented.

Indeed, the collective efforts and analyses presented in this thesis are converging to address our central research question going to a broader scope on chapter 2, all unemployment-related literature, passing through the assessment of unemployment determinants on supra-national levels in chapter 3 up to the refinement directed to BRICS forecasting of unemployment rates and scenario forecast propositions on chapter 4. Each chapter provides nuanced perspectives on different facets of this thesis' inquiry, offering a holistic and in-depth understanding.

Going through all the steps to fulfil and respond the proposed research question we as well met the defined main purpose of this thesis, the proposition of a scenario-based foresight for unemployment rates in a cross-country analysis, having as main scope Brazil, Russia, India, China, and South Africa realities. Before these countries particular assessments, we have directed our research efforts following a tapered approach. Starting with a macroanalysis of the unemployment literature, followed by a still macro cross-country examination of potential determinants for unemployment rates from all around the world to culminate on restricted to BRICS forecasting quantitative and later qualitative proceedings.

On applying this integrative approach, we believe to have achieved another intention on this research. Chapters could, and should, be understandable by an isolate reading whereas as the same time they converge between themselves to accomplish the delimited main purpose specified on our introductory section. Integrating results, insights and conclusions derived from the three chapters we are offering a unified view of unemployment phenomenon, a matter of interest for BRICS nations but for any country that inevitably must cope with such a perennial macroeconomic factor.

Bibliometric analysis developed on chapter 2 was the starting point to us, in writing about it, and for the reader's first in-depth assessment of unemployment phenomenon. We reviewed all available unemployment-related literature on academic research in Scopus and Web of Science databases. By doing this reviewing process our intention was on identify emergent topics discussed on the field as well the ones that long endure on the theme, going back to the initial studies dated from 1971 for the most recent.

Bearing this purpose in sight we believe to have answered the chapter's research question (What are the main topics discussed in the unemployment-related academic literature?) while concomitantly we unveil some potential driving themes that would be useful for the checking of possible determinants composing unemployment rates that was later developed on chapter 3. Before that and adding to the most discussed topics on the research field, our bibliometric analysis enables the unveiling of most cited documents, referential sources for unemployment studies, leading authors on citations, most research-productive countries, and other insights on the broader scope of unemployment theme.

We will not overextend on the outputs obtained by the applied bibliometric analysis (see section 2.4 for that) but some results are worth of mention. There were plenty of tracks within unemployment-related literature emerging from our reviewing proceeding, and we still presume to not have exhausted all of them. Five driving themes by our understanding have been more prominent: Unemployment insurance programs, youth unemployment, monetary and labour policies, COVID-19, entrepreneurship, and self-employment.

These five mentioned themes were highlighted because they are reflecting enduring recurring themes (as youth unemployment) and recently emergent ones (as COVID-19). More than their recurring presence on the analysed studies apart from COVID-19, which effects and exact extension on numbers when we started to develop this research being still incipient, the other four factors may be more directly measurable, something that turns to be useful for put these themes into the potential determinants composing unemployment rates, intention pursued on the chapter that followed the bibliometric one.

Other relevant insight that emerged from the bibliometric analysis is that most of academic research regarding unemployment-related studies are being developed on some of the leading worldwide economies, particularly from United States, Germany, and United Kingdom. From this result we had our first indicative to move away of these nations' context. Comparing with lesser advanced economies it is reasonable to presume that different countries, dealing with higher economic problems, would suffer more with labour market dysfunctions as unemployment.

This perception resonated for us that an alternative scope from the ones most usually analysed countries, which led us for BRICS nations, would be of better interest. First, would be quasi-innovative on academic level, considering that the bibliometric analysis indicates that studies within these countries' reality were scarce. Second, and derived from the first insight, it is our believe that if unemployment-related research on these nations is underdeveloped, managerial and public policies may be missing inputs that could be extracted from our efforts on propose an examination that is more oriented to their specificities.

Within the bibliometric sample of 2334 documents the ones most cited reinforce the five themes highlighted before. Meyer (1990) for example discussed unemployment insurances policies; Blanchflower (2000) argues about self-employment as an alternative to cope with the absence of job-opportunities; and Glaeser, Scheinkman & Shleifer (1995) study assessed economic growth related with labour market characteristics. These three mentioned here are just an illustrative example on how chapter 2 offered to our research useful insights on unemployment and solid starting point to proceed for the sequential studies.

Indeed, from bibliometric analysis we could identify some prominent themes on unemployment-related literature, guiding the selection of variables definition for subsequent study. Third chapter advances with an econometric analysis using the Vector Error Correction Modelling (VECM) technique. The main premise of our usage of the method, the one that respond to the chapter's research question, is to assess the level of influence of selected variables (determinants in an equation) on overall unemployment rates composition.

Be conscious about some of the factors that exerts influence into unemployment as well the direction of this influence, if a positive or negative one, may be a basal knowledge to know in advance which factors must receive attention to a better coping with labour market dysfunctions and unemployment. VECM is the main methodological proceeding on chapter 3 development, but we also used a clustering technique for achieve a better assessment of a final sample of 154 countries' data.

We delve deeper than the composition of unemployment rates checking for worldwide differences of unemployment and how one variable may have a different form of influence in a distinct group (cluster) of nations. Six variables were defined and analysed as exploratory factors for unemployment rates: Gross domestic product (GDP), inflation (INF), self-employment rates (SER), youth unemployment (YU), labour productivity (LP), and savings-rates (SAV).

For these six potential determinants for unemployment rates, we used some of the emerged themes from bibliometric analysis (YU and SER), well-grounded factors on the literature as the cases of GDP by the Okun's (1962) law and inflation by the Phillips (1958) curve concept. Other variables were more discretionary by our assessment of the data. Labour productivity and savings-rates were selected from managerial and economic purposes to assess how efficiency on production and financial security may impact on unemployment levels.

On a general assessment, we have found that GDP, SER and LP have a negative association with unemployment rates, indicating that when these variables indexes are increasing, this implicates on a lowering in unemployment. When economic growth (proxied by GDP) is sustainable, self-employment and productivity are prolific, countries tend to experience less unemployment dysfunctionalities. INF, YU and SAV on opposition had a positive association with unemployment. Meaning that if inflation pressure is rising, number of youth people without a job is growing and people can save more money, unemployment levels will increase.

Results described above are referring to the complete cross-country sample that have 154 nations aggregated. As we mentioned to have applied a clustering proceeding to subgroup this larger sample, three groups were formed and named as Low, Medium, and High labour markets. This formation of clusters' procedure was adopted to have an insight on how, and which, countries would agglomerate without being forced but rather due to inherent similarities they had between themselves. The global and later clustered approach used on chapter 3 was the igniting point to later refine or scope for BRICS countries.

Assessing where BRICS countries were allocated on the three proposed clusters, on dealing with these five nations we are as well covering all groups that were created. Brazil and Russia were inserted on the Medium Labour Market (MLM); India and South Africa are included on the Low Labour Market (LLM) whereas China is the country that represents the High Labour Market (HLM) on our later refinement on the scope of countries.

We assume that further analyses would be more insightful considering the clusters' representation because as we could observe on these groups assessments, each cluster had their own level of influence by each of the six exploratory variables defined. In the LLM group of countries, GDP and inflation have a positive association with unemployment rates composition while self-employment has a negative one; other three variables were not statistically significant. MLM presented all variables having statistical significance, GDP, INF and SAV positively associated and YU, SER and LP negatively. HLM as well have six variables as significant, being GDP, YU and SAV as negatives whereas INF, SER and LB were positively associated with unemployment rates.

Leading influential variable differs by cluster, therefore by countries, and we strongly recommend the reading of the results topic (section 3.4) on chapter 3 for an in-depth analysis of the overall VECM application both on clusters' level and on the 154 complete sample of countries. In the end, VECM usage and outputs offered to our research some avenues to be pursued on the fourth chapter of this thesis.

First and foremost, we perceived that a refinement on the scope with implicate on more assertive results considering that forecasting and studies speculating the future demands a more delimited scope to be assessed. That is the main directive we follow on chapter 4, refining our analyses for BRICS countries once they are representing all clusters' we create whereas share similarities between themselves in a global context. From this premise and, to our understanding a solid portrait about unemployment rates composition, we move forward to check how these rates could present on the future and to, at last, tackle the main research question and objective proposed for this thesis.

The fourth chapter follows a multimethod approach aiming to present a final piece for our unemployment phenomenon assessment. At first, we check for all within our reach available information regarding the unemployment-related literature; later, we focused on understanding how unemployment rates are composed, in a global level and within clusters of nations; finally, we direct our lenses for analyse a group inside the groups before presented, the BRICS countries.

We restrain our scope to have a concise and feasible working data in hand to fulfil a purpose of research that does not turns excessive wider potentially limiting any results and analyses inferred. Defined main objective being the envisioning of possible future scenarios for the unemployment rates in the BRICS countries are fulfilled on chapter 4 as well the research question that orients the chapter, and the thesis in a wider perspective is answered as we unveil foreseeable tendencies of unemployment rates in Brazil, Russia, India, China, and South Africa considering a 10-year into the future forecasting horizon.

Our efforts to offer a reliable scenario forecasting as we believe is presented on chapter 4 followed a two-way path. First, we have extensively projected future unemployment rates for BRICS countries using a pool of quantitative proceedings, namely: Artificial Neural Networks – ANN, Autoregressive Integrated Moving Average – ARIMA, Exponential Smoothing Technique – ETS, and Seasonal and Trend Decomposing using Loess – STL. Having as referential all outputs obtained by these methods applied in isolation, we intend to mitigate biases when directing our research to a qualitative and inherently more subjective approach, the scenario-forecasting.

Some of the reasoning behind our decision to dedicate further analyses on BRICS countries in specific were already depicted here. Dealing with Brazil, Russia, India, China, and South Africa we are covering the three clusters formed on chapter 3 which implicates that distinct contexts regarding unemployment rates composition and labour market characteristics were contemplated. Also, despite their own distinctiveness, these five countries have some level of similarities that drove them to converge and create the BRICS bloc, something we further asses on chapter 4, section 2, topic 3 and recommend being checked for further clarification.

On the methodological proceeding adopted for this chapter fourth our guiding process as well was inferred from chapter 3. On the VECM application we perceived that despite using a same group of exploratory variables, their influential level on the composition of unemployment rates could differ depending on the country that is being observed. That's the reasoning on selecting the countries we have worked as well the justifying our effort to make a somewhat inverse approach from chapter 3, now maintaining the same group of countries and assess their unemployment through different methods and using one variable in isolation, the unemployment rates.

Based on the proposal and results from chapter 4, and by the same on chapters that preceded it, we believe to have offered an extensive and solid depiction of unemployment phenomenon complexities, both at a far-reaching level and in a more delimited context that considers the BRICS countries. It is indeed our intention to be contributory for the unemployment-related research field, not necessarily we aim for innovativeness or to solve all perennial labour market problems; rather we aimed to assess how unemployment was in past literature, the present of this variable and how it could evolve it 10-years into our future, although we may not predict anything with total accuracy.

