LUISS

Department of Business and Management

Degree Program in Data Science and Management

Chair of Data Visualization

Urban Growth and Gendered Workforce: Employment Evolution in Developing and Developed Economies

Prof. Blerina Sinaimeri

SUPERVISOR

Prof. Alessio Martino

CO-SUPERVISOR

Costanza Placanica

759641

CANDIDATE

Academic Year 2023/2024

Acknowledgements

I would like to express my deepest gratitude to Professor Blearina Sinaimeri for their invaluable guidance, insightful feedback, and unwavering support throughout the course of this research. Their expertise and encouragement have been instrumental in shaping the direction of this thesis, and I am truly grateful for the opportunity to have worked under their supervision.

I am also incredibly grateful to my family, friends, and colleagues who have supported me throughout this journey, offering encouragement and motivation during the most challenging moments of the research process. Their belief in my work has been a constant source of inspiration.

Abstract

Urbanization is one of the defining global trends of the 21st century, transforming economies, societies, and labor markets worldwide. This thesis investigates the relationship between urbanization and employment, with a particular focus on women's employment dynamics, by comparing three distinct economic regions: China, India, and the European Union. While urbanization is often associated with economic development and job creation, its impact on women's employment dynamics remains complex and varies between developing and developed economies.

This research leverages data from the World Bank to examine the evolution of urban growth, employment structures, and sectoral shifts in labor markets. The study first explores the patterns of urbanization and population growth, highlighting the rapid urban transformation in China and India compared to the more stabilized urbanization in the European Union. It then examines the employment sector distribution, gender disparities in labor force participation, and the prevalence of precarious employment, including self-employment, family work, and vulnerable employment.

To assess future trends, an ARIMA-based forecasting model has been applied to predict the evolution of female employment in agriculture, industry, and services over the next decade. The projections indicate a continued decline in agricultural employment for women in all three regions, with a concurrent rise in the service sector employment. However, the forecasts also highlight uncertainties due to the absence of external influencing factors in the model, such as policy changes, economic crises, or technological advancements.

This thesis contributes to the broader discourse on urbanization and gendered labor market transformations, providing insights for policymakers seeking to promote inclusive economic growth. The study underscores the need for gender-sensitive urban policies to ensure that women can fully benefit from the opportunities created by urban expansion, particularly in developing economies where gender disparities in employment remain substantial.

Table of Contents

Table of Figures

CHAPTER 1. Introduction

"*Urbanization is not about simply increasing the number of urban residents or expanding the area of cities. More importantly, it's about a complete change from rural to urban style in terms of industry structure, employment, living environment and social security.*"

Li Keqiang

Human populations have continuously expanded throughout history. As the number of people kept growing, small groups came together to form groups and small sedentary communities, which slowly grew into what we now call cities. Those early civilizations were largely rural, but with the advent of agriculture, the sedentary communities evolved into urban settlements: this growth has historically corresponded with shifts in labor organization, as societies transitioned from agrarian economies to more complex, industrialized and service-driven labor markets.

The process of urbanization, defined as the movement of populations from rural to urban areas, leading to their expansion, has been one of the defining forces in shaping modern economies. Although cities have existed for centuries, it was not until the onset of the Industrial Revolution that urbanization in the Western World accelerated at an unprecedented pace: before then, the vast majority of the population was living in a rural context, and was sustained by agriculture, while others concentrated in small towns where they lived from trading and manufacturing on a small scale. With the start of the Industrial Revolution in the late 18th century, an unprecedented growth in urban population took place, both through important migration from the countryside to the city and through demographic expansion, forever changing the course of history.

The 20th century has been regarded as a period of intense transformation regarding the general world's urbanization: indeed, in 1900 only 15% of the world's population lived in cities and it was mostly in the Western World, but the pace of urban growth and urban population growth accelerated quickly after 1950, and in 2007 more than 50% of the world lived in urban areas.

There are various contributing factors that influence urban growth and expansion: it can happen due to natural causes such as urban demographic expansion, it can be planned and be a result of a government action, or it can be a result of the migrations from rural to urban areas. Indeed, migration of people looking to improve their life conditions because of declining agricultural opportunities, lack of access to quality healthcare and education or natural disasters is one of the major forces that drive urban growth.

As cities grew and became hubs for manufacturing, commerce, education, and professional employment, they attracted people in search of economic opportunities and improved living conditions, leading to nearly 56% of the world's population (around 4.4 billion people) living in urban areas, projected to rise to 70% by 2050. As of today, cities generate more than 80% of the global GDP. *Figure 1.1, 1.2* and *1.3,* created using Tableau, provide a visual overview of urbanization trends from 1960 to 2023, highlighting the rapid pace of urban growth, particularly in developing countries. In the charts, it is clearly shown how in 1960 higher levels of urbanization were concentrated primarily in Western countries; however, by 1992 and 2023, the rapid expansion of urban areas becomes evident, particularly across Asia and Africa. This transformation reflects the accelerating pace of development in previously less urbanized economies, driven by infrastructural expansion and demographic shifts.