Our quantitative proceedings are anchored on Hyndman & Athanasopoulos (2021) whereas the scenario writing follows steps presented and developed by Wright & Spers (2006) and Spers, Wright & Amedomar (2013). Following these and the ANN, ARIMA, ETS, and STL outputs, three scenarios foreseeing possible paths for unemployment rates in the future of BRICS countries were proposed: “Inflationary headwinds”, “Youth unemployment resilience” and “Tech-led economic labour diversification”. Scenarios that we create are also in line with results and tendencies unveiled on the second and third chapter of this thesis, offering a consolidate result from everything we developed up to this point.

Again, without overextending about the results we already presented on chapter 4 we have acknowledged a common ground on all scenarios, and consequently, for the future of labour market and for a better confrontation of their dysfunctionalities specially about unemployment levels: Cooperative efforts between BRICS countries is a must for a potentially less harmful future. From a group perspective, collective thinking and effective managerial policies must, in fact, be implemented to involve and benefit all parts of the group.

In general, we noticed in all the results, despite observing the unemployment rates forecasted through ANN, ARIMA, ETS, or STL, that in the defined 10-year into the future forecasting horizon, the historical data, even considering unexpected economic and social shocks, and the still lasting effects of COVID-19 pandemic, they tend to be repeated with small deviations. Taking this perspective into consideration, little or no progress for the BRICS would be sustainable if China’s already consolidate economic leverage in comparison with Brazil, India, South Africa, and even Russia, benefited Chinese people solely.

Going through some results and inferences extracted from each chapter, we can notice the interconnectedness between all studies, while simultaneously, each of them also brings its own contributions for discussions about unemployment theme. Variables unveiled by bibliometric analysis, such as inflation, youth unemployment and self-employment, were later explored as determinants on the composition of unemployment rates, using VECM methodology, and used as driving themes for our scenario-based forecasting proposition presented in the fourth chapter.

Nonetheless, this established connectivity between chapters and themes they cover in isolation up to converge on resolve our main research question and purpose does not eliminate limitations and avenues for future studies that can develop from this one. Our research follows a tapered approach but the scope refinement of the scenarios for BRICS countries naturally imposes a limit on the generalization of results to other regions. Data sources as well is limited, we are focusing on Scopus and Web of Science for the bibliometric part and World Bank and ILO for the unemployment rates assessment, other sources of information may suggest divergent results or at least some comparison would be useful.

Furthermore, the use of forecasting methods and the projection of future scenarios are inevitably subject to uncertainties, especially considering dynamic topics of economic and global interest such as unemployment levels. Unforeseen factors, such as geopolitical changes or unexpected technological advances, can impact the accuracy of our forecasts and recognizing this limitation is crucial to interpret and applications of the scenarios we present. By way of example only, the lingering effects of COVID-19 pandemic may not be fully captured by our research efforts, as the consequences of the pandemic are ongoing and evolving as we write.

Finally, our study focuses mainly on the economic determinants of unemployment. However, cultural, social, and even more subjective factors, are beyond our scope and can significantly influence labour markets, although we will not address them here. Future studies could delve deeper into these aspects to provide a wider and more holistic understanding of unemployment dynamics which, in any case, is no less important than the econometric approach we are following in this thesis.

Still in future research directions that could complement the results presented on this study, the subjectivity approach could be useful to develop a cross-cultural analysis of the specificities of countries. We restrict much of our analysis to the BRICS nations, but we do not go into their cultural and social factors that influence unemployment trends. Understanding how cultural nuances affect labour market dynamics could provide valuable information for policymakers and researchers in the unemployment-related field. Different regions and diverse economic, social, and political contexts could reveal region-specific challenges and contribute to a more nuanced global perspective.

Intersectional perspectives, considering the impact of factors such as gender, age, and socioeconomic status on unemployment, may be other possible path for future research. This approach can unveil disparities within populations, informing more appropriate and targeted interventions. Assessing the potential impact of emerging technologies related with unemployment patterns could be fruitful as well; how automation, artificial intelligence, and other technological advances could influence labour markets and shape the future of employment would be a natural development from what we have presented so far.

In conclusion, although we believe that this thesis contributes significantly to a comprehensive understanding about unemployment phenomenon, recognizing its limitations and suggesting future research directions is essential to promote continued academic exploration and refining policy recommendations in the dynamic landscape of labour markets. In terms of theoretical advances, we understand to have contributed by evaluating the unemployment-related literature and by unveiling the most discussed topics in the area; furthermore, on sequential chapters we were able to establish a macro-to-micro approach to unemployment rates, improving the theoretical framework by elucidating the links between global economic trends and country's employment level.

Considering the empirical contributions, we believe we have addressed a gap revealed by the bibliometric analysis, where it was noticeable that unemployment-related studies, for the Brazilian, Russian, Indian, Chinese, and South African realities, were scarce compared to studies focused on the Europe and USA. For these BRICS countries in particular, the empirical advances lie in generating three scenario-based foresights for their unemployment rates over a 10-year horizon, providing valuable empirical contributions for future policy considerations.

Methodologically, our thesis also offers some contributions. We carry out an integrative use of methods achieving solid results. From clustering analysis and VECM assessments up to the multimethod approach using ANN, ARIMA, ETS, and STL techniques. By following this ensemble of methods, we believe we have achieved an enriched analysis that allows a more granular understanding of the complexities of the unemployment phenomenon whereas offering a solid number of inputs that guided our scenarios' proposition.

The integration of diverse research methods, the nuanced examination of BRICS nations, and the generation of future scenarios serve as valuable contributions to the field. This study not only adds layers to the theoretical foundations of unemployment theme, but also provides actionable empirical insights and methodological frameworks for future research development and real-life policymaking. The interconnectivity of the chapters reinforces the robustness of the study, offering a unified and comprehensive view of the complexities associated with unemployment.

The model developed in this thesis, with its integrated approach that encompasses bibliometric analysis, econometric modelling, and scenario-based forecasting, is promising for replication considering countries outside the BRICS. Researchers can adapt and expand the model by using variables relevant to their specific economic and labour market conditions, extrapolating what we did in the BRICS to other countries and topics beyond unemployment. Detailed documentation of applied techniques, including the use of bibliometric analysis, VECM, clustering, and quantitative forecasting methods, offers a comprehensive guide to replicability.

Although our thesis model, core components and proposed scenarios are designed to be restrictive for the BRICS context, we believe we have offered an adaptable approach that can be tailored to the specific challenges and characteristics of other countries being examined. Acknowledging the dynamic nature of economies and labour market conditions, our methods and thesis model should be seen as a foundation that can be continually refined and improved. We encourage other researchers to adapt and improve our proposals with new data, emerging theories, and by tracking evolving economic scenarios around the world, promoting a collaborative and iterative research process that does not end here.

In conclusion, everything presented and proposed in this thesis demonstrates a solid potential for replicability beyond the context of BRICS nations. We followed a clear research protocol and methodological transparency, paving the way for future researchers to build upon the presented in our study. As we stand at the nexus of past, present, and future labour market dynamics, the relevance of this study may resonate beyond BRICS borders, advocating a collective pursuit of better-informed policies and collaborative efforts to promote sustainable economic growth and eventually mitigate the impacts of unemployment on societies worldwide.

REFERENCES

Ábrahám, Á., de Sousa, J. B., Marimon, R. & Mayr, L. (2023). On the design of a European unemployment insurance system. *European Economic Review*, 156, 104469. Available at: https://econ-papers.upf.edu/papers/1826.pdf

Acemoglu, D. (1995). Public policy in a model of long-term unemployment. *Economica*, 62(246), 161-178. https://doi.org/10.2307/2554901

Acemoglu, D. & Shimer, R. (1999). Efficient unemployment insurance. *Journal of Political Economy*, 107(5), 893-928. https://doi.org/10.1086/250084

Aggarwal, R., Mohanty, S. & Song, F. (1995). Are survey forecasts of macroeconomic variables rational? *Journal of Business*, 99-119. https://doi.org/10.1086/296655

Agnew, C. E. (1985). Bayesian consensus forecasts of macroeconomic variables. *Journal of Forecasting*, 4(4), 363-376. https://doi.org/10.1002/for.3980040405

Agrahari, A., Singh, P., Veer, A., Singh, A., Vidyarthi, A. & Khan, B. (2021). Prognosticating the effect on unemployment rate in the post-pandemic India via time-series forecasting and least squares approximation. *Pattern Recognition Letters*, 152, 172-179. https://doi.org/10.1016/j.patrec.2021.10.012

Albanesi, S. & Sahin, A. (2018). The gender unemployment gap. *Review of Economic Dynamics*, 30, 47-67. https://doi.org/10.1016/j.red.2017.12.005

Alisa, M. (2015). The Relationship between inflation and unemployment: a theoretical discussion about the Philips Curve. *Journal of International Business and Economics*, 3(2), 89-97. http://dx.doi.org/10.15640/

Almeida, F. & Santos, J. D. (2020). The effects of COVID-19 on job security and unemployment in Portugal. *International Journal of Sociology and Social Policy*, 40(9), 995-1003. https://doi.org/10.1108/IJSSP-07-2020-0291

Altindag, D. T., Dursun, B. & Filiz, E. S. (2022). The effect of education on unemployment duration. *Economic Inquiry*, 60(1), 21–42. https://doi.org/10.1111/ecin.13027

Aria, M. & Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, 11(4), 959-975, Elsevier.

Askitas, N. & Zimmermann, K. F. (2009). Google econometrics and unemployment forecasting. *Applied Economics Quarterly*, 55(2), 107–120. https://doi.org/10.3790/aeq.55.2.107

Awogbenle, A. C. & Iwuamadi, K. C. (2010). Youth unemployment: Entrepreneurship development programme as an intervention mechanism. *African Journal of Business Management*, 4(6), 831-835. Available at: http://www.academicjournals.org/AJBM

Baily, M. N. (1978). Some aspects of optimal unemployment insurance. *Journal of Public Economics*, 10(3), 379-402. https://doi.org/10.1016/0047-2727(78)90053-1

Ball, L. M. & Mankiw, N. G. (2002). The NAIRU in theory and practice. *Journal of Economic Perspectives*, 16(4), 115-136. https://doi.org/10.1257/089533002320951000

Balliester, T. & Elsheikhi, A. (2018). *The future of work: A literature review*. Research Department Working Paper, 29, International Labour Organization, 1-62.

Baloch, M. A. & Wang, B. (2019). Analyzing the role of governance in CO2 emissions mitigation: The BRICS experience. *Structural Change and Economic Dynamics*, 51, 119-125. https://doi.org/10.1016/j.strueco.2019.08.007

Banerjee, A., Galiani, S., Levinsohn, J., McLaren, Z. & Woolard, I. (2008). Why has unemployment risen in the New South Africa? *Economics of Transition*, 16(4), 715–740. https://doi.org/10.1111/j.14680351.2008.00340.x

Baptista, R. & Preto, M. T. (2007). The dynamics of causality between entrepreneurship and unemployment. *International Journal of Technology, Policy, and Management*, 7(3), 215-224. https://doi.org/10.1504/IJTPM.2007.015107

Barbosa, C. E., de Lima, Y. O., Costa, L. F. C., dos Santos, H. S., Lyra, A., Argôlo, M., ... & de Souza, J. M. (2022). Future of work in 2050: thinking beyond the COVID-19 pandemic. *European Journal of Futures Research*, 10(1), 1-19. https://doi.org/10.1186/s40309-022-00210-w

Barford, A., Coutts, A. & Sahai, G. (2021). *Youth employment in times of COVID: A global review of COVID-19 policy responses to tackle (un)employment and disadvantage among young people*. ILO: Geneva.

Barnichon, R., Nekarda, C. J., Hatzius, J., Stehn, S. J. & Petrongolo, B. (2012). The ins and outs of forecasting unemployment: Using labor force flows to forecast the labor market [with comments and discussion]. *Brookings Papers on Economic Activity*, 83-131. Available at: https://www.jstor.org/stable/41825365

Barnichon, R. & Figura, A. (2015). *Declining desire to work and downward trends in unemployment and participation*. Working Paper 21252, National Bureau of Economic Research, 1-51. Available at: http://www.nber.org/papers/w21252

Bates, J. M. & Granger, C. W. J. (1969). The combination of forecasts. *Journal of the Operational Research Society*, 20(4), 451-468. https://doi.org/10.1057/jors.1969.103

Battistella, C. & De Toni, A. F. (2015). A methodology of technological foresight: A proposal and field study. *Technological Forecasting and Social Change*, 78(6), 1029-1048. https://doi.org/10.1016/j.techfore.2011.01.006

Bauer, A. & Weber, E. (2020). COVID-19: How much unemployment was caused by the shutdown in Germany? *Applied Economics Letters*, 1-6. https://doi.org/10.1080/13504851.2020.1789544

Bayrak, R. & Tatlı, H. (2018). The determinants of youth unemployment: A panel data analysis of OECD countries. *European Journal of Comparative Economics*, 15(2), 231-248. Available at: https://ejce.liuc.it/18242979201802/182429792018150202.pdf

Bell, D. N. F. & Blanchflower, D. G. (2010). Youth unemployment: Déjà vu? *IZA Discussion Paper*, 4705. Available at: https://www.iza.org/publications/dp/4705/youth-unemployment-deja-vu

Bell, D. N. & Blanchflower, D. G. (2011). Young people and the Great Recession. *Oxford Review of Economic Policy*, 27(2), 241-267. https://doi.org/10.1093/oxrep/grr011

Bell, W. (1997). The purposes of futures studies. *The Futurist*, 31(6), 42-45. Available at: https://www.proquest.com/openview/1bc92ecb43d718fe28dfdd10110d86bd/

Bell, W. (2001). Futures studies comes of age: twenty-five years after. The limits to growth. *Futures*, 33(1), 63-76. https://doi.org/10.1016/S0016-3287(00)00054-9

Betul, G. (2015). An analysis of unemployment determinants in BRIC countries. *International Journal of Business and Social Science*, 6(1), 192-198. Available at: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=5e44ab313a18fcfbc765489f3ece20fcce88ed9a

Bhaduri, A. & Marglin, S. A. (1990). Unemployment and the real wage: The economic basis for contesting political ideologies. *Cambridge Journal of Economics*, 14(4), 375-393. https://doi.org/10.1093/oxfordjournals.cje.a035141

Blanchard, O. J. & Leigh, D. (2013). Growth forecast errors and fiscal multipliers. *The American Economic Review*, 103(3), 117–120. https://doi.org/10.1257/aer.103.3.117

Blanchard, O. J. & Summers, L. H. (1986). Hysteresis and the European unemployment problem. *NBER Macroeconomics Annual*, 1, 15–78. https://doi.org/10.2307/3585159

Blanchard, O. & Gali, J. (2010). Labor markets and monetary policy: A new Keynesian model with unemployment. *American Economic Journal-Macroeconomics*, 2(2), 1-30. https://doi.org/10.1257/mac.2.2.1

Blanchard, O. & Wolfers, J. (2000). The role of shocks and institutions in the rise of European unemployment: The aggregate evidence. *The Economic Journal*, 110(462), C1–C33. Available at: http://www.jstor.org/stable/2565720

Blanchflower, D. G. (2000). Self-employment in OECD countries. *Labour Economics*, 7, 471-505. https://doi.org/10.1016/S0927-5371(00)00011-7

Blanchflower, D. G. & Meyer, B. D. (1994). A longitudinal analysis of the young self-employed in Australia and the United States. *Small Business Economics*, 6, 1–19. https://doi.org/10.1007/BF01066108

Blanchflower, D. G. & Oswald, A. J. (1998). What makes an entrepreneur? *Journal of Labor Economics*, 16(1), 26-60. https://doi.org/10.1086/209881

Blinder, A. (1997). Is there a core of practical macroeconomics that we all should believe? *American Economic Review*, 87(2), 240-243. Available at: https://www.jstor.org/stable/2950924

Blustein, D. L., Duffy, R., Ferreira, J. A., Cohen-Scali, V., Cinamon, R. G. & Allan, B. A. (2020). Unemployment in the time of COVID-19: A research agenda. *Journal of Vocational Behavior*, 19. https://doi.org/10.1016/j.jvb.2020.103436

Boga, S. (2020). Investigating the asymmetry between economic growth and unemployment in Turkey: A hidden cointegration approach. *Journal of Economics Finance and Accounting*, 7(1), 22-33. https://doi.org/10.17261/Pressacademia.2020.1178

Boianovsky, M. & Trauntwein, H-M. (2010). Schumpeter on unemployment. *Journal of Evolutionary Economics*, 20, 233-263. https://doi.org/10.1007/s00191-009-0154-8

Bookstein, A. (1980). Explanations of the bibliometric laws. *Collection Management*, 3(2-3), 151-162. https://doi.org/10.1300/J105v03n02_04

Borger, J. (2023, August 25). *Brics to more than double with admission of six new countries*. The Guardian. Available at: https://www.theguardian.com/business/2023/aug/24/five-brics-nations-announce-admission-of-six-new-countries-to-bloc

Börner, K., Chen, C. & Boyack, K. W. (2003). Visualizing knowledge domains. *Annual Review of Information Science and Technology*, 37(1), 179–255. Available at: https://cns.iu.edu/docs/publications/2003-borner-arist.pdf

Boskin, M. J., Dulberger, E. L., Gordon, R. J., Griliches, Z & Jorgenson, D. W. (1998). Consumer prices, the consumer price index, and the cost of living. *Journal of Economic Perspectives*, 12(1), 3-26. https://doi.org/10.1257/jep.12.1.3

Bosna, J. (2022). Panel data analysis of CEE and SEE: Determinants of unemployment. SciPap, 30(1). https://doi.org/10.46585/sp30011384

Box, G. & Jenkins, G. (1970) *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco.

Box, G. E. P. & Pierce, D. A. (1970). Distribution of residual autocorrelations in Autoregressive-Integrated Moving Average Time Series Models. *Journal of the American Statistical Association*, 65(332), 1509-1526. https://doi.org/10.1080/01621459.1970.10481180

Box, G. E. P., Jenkins, G. M. & Reinsel. (2008). *Time Series Analysis: Forecasting and Control*. John Wiley & Sons Inc., New York.

Breen, R. (2005). Explaining cross-national variation in youth unemployment - Market and institutional factors. *European Sociological Review*, 21(2), 125-134. https://doi.org/10.1093/esr/jci008

Breusch, T. S. & Pagan, A. R. (1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica*, 47(5), 1287–1294. https://doi.org/10.2307/1911963

Broadus, R. N. (1987). Toward a definition of "bibliometrics". *Scientometrics*, 12(5-6), 373-379. https://doi.org/10.1007/BF02016680

Brown, R. G. (1959). *Statistical forecasting for inventory control*. McGraw/Hill.

Brüggemann, R. (2006). Sources of German unemployment: A structural vector error correction analysis. *Empirical Economics*, 31, 409–431. https://doi.org/10.1007/s00181-005-0021-x

Brunner, K., Cukierman, A. & Meltzer, A. (1980). Stagflation, persistent unemployment, and the permanence of economic shocks. *Journal of Monetary Economics*, 6(4), 467–492. https://doi.org/10.1016/0304-3932(80)90002-1

Bruno, M. & Sachs, J. D. (1985). *Economics of worldwide stagflation*. Harvard University Press.

Bryan, M. F. & Cecchetti, S. (1993). The Consumer Price Index as a measure of inflation. *Federal Reserve Bank of Cleveland, Economic Review*, 29(4), 15-24. https://doi.org/10.3386/w4505

Bughin, J. R. (2011). *Nowcasting the Belgian Economy*. Available at SSRN 1903791. http://dx.doi.org/10.2139/ssrn.1903791

Bunn, D. W. & Salo, A. A. (1993). Forecasting with scenarios. *European Journal of Operational Research*, 68(3), 291-303. https://doi.org/10.1016/0377-2217(93)90186-Q

Butkus, M., Matuzeviciute, K., Rupliene, D. & Seputiene, J. (2020). Does Unemployment responsiveness to output change depend on age, gender, education, and the phase of the business cycle? *Economies*, 8(4), 98. https://doi.org/10.3390/economies8040098

Caliendo, M. & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, 22(1), 31-72. https://doi.org/10.1111/j.1467-6419.2007.00527.x

Caliendo, M. & Schmidl, R. (2016). Youth unemployment and active labor market policies in Europe. *IZA J Labor Policy*, 5(1). https://doi.org/10.1186/s40173-016-0057-x

Caliendo, M., Fossen, F. M. & Kritikos, A. S. (2009). Risk attitudes of nascent entrepreneurs–new evidence from an experimentally validated survey. *Small Business Economics*, 32, 153–167. https://doi.org/10.1007/s11187-007-9078-6

Caliendo, M., Künn, S. & Uhlendorff, A. (2016). Earnings exemptions for unemployed workers: The relationship between marginal employment, unemployment duration and job quality. *Labour Economics*, 42, 177-193. https://doi.org/10.1016/j.labeco.2016.07.003

Camba, A. L. (2020). Estimating the nature of relationship of entrepreneurship and business confidence on youth unemployment in the Philippines. *Journal of Asian Finance Economics and Business*, 7(8), 533-542. DOI: 10.13106/jafeb.2020.vol7.no8.533

Caperna, G., Colagrossi, M., Geraci, A., & Mazzarella, G. (2022). A babel of web-searches: Googling unemployment during the pandemic. *Labour Economics*, 74, 102097. https://doi.org/10.1016/j.labeco.2021.102097

Castells, M (1996). *The information age: economy, society, and culture*. Volume I: The rise of the network society. Blackwell, Oxford

Celbiş, M. G. (2023). Unemployment in Rural Europe: A Machine Learning Perspective. *Applied Spatial Analysis and Policy*, 16, 1071–1095. https://doi.org/10.1007/s12061-022-09464-0

Çelik, K. & Lüküslü, G. D. (2018) Unemployment as a chronic problem facing young people in Turkey. *Research and Policy on Turkey*, 3(2), 155-172, https://doi.org/10.1080/23760818.2018.1517449

Cepal, N. (2020). *Producing the consumer price index (CPI) and the COVID-19 pandemic in Latin America and the Caribbean*. Available at: https://hdl.handle.net/11362/45490

Chadwick, M. G. & Sengül, G. (2015). Nowcasting the unemployment rate in Turkey: Let's ask google. *Central Bank Review*, 15(3), 15-40.