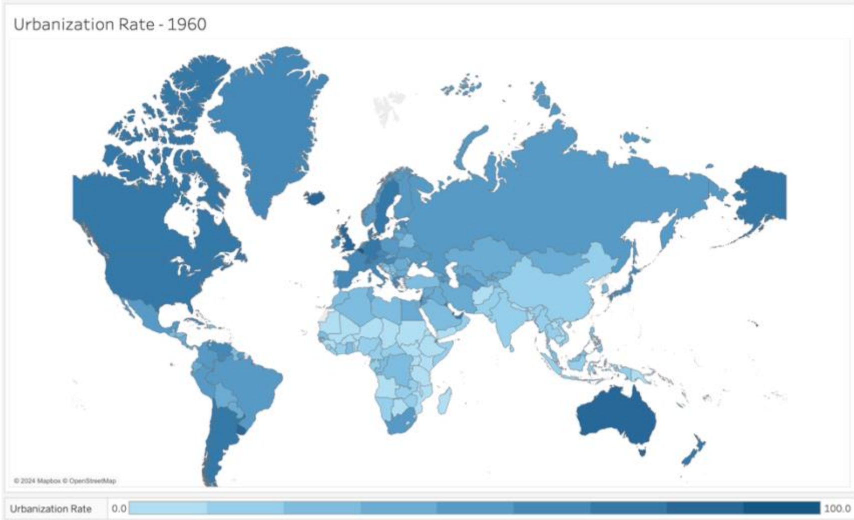

Figure 1.1 - Urbanization in 1960

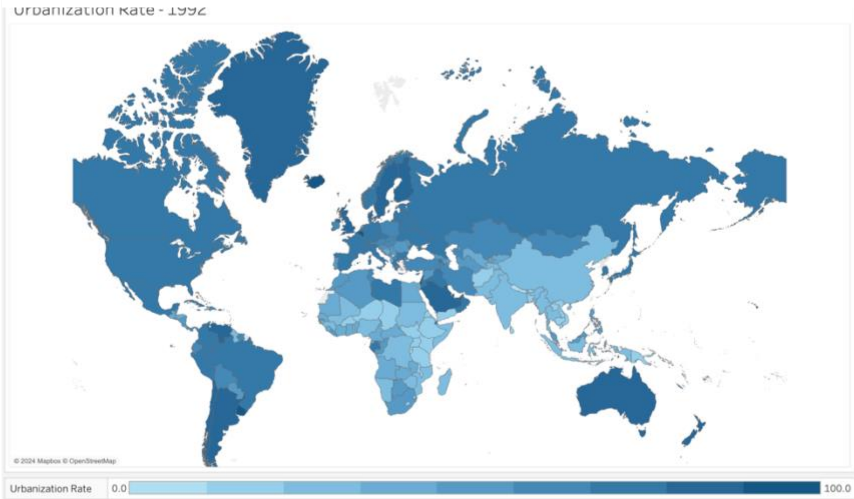

Figure 1.2 - Urbanization in 1992

Figure 1.3 - Urbanization in 2023

Nonetheless, despite urbanization often being associated with diversification, economic growth, higher productivity and increased opportunities, the scale and speed of urbanization also pose complex challenges, particularly in rapidly growing regions. Indeed, the so-called *Megacities*[1] are becoming increasingly common, and urban sprawl, a trend of over expansion of cities at a faster rate than their population, is at risk of occupying important land used for agriculture as well as straining transportation infrastructure. Moreover, urbanization frequently contributes to inequality, housing shortages (with slums formation as a consequence) and environmental degradation, as well as employment instability. As urban populations continue to grow, many are concerned with sustainable development and increasing levels of pollution, as the existing infrastructure and city planning practices are still not environmentally friendly.

However, in general, urbanization has also provided greater economic and social opportunities for women, although these benefits are unequally distributed across different economic contexts. Urban areas provide better access to education, healthcare and employment, but in many development countries women continue to face significant barriers to workforce participation and limited access to well-paying jobs, and women too often need to rely on low-wage, informal or unpaid family work.

The objective of this thesis is to explore the evolution and the relationship between urbanization and employment, specifically focusing on gender disparities and on how urbanization has shaped and reformed women's role in the labor market, through the comparison of three distinct economic landscapes: China, India and the European Union.