Chakraborty, T., Chakraborty, A.K., Biswas., Banerjee, S. & Bhattacharya, S. (2021). Unemployment rate forecasting: A hybrid approach. *Computational Economics*, 57, 183–201. https://doi.org/10.1007/s10614-020-10040-2

Chen, C. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the Association for Information Science and Technology*, 57(3), 359–377.

Cheratian, I., Golpe, A., Goltabar, S., & Iglesias, J. (2019). The unemployment-entrepreneurship nexus: new evidence from 30 Iranian provinces. *International Journal of Emerging Markets*, 15(3), 469–489. https://doi.org/10.1108/ijoem-02-2019-0084

Chereshnev, V. A. & Vasilyeva, A. V. (2013). Model complex of forecasting of interdependent development of migration processes and region labour market. *Economy of Regions*, (3), 272-281. https://doi.org/10.17059/2013-3-25

Cheung, Y. W. & Lai, K. S. (1995). Lag order and critical values of the augmented Dickey–Fuller test. *Journal of Business & Economic Statistics*, 13(3), 277-280. https://doi.org/10.1080/07350015.1995.10524601

Chirici, G. (2012). Assessing the scientific productivity of Italian forest researchers using the Web of Science, SCOPUS and SCIMAGO databases. *Forest-Biogeosciences and Forestry*, 5(3), 101–107.

Chitiga, M., Henseler, M., Mabugu, R. E. & Maisonnave, H. (2021). How COVID-19 pandemic worsens the economic situation of women in South Africa. *The European Journal of Development Research*. https://doi.org/10.1057/s41287-021-00441-w

Choi, H. & Varian, H. (2012). Predicting the present with Google Trends. *Economic Record*, 88, 2–9. https://doi.org/10.1111/j.1475-4932.2012.00809.x

Choi, I. (2001). Unit root tests for panel data. *Journal of International Money and Finance*, 20(2), 249-272. https://doi.org/10.1016/S0261-5606(00)00048-6

Choudhry, M. T., Marelli, E. & Signorelli, M. (2012). Youth unemployment rate and impact of financial crises. *International Journal of Manpower*, 33(1), 76-95. https://doi.org/10.1108/01437721211212538

Clark, A. E. & Oswald, A. J. (1994). Unhappiness and unemployment. *The Economic Journal*, 104(424), 648-659.

Clemen, R. T. (1989). Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting*, 5(4), 559-583. https://doi.org/10.1016/0169-2070(89)90012-5

Cleveland, R. B., Cleveland, W. S., McRae, J. E., & Terpenning, I. (1990). STL: A seasonal-trend decomposition. *Journal of Official Statistics*, 6(1), 3-73. Available at: http://bit.ly/stl1990

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, 74(368), 829-836. https://doi.org/10.1080/01621459.1979.10481038

Coates, J. F. (2000). Scenario Planning. *Technological Forecasting and Social Change*, 65(1), 115–123.

Cobo, M. J., Lopez-Herrera, A. G., Herrera-Viedma, E. & Herrera, F. (2011). Science Mapping Software Tools: Review, analysis, and cooperative study among tools. *Journal of the American Society for Information Science and Technology*, 62(7), 1382-1402. https://doi.org/10.1002/asi.21525

Constant, A. & Zimmermann, K. F. (2006). The making of entrepreneurs in Germany: Are native men and immigrants alike? *Small Business Economics*, 26, 279–300. https://doi.org/10.1007/s11187-005-3004-6

Cook, T. R. & Smalter Hall, A. (2017). Macroeconomic indicator forecasting with deep neural networks. *The Federal Reserve Bank of Kansas City*, Working Paper, (17-11). https://dx.doi.org/10.18651/RWP2017-11

Couch, K. A., Fairlie, R. W. & Xu, H. (2020). Early evidence of the impacts of COVID-19 on minority unemployment. *Journal of Public Economics*, 192, 104287. https://doi.org/10.1016/j.jpubeco.2020.104287

Cromwell, J. B., Labys, W. C. & Terraza, M. (1994): *Univariate Tests for Time Series Models*. Sage, Thousand Oaks, CA, p. 20–22.

Crowder, W. J. & Smallwood, A. (2019). Volatility in productivity and the impact on unemployment. *Applied Economics*, 51(56), 6034-6039. https://doi.org/10.1080/00036846.2019.1654079

Cuaresma, J. C. (2003). Okun's Law revisited. *Oxford Bulletin of Economics and Statistics*, 65(4), 439-451. https://doi.org/10.1111/1468-0084.t01-1-00056

Cuhls, K. (2003). From forecasting to foresight processes—new participative foresight activities in Germany. *Journal of Forecasting*, 22(2-3), 93-111. https://doi.org/10.1002/for.848

D'Amuri, F. & Marcucci, J. (2010). 'Google It!' Forecasting the US Unemployment Rate with A Google Job Search Index. *SSRN Electronic Journal*, https://doi.org/10.2139/ssrn.1594132

Da, Z., Engelberg, J. & Gao, P. (2015). The Sum of all FEARS investor sentiment and asset prices. *The Review of Financial Studies*, 28(1), 1–32, https://doi.org/10.1093/rfs/hhu072

Đalić, I. & Terzić, S. (2021). Violation of the assumption of homoscedasticity and detection of heteroscedasticity. *Decision Making: Applications in Management and Engineering*, 4(1), 1-18. https://doi.org/10.31181/dmame2104001d

De Gooijer, J. G. & Hyndman, R. J. (2006). 25 years of time series forecasting. *International Journal of Forecasting*, 22(3), 443-473. https://doi.org/10.1016/j.ijforecast.2006.01.001

de'Robertis, A. (2016). *La Russia ei BRICS e l'Ordine Internazionale* (Vol. 15). Edizioni Nuova Cultura.

Devonshire-Ellis, C. (2023, March 27). The BRICS Has Overtaken The G7 In Global GDP. *Silk Road Briefing*. Available at: https://www.silkroadbriefing.com/news/2023/03/27/the-brics-has-overtaken-the-g7-in-global-gdp/

Dickey, D. A. & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366a), 427-431. https://doi.org/10.1080/01621459.1979.10482531

Dilmaghani, M. (2019). Workopolis or The Pirate Bay: what does Google Trends say about the unemployment rate? *Journal of Economic Studies*, 46(2), 422-445. https://doi.org/10.1108/JES-11-2017-0346

Dobre, I. & Alexandru, A. A. (2008). Modelling unemployment rate using Box-Jenkins procedure. *Journal of Applied Quantitative Methods*, 3(2), 156-166.

Donthu, N., Kumar, S., Mukherjee, D., Pandey, N. & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research*, 133, 285-296. https://doi.org/10.1016/j.jbusres.2021.04.070

Draskovic, V., Pupavac, J., Delibasic, M. & Koltun, L. (2021). Employment in Croatia: Insigths into the past and the future. *Journal of International Studies*, 14(1), 117-128. doi:10.14254/2071-8330.2021/14-1/8

du Toit, A. (2021). Harnessing education through entrepreneurship in consumer studies to address youth unemployment in South Africa. *Journal of Consumer Sciences*, 49, 1-14.

Duggan, N. & Azalia, J. C. L. (2020). From Yekaterinburg to Brasilia: BRICS and the G20, road to nowhere? *Revista Brasileira De Política Internacional*, 63(1), e009. https://doi.org/10.1590/0034-7329202000109

Durbin, J. & Watson, G. S. (1971). Testing for serial correlation in Least Squares Regression. III. *Biometrika*, 58(1), 1–19. https://doi.org/10.2307/2334313

Earle, J. S. & Sakova, Z. (2000). Business start-ups or disguised unemployment? Evidence on the character of self-employment from transition economies. *Labour Economics*, 7(5), 575-601. https://doi.org/10.1016/S0927-5371(00)00014-2

Echchakoui, S. (2020). Why and how to merge Scopus and Web of Science during bibliometric analysis: the case of sales force literature from 1912 to 2019. *Journal of Marketing Analytics*, 1-20. https://doi.org/10.1057/s41270-020-00081-9

Egghe, L. (2006). Theory and practice of the G-index. *Scientometrics*, 69(1), 131–152.

Eisenberg, P. & Lazarsfeld, P. F. (1938). The psychological effects of unemployment. *Psychological Bulletin*, 35(6), 358–390. https://doi.org/10.1037/h0063426

Engen, E. M. & Gruber, J. (2001). Unemployment insurance and precautionary saving. *Journal of Monetary Economics*, 47(3), 545-579. https://doi.org/10.1016/S0304-3932(01)00051-4

Fana, M., Pérez, S. T. and Fernández-Macías, E. (2020). Employment impact of Covid-19 crisis: From short term effects to long terms prospects. *Journal of Industrial and Business Economics*, 47, 391–410. https://doi.org/10.1007/s40812-020-00168-5

Farber, H. S., Herbst, C. M., Silverman, D. & Von Wachter, T. (2019). Whom do employers want? The role of recent employment and unemployment status and age. *Journal of Labor Economics*, 37(2), 323-349. https://doi.org/10.1086/700184

Faria, J. R., Cuestas, J. C., Gil-Alana, L. A. (2009). Unemployment and entrepreneurship: A cyclical relation? *Economics Letters*, 105(3), 318-320. https://doi.org/10.1016/j.econlet.2009.09.004

Farrar, D. E. & Glauber, R. R. (1967). Multicollinearity in regression analysis: the problem revisited. *The Review of Economic and Statistics*, 49(1), 92-107. https://doi.org/10.2307/1937887

Fergnani, A. (2018). Mapping futures studies scholarship from 1968 to present: A bibliometric review of thematic clusters, research trends, and research gaps. *Futures*, 105, 104-123. https://doi.org/10.1016/j.futures.2018.09.007

Fergnani, A. (2022). Corporate foresight: A new frontier for strategy and management. *Academy of Management Perspectives*, 36(2), 820-844. https://doi.org/10.5465/amp.2018.0178

Floros, C. (2005). Forecasting the UK unemployment rate: model comparisons. *International Journal of Applied Econometrics and Quantitative Studies*, 2(4), 57-72.

Fondeur, Y. & Karamé, F. (2013). Can google data help predict French youth unemployment? *Economic Modelling*, 30, 117–125. https://doi.org/10.1016/j.econmod.2012.07.017

Fratianni, M., Gallegati, M. & Giri, F. (2022). The medium-run Phillips curve: A time-frequency investigation for the UK. *Journal of Macroeconomics*, 73. https://doi.org/10.1016/j.jmacro.2022.103450

Frey, B. J. & Dueck, D. (2007) Clustering by passing messages between data points. *Science*, 315(5814), p. 972-976. https://doi.org/10.1126/science.1136800

Friedman, M. (1968). The role of monetary policy. *American Economic Review*, 58(1), 1–17.

Friedman, M. (1977). Inflation and unemployment. *Journal of Political Economy*, 85(3), 451-472. Available at: https://www.jstor.org/stable/1830192

Funke, M. (1992). Time-series forecasting of the German unemployment rate. *Journal of Forecasting*, 11(2), 111-125. https://doi.org/10.1002/for.3980110203

Gagolewski, M. (2011). Bibliometric impact assessment with R and the CITAN package. *Journal of Informetrics*, 5(4), 678–692.

Galea, S. & Abdalla. S. M. (2020). COVID-19 pandemic, unemployment, and civil unrest: Underlying deep racial and socioeconomic divides. *JAMA*, 324(3), 227–228. https://doi.org/10.1001/jama.2020.11132

Gardner Jr, E. S., & McKenzie, E. D. (1985). Forecasting trends in time series. *Management Science*, 31(10), 1119-1334. https://doi.org/10.1287/mnsc.31.10.1237

Garfield, E. (1971). The mystery of the transposed journal lists: Wherein Bradford's law of scattering is generalized according to Garfield's law of concentration. *Current Contents*, (1) 222–223.

Germay, M. D. (2016). *Modeling and Forecasting Unemployment Rate in Sweden Using Various Econometric Measures* [Master's Thesis, Örebro University]. Retrieved from: http://urn.kb.se/resolve?urn=urn:nbn:se:oru:diva-51425

Giulietti, C., Guzi, M., Kahanec, M. & Zimmermann, K. F. (2013). Unemployment benefits and immigration: evidence from the EU. *International Journal of Manpower*, 34(1). https://doi.org/10.1108/01437721311319638

Glaeser, E. L., Scheinkman, J. A. & Shleifer, A. (1995). Economic growth in a cross-section of cities, *Journal of Monetary Economics*, 36, 117-143. https://doi.org/10.1016/0304-3932(95)01206-2

Glynatsi, N. E. & Knight, V.A. (2021). A bibliometric study of research topics, collaboration, and centrality in the iterated prisoner's dilemma. *Humanities & Social Sciences Communications*, 8(45). https://doi.org/10.1057/s41599-021-00718-9

Gneiting, T. & Katzfuss, M. (2014). Probabilistic forecasting. *Annual Review of Statistics and Its Application*, 1, 125-151. https://doi.org/10.1146/annurev-statistics-062713-085831

Gogas, P., Papadimitriou, T. & Sofianos, E. (2022). Forecasting unemployment in the euro area with machine learning. *Journal of Forecasting*, 41(3), 551-566. https://doi.org/10.1002/for.2824

Görlich, D., Stepanok, I. & Al-Hussami, F. (2013). Youth unemployment in Europe and the world: Causes, consequences, and solutions. *Kiel Policy Brief*, (No. 59). Available at: http://hdl.handle.net/10419/70102

Gough, K., Langevang, T. & Owusu, G. (2013). Youth employment in a globalising world. *International Development Planning Review*, 35(2), 91–102. https://doi.org/10.3828/idpr.2013.7

Gowda, K. C. & Diday, E. (1991). Symbolic clustering using a new dissimilarity measure. *Pattern Recognition*, 24(6), 567-578. https://doi.org/10.1016/0031-3203(91)90022-W

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3), 424–438. https://doi.org/10.2307/1912791

Greblikaitė, J., Sroka, W. & Grants, J. (2015). Development of social entrepreneurship in European Union: policy and situation of Lithuania and Poland. *Transformations in Business & Economics*, 14, 376-396.

Gregg, P. (2001). The impact of youth unemployment on adult unemployment in the NCDS. *The Economic Journal*, 111(475), 626-653. https://doi.org/10.1111/1468-0297.00666

Grima, S. & Caruana, L. (2017). The effect of the financial crisis on emerging markets: A comparative analysis of the stock market situation before and after. *European Research Studies Journal*, 20(4), 727-753. https://doi.org/10.35808/ersj/923

Guisinger, A. Y., Hernandez-Murillo, R., Owyang, M. T. & Sinclair, T. M. (2018). A state-level analysis of Okun's law. *Regional Science and Urban Economics*, 68, 239-248. https://doi.org/10.1016/j.regsciurbeco.2017.11.005

Gupta, R., Pierdzioch, C. & Salisu, A. A. (2022). Oil-price uncertainty and the UK unemployment rate: A forecasting experiment with random forests using 150 years of data. *Resources Policy*, 77, 102662. https://doi.org/10.1016/j.resourpol.2022.102662

Hammarström, A. (1994). Health consequences of youth unemployment—review from a gender perspective. *Social Science & Medicine*, 38(5), 699-709. https://doi.org/10.1016/0277-9536(94)90460-X

Hartigan, J. A. & Wong, M. A. (1979). Algorithm AS 136: A K-Means clustering algorithm. Journal of the Royal Statistical Society. *Series C (Applied Statistics)*, 28(1), 100–108. https://doi.org/10.2307/2346830

Hashimi, S. M., Khushik, A. G., Gilal, M. A. & Yongliang, Z. (2021). The impact of GDP and its expenditure components on unemployment within BRICS countries: Evidence of Okun's law from aggregate and disaggregated approaches. *SAGE Open*, 11(2), 1-11. https://doi.org/10.1177/21582440211023423

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences*, 102(46), 16569-16572.

Hjazeen, H., Seraj, M. & Ozdeser, H. (2021). The nexus between the economic growth and unemployment in Jordan. *Future Business Journal*, 7, 42. https://doi.org/10.1186/s43093-021-00088-3

Hogarth, R. M., & Makridakis, S. (1981). Forecasting and Planning: An Evaluation. *Management Science*, 27(2), 115–138. https://doi.org/10.1287/mnsc.27.2.115

Holt, C. C. (2004). Forecasting seasonals and trends by exponentially weighted moving averages. *International Journal of Forecasting*, 20(1), 5-10. https://doi.org/10.1016/j.ijforecast.2003.09.015

Horvath, J. & Zhong, J. (2019). Unemployment dynamics in emerging countries: Monetary policy and external shocks. *Economic Modelling*, 76, 31-49.

Huang, G., Huang, H-C., Liu, X. & Zhang, J. (2019). Endogeneity in Okun's law. *Applied Economics Letters*. https://doi.org/10.1080/13504851.2019.1646861

Huang, X. (2015). *Forecasting the U.S. unemployment rate with Job Openings Index* [Master's thesis, University of Rhode Island]. The University of Rhode Island Commons. https://digitalcommons.uri.edu/theses/699

Hyndman, R. J. & Khandakar, Y. (2008). Automatic time series forecasting: The forecast Package for R. *Journal of Statistical Software*, 27(3), 1–22. https://doi.org/10.18637/jss.v027.i03

Hyndman, R. J. & Athanasopoulos, G. (2021) *Forecasting: principles and practice*, 3rd edition, OTexts: Melbourne, Australia. OTexts.com/fpp3. Accessed on <3rd of March, 2022>.

Ignaszak, M., Jung P. & Kuester, K. (2020). Federal unemployment reinsurance and local labor-market policies. *ECONtribute Discussion Papers Series* 040, University of Bonn and University of Cologne, Germany.

Im, K. S., Pesaran, M. H. & Shin, Y. (2003). Testing for unit roots in heterogenous panels. *Journal of Econometrics*, 115(1), 53-74. https://doi.org/10.1016/S0304-4076(03)00092-7

International Labour Organization (ILO) (2015). *World employment and social outlook: Trends 2015*. Geneva: ILO Publications.

International Labour Organization (ILO) (2022). *Indicator description: Unemployment rate*. accessed on 21 April 2022, https://ilostat.ilo.org/resources/concepts-and-definitions/description-unemployment-rate/

International Labour Organization (ILO) (2022). *Statistics on the working-age population and labour force*. accessed 14 April 2022, https://ilostat.ilo.org/topics/population-and-labour-force/

Jahoda, M. (1981). Work, employment, and unemployment: Values, theories, and approaches in social research. *American Psychologist*, 36(2), 184–191. https://doi.org/10.1037/0003-066X.36.2.184

Jain, A. K., Murty, M. N. & Flynn, P. J. (1999). Data clustering: A review. *ACM Computing Surveys*, 31(3), 264-323. https://doi.org/10.1145/331499.331504

Jarque, C. M. & Bera, A. K. (1980). Efficient test for normality, homoscedasticity, and serial independence of residuals. *Economic Letters*, 6(3), 255-259. https://doi.org/10.1016/0165-1765(80)90024-5

Javaid, A., Akbar, A. & Nawaz, S. (2018) A Review on Human Development Index. *Pakistan Journal of Humanities and Social Sciences*, 6(3), 357–369. Available at SSRN: https://ssrn.com/abstract=3432413

Jefferson, M. (2012). Shell scenarios: What really happened in the 1970s and what may be learned for current world prospects. *Technological Forecasting and Social Change*, 79(1), 186-197. https://doi.org/10.1016/j.techfore.2011.08.007

Jia, S. & Meng, X. (1997). Forecasting analysis of employment in China. *Chinese Journal of Population Science*, 9(2), 115–121.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian Vector Autoregressive Models. *Econometrica*, 59(6). https://doi.org/10.2307/2938278

Johansen, S. (1995). *Likelihood-based inference in cointegrated Vector Autoregressive Models*, Oxford University Press, Oxford.

Johnson, G. R. (1917). Unemployment and feeblemindedness. *The Journal of Delinquency*, 2(2), 59-72.

Kao, C. (1999). Spurious regression and residual-based tests for cointegration in panel data. *Journal of Econometrics*, 90(1), 1–44. https://doi.org/10.1016/S0304-4076(98)00023-2

Karanassou, M., Sala, H. & Snower, D. J. (2010). Phillips curves and unemployment dynamics: A critique and a holistic perspective. *Journal of Economic Surveys*, 24(1), 1-51. https://doi.org/10.1111/j.1467-6419.2009.00598.x

Katris, C. (2020). Prediction of unemployment rates with time series and machine learning techniques. *Computational Economics*, 55, 673–706. https://doi.org/10.1007/s10614-019-09908-9

Kawaguchi, D. & Murao, T. (2014). Labor-market institutions and long-term effects of youth unemployment. *Journal of Money, Credit and Banking*, 46(2), 95-116. https://doi.org/10.1111/jmcb.12153

Kawohl, W. & Nordt, C. (2020). COVID-19, unemployment, and suicide. *The Lancet (Psychiatry)*, 7(5), 389-390. https://doi.org/10.1016/S2215-0366(20)30141-3

Kessler, M. M. (1963). Bibliographic coupling between scientific papers. *American Documentation*, 14(1), 10-25. https://doi.org/10.1002/asi.5090140103

Kim, J. H. (2019). Multicollinearity and misleading statistical results. *Korean journal of Anaesthesiology*, 72(6), 558–569. https://doi.org/10.4097/kja.19087

Klein, L. R. (1984). The importance of the forecast. *Journal of Forecasting*, 3(1), 1-9. https://doi.org/10.1002/for.3980030102

Kumar, S., Viral, R., Deep, V., Sharma, P., Kumar, M., Mahmud, M. & Stephan, T. (2021). Forecasting major impacts of COVID-19 pandemic on country-driven sectors: challenges, lessons, and future roadmap. *Personal and Ubiquitous Computing*, 27, 807-830. https://doi.org/10.1007/s00779-021-01530-7

Kuosa, T. (2011). Evolution of futures studies. *Futures*, 43(3), 327-336. https://doi.org/10.1016/j.futures.2010.04.001

Kurita, T. (2010). A Forecasting Model for Japan's Unemployment Rate. *Eurasian Journal of Business and Economics*, 3(5), 127-134.

Laffineur, C., Barbosa, S. D., Fayolle, A. & Nziali, E. (2017). Active labor market programs' effects on entrepreneurship and unemployment. *Small Business Economics*, 49(4), 889-918. https://doi.org/10.1007/s11187-017-9857-7

Lalthapersad-Pillay, P. (2014). Gender influences in the labour market: The case of BRICS. Mediterranean *Journal of Social Sciences*, 5(10), 146. https://doi.org/10.5901/mjss.2014.v5n10p146

Lasso, F. & Snijders, S. (2016). The power of Google search data; an alternative approach to the measurement of unemployment in Brazil. *Student Undergraduate Research E-journal*, 2.

Lavoie, M. & Stockhammer, E. (2013). *Wage-led growth: Concept, theories, and policies*. In Wage-led growth: An equitable strategy for economic recovery (p. 13-39). London: Palgrave Macmillan, UK.

Lee, J. (2000). The robustness of Okun's Law: Evidence from OECD countries. *Journal of Macroeconomics*, 22(2), 331-356. https://doi.org/10.1016/S0164-0704(00)00135-X

Lee, S., Schmidt-Klau, D. & Verick, S. (2020). The labour market impacts of the COVID-19: A global perspective. *The Indian Journal of Labour Economics*, 63, 11-15. https://doi.org/10.1007/s41027-020-00249-y

Lentz, R. & Tranaes, T. (2005). Job search and savings: Wealth effects and duration dependence. *Journal of Labour Economics*, 23(3), 467-489. https://doi.org/10.1086/430284

Lentz, R. & Mortensen, D. T. (2005). Productivity growth and worker reallocation. *International Economic Review*, 46(3), 731–749. Available at: http://www.jstor.org/stable/3663491

Levenko, N. (2020). Perceived uncertainty as a key driver of household saving. International *Review of Economics & Finance*, 65, 126-145. https://doi.org/10.1016/j.iref.2019.10.005

Levin, A., Lin, C. F. & Chu, C. S. J. (2002). Unit root tests in panel data: asymptotic and finite-sample properties. *Journal of Econometrics*, 108(1), p. 1-24. https://doi.org/10.1016/S0304-4076(01)00098-7

Li, G., Wu, D. C., Zhou, M. & Liu, A. (2019). The combination of interval forecasts in tourism. *Annals of Tourism Research*, 75, 363-378. https://doi.org/10.1016/j.annals.2019.01.010

Liew, V. K-S. (2004) Which lag length selection criteria should we employ? *Economics Bulletin*, 3(33), p. 1−9.

Lila, M. F., Meira, E. & Oliveira, F. L. C. (2022). Forecasting unemployment in Brazil: A robust reconciliation approach using hierarchical data. *Socio-Economic Planning Sciences*, 82, 101298. https://doi.org/10.1016/j.seps.2022.101298

Linnenluecke, M. K., Marrone, M. & Singh, A. K. (2020). Conducting systematic literature reviews and bibliometric analyses. *Australian Journal of Management*, 45(2), 175-194. https://doi.org/10.1177/0312896219877678.

Ljung, G. M. & Box, G. E. P. (1978). On a measure of lack of fit in time series models, *Biometrika*, 65(2), 297–303, https://doi.org/10.1093/biomet/65.2.297

Lockwood, B., Miller, M. & Zhang, L. (1998). Designing monetary policy when unemployment persists. *Economica*, 65(259), 327-345. https://doi.org/10.1111/1468-0335.00132

Long, C. D. (1942). The concept of unemployment. *The Quarterly Journal of Economics*, 57(1), 1-30. Available at: http://www.jstor.org/stable/1881811

Long, J. S. & Ervin, L. H. (2000). Using heteroscedasticity consistent standard errors in the linear regression model. *The American Statistician*, 54(3), 217-224. https://doi.org/10.1080/00031305.2000.10474549

Maas, B. (2020). Short-term forecasting of the US unemployment rate. *Journal of Forecasting*, 39(3), 394–411. https://doi.org/10.1002/for.2630

Macgregor, D. H. (1907). Labour exchanges and unemployment. *The Economic Journal*, 17(68), 585-589. Available at: https://www.jstor.org/stable/2220875

Madley-Dowd, P., Hughes, R., Tilling, K. & Heron, J. (2019). The proportion of missing data should not be used to guide decisions on multiple imputation. *Journal of Clinical Epidemiology*, 110, 63-73. https://doi.org/10.1016/j.jclinepi.2019.02.016

Major, E., Asch, D. & Cordey-Hayes, M. (2001). Foresight as a core competence. *Futures*, 33(2), 91-107. https://doi.org/10.1016/S0016-3287(00)00057-4

Makridakis, S. (1993). Accuracy measures: Theoretical and practical concerns. *International Journal of Forecasting*, 9(4), 527-529. https://doi.org/10.1016/0169-2070(93)90079-3

Mansfield, E. R. & Helms, B. P. (1982). Detecting multicollinearity. *The American Statistician*, 36(3a), 158-160. https://doi.org/10.1080/00031305.1982.10482818

Marcellino, M. (2004). Forecasting EMU macroeconomic variables. *International Journal of Forecasting*, 20(2), 359-372. https://doi.org/10.1016/j.ijforecast.2003.09.003

Marconi, G., Beblavý, M. & Maselli, I. (2016) Age effects in Okun's law with different indicators of unemployment. *Applied Economics Letters*, 23(8), 580-583. https://doi.org/10.1080/13504851.2015.1090540

Margo, R. A. (1993). Employment and unemployment in the 1930s. *Journal of Economic Perspectives*, 7(2), 41-59. https://doi.org/10.1257/jep.7.2.41

Marinković, M., Al-Tabbaa, O., Khan, Z. & Wu, J. (2022). Corporate foresight: A systematic literature review and future research trajectories. *Journal of Business Research*, 144, 289-311. https://doi.org/10.1016/j.jbusres.2022.01.097

Marques, P. & Hoerisch, F. (2020). Understanding massive youth unemployment during the EU sovereign debt crisis: a configurational study. *Comparative European Politics*, 18(2), 233-255. https://doi.org/10.1057/s41295-019-00184-3

McClure, P. K. (2018). "You're Fired," says the robot: The rise of automation in the workplace, technophobes, and fears of unemployment. *Social Science Computer Review*, 36(2), 139–156. https://doi.org/10.1177/0894439317698637

McGeever, J. (2021, June 30). *Brazil's unemployment rate stuck at historic high 14.7%.* Reuters. https://www.reuters.com/world/americas/brazils-unemployment-rate-historic-high-147-quarter-through-april-2021-06-30/

McLaren, N. & Shanbhogue, R. (2011). Using internet search data as economic indicators. *Bank of England Quarterly Bulletin*, (2011), Q2. http://dx.doi.org/10.2139/ssrn.1865276

Menard, S. (1995). *Applied Logistic Regression Analysis*: Sage University Series on Quantitative Applications in the Social Sciences. Thousand Oaks, CA: Sage.

Meng, X. (2003). Unemployment, consumption smoothing, and precautionary saving in urban China. *Journal of Comparative Economics*, 31, 465–485.

Merz, M. (1995). Search in the labor market and the real business cycle. *Journal of Monetary Economics*, 36(2), 269-300. https://doi.org/10.1016/0304-3932(95)01216-8

Metzler, L. A. (1951). Wealth, saving, and the rate of interest. *Journal of Political Economy*, 59(2), 93–116. https://doi.org/10.1086/257052

Meyer, B. D. (1990). Unemployment insurance and unemployment spells. *Econometrica*, 58(4), 757–782. https://doi.org/10.2307/2938349

Michaelides, M., & Davis, S. (2020). From unemployment to self-employment: The role of entrepreneurship training. *IZA Journal of Labor Policy*, 10(1). https://doi.org/10.2478/izajolp-2020-0018

Miles, I. (2010). The development of technology foresight: A review. *Technological Forecasting and Social Change*, 77(9), 1448-1456. https://doi.org/10.1016/j.techfore.2010.07.016

Millán, J. M., Congregado, E. & Román, C. (2012). Determinants of self-employment survival in Europe. *Small Business Economics*, 38, 231-258. https://doi.org/10.1007/s11187-010-9260-0

Mingers, J. & Lipitakis, E. (2010). Counting the citations; a comparison of Web of Science and Google Scholar in the field of business and management. *Scientometrics*, 85(2), 613–625.

Miyamoto, H. & Suphaphiphat, N. (2021). Mitigating long-term unemployment in Europe. *IZA Journal of Labor Policy*, 11(1). https://doi.org/10.2478/izajolp-2021-0003

Mody, A., Ohnsorge, F. & Sandri, D. (2012). Precautionary savings in the great recession. *IMF Economic Review*, 60(1), 114-138. https://doi.org/10.1057/imfer.2012.5

Montgomery, A. L., Zarnowitz, V., Tsay, R. S. & Tiao, G. C. (1998). Forecasting the U.S. unemployment rate. *Journal of the American Statistical Association*, 93(442), 478–493. https://doi.org/10.2307/2670094

Morgan, I. G. (1976). Stock prices and heteroscedasticity. *The Journal of Business*, 49(4), 496-508. https://doi.org/10.1016/S0167-6296(98)00025-3

Mortensen, D. T. & Pissarides, C. A. (1994). Job creation and job destruction in the Theory of Unemployment. *The Review of Economic Studies*, 61(3), 397–415. https://doi.org/10.2307/2297896

Moyen, S., Stähler, N. & Winkler, F. (2019). Optimal unemployment insurance and international risk sharing. *European Economic Review*, 115, 144-171. https://doi.org/10.1016/j.euroecorev.2019.03.004

Mroz, T. A. & Savage, T. H. (2006). The long-term effects of youth unemployment. *Journal of Human Resources*, 41(2), 259-293.

Murtin, F. & Robin, J-M. (2018). Labor market reforms and unemployment dynamics. *Labour Economics*, 50, 3-19. https://doi.org/10.1016/j.labeco.2016.05.025

Mushtaq, R. (2011). *Augmented Dickey-Fuller Test*. Available at SSRN: https://ssrn.com/abstract=1911068 or http://dx.doi.org/10.2139/ssrn.1911068

Naccarato, A., Falorsi, S., Loriga, S. & Pierini, A. (2018). Combining official and Google Trends data to forecast the Italian youth unemployment rate. *Technological Forecasting and Social Change*, 130, 114-122. https://doi.org/10.1016/j.techfore.2017.11.022

Nakajima, T. (2006). Unemployment and indeterminacy. *Journal of Economic Theory*, 126(1), 314-327. https://doi.org/10.1016/j.jet.2004.08.003

Newbold, P. (1983). ARIMA model building and the time series analysis approach to forecasting. *Journal of Forecasting*, 2(1), 23-35. https://doi.org/10.1002/for.3980020104

Nguyen, A., Guttentag, A.; Li, D., & van Meijgaard, J. (2022). The impact of job and insurance loss on prescription drug use: a panel data approach to quantifying the health consequences of unemployment during the Covid-19 pandemic. *International Journal of Health Services*, 1-11. https://doi.org/10.1177/00207314221078749

Nickell, S. (1979). Education and lifetime patterns of unemployment. *Journal of Political Economy*, 87(5), 117-131. https://doi.org/10.1086/260825

Nickell, S. (1997). Unemployment and labor market rigidities: Europe versus North America. *Journal of Economic Perspectives*, 11(3), 55-74. https://doi.org/10.1257/jep.11.3.55

Nickell, S., Nunziata, L. & Wolfgang, O. (2005). Unemployment in the OECD since the 1960s. What do we know? *The Economic Journal*, 115(500), 1-27. https://doi.org/10.1111/j.1468-0297.2004.00958.x

Nkoane, S. S. & Seeletse, S. M. (2021). Forecasting unemployment rate in South Africa with unexpected events using robust estimators. *International Journal of Economics and Finance Studies*, 13(2), 199-222. https://doi.org/10.34109/ijefs.20212010

O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, 41(5), 673–690. https://doi.org/10.1007/s11135-006-9018-6

O'Campo, P.; Molnar, A.; Ng, E.; Renahy, E.; Mitchell, C.; Shankardass, K.; St. John, A.; Bambra, C.; Muntaner, C. (2015). Social welfare matters: A realist review of when, how, and why unemployment insurance impacts poverty and health. *Social Science and Medicine*, 132, 88-94. https://doi.org/10.1016/j.socscimed.2015.03.025

O'Neill, J. (2001). *Building better global economic BRICs* (Vol. 66, pp. 1-16). New York: Goldman Sachs.

Okun, A. M. (1962). *Potential GNP: Its measurement and significance*. Proceedings of the Business and Economic Statistics Section, American Statistical Association.

Ormerod, P., Rosewell, B. & Phelps, P. (2013). Inflation/unemployment regimes and the instability of the Phillips curve. *Applied Economics*, 45(12), 1519-1531. https://doi.org/10.1080/00036846.2011.628299

Orphanides, A. (2002). Monetary-policy rules and the Great Inflation. *American Economic Review*, 92(2), 115-120. https://doi.org/10.1257/000282802320189104

Örsal, D. D. K. (2007). *Comparison of panel cointegration tests*, SFB 649 Discussion Paper, No. 029, Humboldt University of Berlin, Collaborative Research Center 649 - Economic Risk, Berlin. Available at: http://hdl.handle.net/10419/25201

Patel, D. P. C. & Choga, I. (2018). *Determinants of youth unemployment in South Africa*. Proceedings of Economics and Finance Conferences 7108913, International Institute of Social and Economic Sciences. https://doi.org/10.20472/EFC.2018.009.016

Pattanaik, F. & Nayak, N. C. (2013). Trends and forecasting of employment intensity of growth in India. *Journal of the Asia Pacific Economy*, 18(3), 438-459. https://doi.org/10.1080/13547860.2012.742693

Paul, K. I. & Moser, K. (2009). Unemployment impairs mental health: Meta-analyses. *Journal of Vocational Behavior*, 74(3), 264-282.

Pedroni, P. (1999). Critical values for cointegration tests in heterogeneous panels with multiple regressors. *Oxford Bulletin of Economics & Statistics*, 61, 653–670. https://doi.org/10.1111/1468-0084.0610s1653

Petropoulos, F., Apiletti, D., Assimakopoulos, V., Babai, M. Z., Barrow, D. K., Taieb, S. B., ... & Ziel, F. (2022). Forecasting: theory and practice. *International Journal of Forecasting*, 38(3), 705-871. https://doi.org/10.1016/j.ijforecast.2021.11.001

Phillips, A. W. (1958). The relation between unemployment and the rate of change of money wage rates in the United Kingdom, 1861-1957. *Economica*, 25(100), 283-299. https://doi.org/10.2307/2550759

Pissarides, C. A. & McMaster, I. (1990). Regional migration, wages, and unemployment: Empirical evidence and implications for policy. *Oxford Economic Papers*, 42(4), 812-831. Available at: https://www.jstor.org/stable/2663123

Pissarides, C. & Vallanti, G. (2004). *Productivity growth and employment: Theory and panel estimates*. Centre for Economic Performance, 663, London School of Economics and Political Science, London, United Kingdom.

Prachowny, M. F. J. (1993). Okun's Law: Theoretical foundations and revisited estimates. *Review of Economic and Statistics*, 75(2), 331-336. https://doi.org/10.2307/2109440

Pritchard, A. (1969). Statistical bibliography or bibliometrics? *Journal of Documentation*, 25(4), 348–349.

Radulescu, I. G., Panait, M. & Voica, C. (2014). BRICS countries challenge to the world economy new trends. *Procedia Economics and Finance*, 8, 605-613. https://doi.org/10.1016/S2212-5671(14)00135-X

Ribaj, A. & Mexhuani, F. (2021) The impact of savings on economic growth in a developing country (The case of Kosovo). *Journal of Innovation and Entrepreneurship*, 10. https://doi.org/10.1186/s13731-020-00140-6

Riddell, W. C. & Song, X. (2011). The impact of education on unemployment incidence and re-employment success: Evidence from the U.S. labour market. *Labour Economics*, 18(4), 453-463. https://doi.org/10.1016/j.labeco.2011.01.003

Ridha, R. N., Burhanuddin, B. & Wahyu, B. P. (2017). Entrepreneurship intention in agricultural sector of young generation in Indonesia. *Asia Pacific Journal of Innovation and Entrepreneurship,* 11(1), 76-89. https://doi.org/10.1108APJIE-04-2017-022

Rinne, U. & Zimmermann, K. F. (2012). Another economic miracle? The German labor market and the Great Recession. *IZA Journal of Labor Policy*, 1(3).

Rohrbeck, R., Battistella, C. & Huizingh, E. (2015). Corporate foresight: An emerging field with a rich tradition. *Technological Forecasting and Social Change*, 101, 1-9. https://doi.org/10.1016/j.techfore.2015.11.002

Rosenbaum, P. R. & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41–55, https://doi.org/10.1093/biomet/70.1.41

Rubio-Aparicio, M., López-López, J. A., Viechtbauer, W., Marín-Martínez, F., Botella, J. & Sánchez-Meca, J. (2020). Testing categorical moderators in mixed-effects meta-analysis in the presence of heteroscedasticity. *The Journal of Experimental Education*, 88(2), 288-310. https://doi.org/10.1080/00220973.2018.1561404

Sahnoun, M. & Abdennadher, C. (2019). Causality between inflation, economic growth, and unemployment in North African countries. *Economic Alternatives*, 1, 77-92.

Sánchez, A. D.; Rama, M. del R. & García, J. A. (2017). Bibliometric analysis of publications on wine tourism in the databases Scopus and WoS. *European Research on Management and Business Economics*, 231, 8–15.

Schaars, S. P. (1987). How to develop and use scenarios. *Futures*, 20(1), 105-114. https://doi.org/10.1016/0024-6301(87)90038-0

Schmid, G. & Gazier, B. (Eds.). (2002). *The dynamics of full employment: Social Integration through transitional labour markets*. Edward Elgar Publishing.

Schmillen, A. & Umkehrer, M. (2017). The scars of youth: Effects of early-career unemployment on future unemployment experience. *International Labour Review*, 156(3–4), 465-494.

Schreiber, M. (2008). An empirical investigation of the g-index for 26 physicists in comparison with the h-index, the A-index, and the R-index. *Journal of the American Society for Information Science and Technology*, 59(9), 1513.

Schumpeter, J. A. (1939). *Business Cycles*. McGraw-Hill, New York.

Sevencan, A. (2023) Remittances, unemployment, growth, and development: A panel cointegration approach. *Applied Economics Letters*, 30(5), 663-668. https://doi.org/10.1080/13504851.2021.2009755

Shabbir, A., Kousar, S. & Alam, M. Z. (2021). Factors affecting level of unemployment in South Asia. *Journal of Economic and Administrative Sciences*, 37(1), 1-25. https://doi.org/10.1108/JEAS-03-2019-0040

Shimer, R. (1998). Why Is the U.S. unemployment rate so much lower? *NBER Macroeconomics Annual*, 13, 11–61. Available at: http://www.jstor.org/stable/4623732

Shimer, R. (2005). The cyclical behavior of equilibrium unemployment and vacancies. *The American Economic Review*, 95(1), 25-49. https://doi.org/10.1257/0002828053828572

Simionescu, M. (2020). Improving unemployment rate forecasts at regional level in Romania using Google Trends. *Technological Forecasting and Social Change*, 155, 120026. https://doi.org/10.1016/j.techfore.2020.120026

Slaughter, R. A. (1990). The foresight principle. *Futures*, 22(8), 801-819. https://doi.org/10.1016/0016-3287(90)90017-C

Slaughter, R. A. (1998). Future studies as an intellectual and applied discipline. *American Behavioral Scientist*, 42(3), 372-385. https://doi.org/10.1177/0002764298042003008

Smit, O. O. (2018). *Unemployment rate forecasting using Google Trends* [Bachelor Thesis, Erasmus University Rotterdam]. Retrieved from: https://thesis.eur.nl/pub/43899/Smit.pdf

Smith, J. A. (2011). BRIC becomes BRICS: Changes on the geopolitical chessboard. *Foreign Policy Journal*, 21, 2011. Available at: https://www.foreignpolicyjournal.com/2011/01/21/bric-becomes-brics-changes-on-the-geopolitical-chessboard/

Spence, M. (1973). Job market signalling. *The Quarterly Journal of Economics*, 87(3), 355–374. https://doi.org/10.2307/1882010

Spers, R. G., Wright, J. T. C. & Amedomar, A de A. (2013). Scenarios for the milk production chain in Brazil in 2020. *Revista de Administração*, 48(2), 254-267. https://doi.org/10.5700/rausp1086

Stock, J. H. & Watson, M. W. (1999). Forecasting inflation. *Journal of Monetary Economics*, 44(2), 293-335. https://doi.org/10.1016/S0304-3932(99)00027-6

Stock, J. H. & Watson, M. W. (2012). *Introduction to Econometrics*, Global Ed., Pearson Education, Boston, MA

Stockhammer, E. & Sturn, S. (2012). The impact of monetary policy on unemployment hysteresis. *Applied Economics*, 44(21), 2743-2756. https://doi.org/10.1080/00036846.2011.566199

Su, C-W., Dai, K., Ullah, S. & Andlib, Z. (2021). COVID-19 pandemic and unemployment dynamics in European economies, *Economic Research-Ekonomska Istraživanja*, https://doi.org/10.1080/1331677X.2021.1912627

Su, Z. (2014). Chinese online unemployment-related searches and macroeconomic indicators. *Frontiers of Economics in China*, 9(4), 573-605. https://doi.org/10.3868/s060-003-014-0027-3

Suchodolski, S. G. & Demeulemeester, J. M. (2018). The BRICS coming of age and the New Development Bank. *Global Policy*, 9(4), 578-585. https://doi.org/10.1111/1758-5899.12600

Syed, A. & Tripathi, R. (2020). Nonperforming loans in BRICS nations: Determinants and macroeconomic impact. *Indian Journal of Finance*, 13(2), 22-35. https://doi.org/10.17010/ijf/2019/v13i2/141684

Tahir, M. & Burki, U. (2023). Entrepreneurship and economic growth: Evidence from the emerging BRICS economies. *Journal of Open Innovation: Technology, Market, and Complexity*, 9(2), 100088. https://doi.org/10.1016/j.joitmc.2023.100088

Tenzin, U. (2019). The nexus among economic growth, inflation, and unemployment in Bhutan. *South Asia Economic Journal*, 20(1), 94-105. https://doi.org/10.1177/1391561418822204

Thomas, J. P. & Worrall, T. (2007). Unemployment insurance under moral hazard and limited commitment: Public versus private provision. *Journal of Public Economic Theory*, 9(1), 151–181. https://doi.org/10.1111/j.1467-9779.2007.00302.x

Thurik, A. R., Carree, M. A., Stel, A. & Audretsch, D. B. (2008). Does self-employment reduce unemployment? *Journal of Business Venturing*, 23(6), 673-686. https://doi.org/10.1016/j.jbusvent.2008.01.007

Todaro, M. P. (1969). A model for labor migration and urban unemployment in less developed countries. *American Economic Review*, 59(1), 138-148.

Tomić, I. (2018). What drives youth unemployment in Europe? Economic vs non-economic determinants. *International Labour Review*, 157(3), 379-408. https://doi.org/10.1111/ilr.12113

Tsai, C. F., Lin, W. C., Hu, Y. H. & Yao, G. T. (2019). Under-sampling class imbalanced datasets by combining clustering analysis and instance selection. *Information Sciences*, 477, 47-54. https://doi.org/10.1016/j.ins.2018.10.029

Umar, M. & Sun, G. (2016). Determinants of different types of bank liquidity: evidence from BRICS countries. *China Finance Review International*, 6(4), 380-403. https://doi.org/10.1108/CFRI-07-2015-0113

Uzay N. (2005). *Productivity and Growth*, Ankara: Nobel Publication.

van der Heijden, K. (2000). Scenarios and forecasting. *Technological Forecasting and Social Change*, 65(1), 31–36. https://doi.org/10.1016/s0040-1625(99)00121-3

van Eck, N. J.; Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523–538.

van Eck, N. J.; Waltman, L. (2014). CitNetExplorer: A new software tool for analyzing and visualizing citation networks. *Journal of Informetrics*, 8(4), 802–823.

Van Roy, V., Vértesy, D. & Vivarelli, M. (2018). Technology and employment: Mass unemployment or job creation? Empirical evidence from European patenting firms. *Research Policy*, 47(9), 1762-1776. https://doi.org/10.1016/j.respol.2018.06.008

Vasilyeva, A. V. (2017). The forecast of labour migration, reproduction of the population and economic development of Russia. *Economy of Regions*, 13(3), 812-826. DOI: 10.17059/2017-3-14

Vasilyeva, A. V. & Tarasyev, A. A. (2014). Future development of migration processes and labour market in the regions of Russia. *Economy of Regions*, 40(4), 283-298. https://doi.org/10.17059/2014-4-25

Venables, W. N. & Ripley, B. D. (2002) *Modern Applied Statistics with S*. Fourth Edition. Springer-Verlag.

Verma, S. & Gustafsson, A. (2020). Investigating the emerging COVID-19 research trends in the field of business and management: A bibliometric analysis approach. *Journal of Business Research*, 118, 253-261. https://doi.org/10.1016/j.jbusres.2020.06.057

Vicente, M. R., López-Menéndez, A. J. & Pérez, R. (2015). Forecasting unemployment with internet search data: Does it help to improve predictions when job destruction is skyrocketing? *Technological Forecasting and Social Change*, 92, 132-139. https://doi.org/10.1016/j.techfore.2014.12.005

Victor, V., Karakunnel, J. J., Loganathan, S. & Meyer, D. F. (2021). From a recession to the COVID-19 pandemic: Inflation-unemployment comparison between the UK and India. *Economies*, 9(2). 10.3390/economies9020073

Wanberg, C. R. (2012). The individual experience of unemployment. *Annual Review of Psychology*, 63, 369-336. https://doi.org/10.1146/annurev-psych-120710-100500

Wang, J., Reinhilde, V. & Stephan, P. (2017). Bias against novelty in science: A cautionary tale for users of bibliometric indicator. *Research Policy*, 46, 1416-1436. http://dx.doi.org/10.1016/j.respol.2017.06.006

Wang, Q. & Zhang, F. (2020). Does increasing investment in research and development promote economic growth decoupling from carbon emission growth? An empirical analysis of BRICS countries. *Journal of Cleaner Production*, 252, 119853. https://doi.org/10.1016/j.jclepro.2019.119853

Wang, X., Hyndman, R. J., Li, F. & Kang, Y. (2023). Forecast combinations: An over 50-year review. *International Journal of Forecasting*, 39(4), 1518-1547. https://doi.org/10.1016/j.ijforecast.2022.11.005

Warsame, A. A., Ali, A. O., Hassan, A. Y. & Mohamed, M. O. (2022). Macroeconomic determinants of unemployment in Somalia: The case of Okun's law and the Phillips Curve. *Asian Economic and Financial Review*, 12(11), 938-949. https://doi.org/10.55493/5002.v12i11.4636

West, P. (1987). The implementation of a youth unemployment program. Australian *Journal of Public Administration*, 46(1), p. 66–76. https://doi.org/10.1111/j.1467-8500.1987.tb01411.x

Widrow, B., Rumelhart, D. E. & Lehr, M. A. (1994). Neural networks: applications in industry, business, and science. *Communications of the ACM*, 37(3), 93-106. Available at: https://link.gale.com/apps/doc/A15061353/AONE?u=anon~56a7dc23&sid=googleScholar&xid=bdd219c6

Winkelmann, L. & Winkelmann, R. (1998). Why are the unemployed so unhappy? Evidence from panel data. *Economica*, 65(257), 1–15. Available at: http://www.jstor.org/stable/2555127

Winters, P. R. (1960). Forecasting sales by exponentially weighted moving averages. *Management Science*, 6(3), 231-362. https://doi.org/10.1287/mnsc.6.3.324

Wong, J. M. W., Chan, A. P. C. & Chiang, Y. H. (2005) Time series forecasts of the construction labour market in Hong Kong: The Box-Jenkins approach. *Construction Management and Economics*, 23(9), 979-991. https://doi.org/10.1080/01446190500204911

World Bank (2023) About us. Retrieved on April 18, 2023, from https://data.worldbank.org/about

World Bank. (2023). China Overview. Retrieved on August 24, 2023, from https://www.worldbank.org/en/country/china/overview

Wright, J. T. C., & Spers, R. G. (2006). O país no futuro: aspectos metodológicos e cenários. *Estudos Avançados*, 20, 13-28. https://doi.org/10.1590/S0103-40142006000100003

Wu, D. C., Cao, Z., Wen, L. & Song, H. (2021). Scenario forecasting for global tourism. *Journal of Hospitality & Tourism Research*, 45(1), 28-51. https://doi.org/10.1177/1096348020919990

Xu, W. (2022). Social insurance and entrepreneurship: The effect of unemployment benefits on new-business formation. *Strategic Entrepreneurship Journal*, https://doi.org/10.1002/sej.1419

Xu, W., Li, Z., Cheng, C. & Zheng, T. (2013). Data mining for unemployment rate prediction using search engine query data. *Service Oriented Computing and Applications*, 7(1), 33–42. https://doi.org/10.1007/s11761-012-0122-2

Yao, S., Li, T. & Li, Y. (2023). Promoting sustainable fossil fuels resources in BRICS countries: Evaluating green policies and driving renewable energy development. *Resources Policy*, 85, 103990. https://doi.org/10.1016/j.resourpol.2023.103990

Yoshida, N. D, Wright, J. T. C. & Spers, R. G. (2013). A prospecção do futuro como suporte à busca de informações para a decisão empresarial. *Revista Ibero Americana de Estratégia*, 12(1), 208-235.

Zanin, L. (2014). On Okun's law in OECD countries: An analysis by age cohorts. *Economic Letters*, 125(2), 243-248. https://doi.org/10.1016/j.econlet.2014.08.030

Zanin, L. (2018). The pyramid of Okun's coefficient for Italy. *Empirica* 45, 17–28. https://doi.org/10.1007/s10663-016-9343-5

Zhang, G., Patuwo, B. E. & Hu, M. Y. (1998). Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting*, 14(1), 35-62. https://doi.org/10.1016/S0169-2070(97)00044-7

Zhartay, Z.; Khussainova, Z. & Yessengeldin, B. (2020). Development of the youth entrepreneurship: Example of Kazakhstan. *Entrepreneurship and Sustainability Issues*, 8(1), 1190-1208. https://doi.org/10.9770/jesi.2020.8.1(80)

Zupic, I. & Cater, T. (2015). Bibliometric methods in management and organization. *Organizational Research Methods*, 18(3), 429–472.