These three regions offer a compelling study as they represent very complex and different stages of economic development and urban growth, providing interesting comparisons between developing and developed economies:

China has had an unprecedented economic growth, making it the second-largest economy in the world, with a GDP of $18.76 trillion in 2024 (the first being the United Stated of America, with a GDP of $29.17 trillion).[2] When China started its urbanization process in 1949, the year when the People's Republic of China was founded, less than the 10% of the population in China was urban (Cook et al., 2017); as of today, in 2024, China has an urbanization rate of 67%, and is expected to reach approximately 80% in the next ten years.

India stands as one of the five largest economies in 2024, with a GDP of $3.90 trillion. Nevertheless, India is experiencing a more gradual urbanization process, with only 35% of its population living in urban areas as of 2023, due to the still large rural workforce, making agriculture a major employer.

[1] Cities with more than 10 million people

[2] Data gathered at the following link: Data Commons

The *European Union,* a highly developed and urbanized region, is a political and economic union of countries which are located primarily in Europe. Its GDP has been estimated to be $19.40 trillion in 2024, and its largest economies, Germany, France and Italy, figure among the most important economies in the world. The Member States' urbanization process, started in the 18th century, is advanced and is maintaining a stable trend, with over 75% of its population living in cities.
Despite having 27 Member States, each with different economies, politics, and history, for the purpose of this thesis, the European Union will be treated as a country in order to facilitate the analyses and provide clearer results.

In this thesis, the intricate relationship between urbanization, employment and gender disparities will be examined: starting with an overview of the evolution of urbanization in China, India, and European Union, it will proceed with highlighting the changes in employment for both male and female workers. It will then provide the results of the interactions and impacts of urbanization on employment with regards to women, and, finally, will present predictions on the engagement of women in the three main employment sectors (agriculture, industry, and services) in the next 10 years. The research employs a quantitative analytical framework, utilizing data-driven methods such as Exploratory Data Analysis, data visualization techniques and forecasting models, to study data on urbanization and employment indicators in the three countries.

The findings of this thesis will contribute to a broader understanding of the economic and social consequences of urbanization, with implications for policymaking, labor market reforms and gender-inclusive economic strategies. This research will highlight the diverse pathways through which urbanization affects gender disparities and women's employment dynamics, enhancing awareness and offering insights into the challenges and opportunities of the different economies.
Because urban populations are continuously expanding, governments must adopt integrated approaches to ensure that it leads to an inclusive growth, with equitable employment opportunities, and sustainable development.

The following chapter will provide a comprehensive review of the existing literature, examining academic research which will contextualize the study by outlining the key theories, previous findings and possible gaps in the literature, helping to establish the foundation for the research framework. It will also present the research questions that guide the study, framing the key inquiries that will be explored throughout the analysis.

By synthesizing insights from economic development, labor economics and gender studies, the review will position this thesis within the broader academic discourse and highlight the unique contributions of this research.

CHAPTER 2. Review of the Literature

This literature review seeks to provide a global overview of the research that has been conducted on the intersection between urbanization and women's employment. It especially aims to explore the broader dynamics of urbanization and its social implications, particularly regarding women's economic participation. The analysis will firstly address urbanization as a global phenomenon, highlighting the historical trends, the driving forces and the consequences on society as a whole; subsequently, the review will narrow its focus to examine how the gender dynamics have been shaped by the rising urbanization, particularly in the labour market, and, finally, it will delve deeply into the specific context of the three regions considered in the analysis: China, India and the European Union. This last section will offer insights into how the interplay between urban growth and women's employment manifests in these regions.

The works cited in this chapter have been selected for their relevance in the fields of sociology, gender studies and urban development. Many keywords were used for the search in academic databases and journal archives, while several works were included due to them being prominently referenced in other influential studies. This method of selection ensured both the reliability of the sources and the diversification of perspectives and methodologies.
It is crucial to note that the specific topic of this thesis has not been extensively studied as a standalone subject, despite there being a substantial body of literature on urbanization's social and economic impacts and a good amount on gender and employment in urban settings.
Indeed, this literature review identifies the current state of knowledge while also highlighting a gap in the current research, where further investigation is needed.

2.1 Urbanization

Characterized by populations migrating from rural to urban areas, urbanization is one of the truly defining phenomena of our era, and with it, the transformation of economic and social structures.
In recent decades, the world has witnessed an unprecedented increase in urban populations.
Extensive research has been done on this phenomenon and how it has shaped the modern society and its development. In 1965 Kingsley Davis (1908-1997), a famous American sociologist and demographer, published an article in the journal *Scientific America* titled *The Urbanization of the*

Human Population, which has been regarded as a foundational text in urban studies. In the article, Davis gives a comprehensive analysis of urbanization[3], linking social and economic changes to the rise of urban centres. In particular, he highlights how urbanization is not simply a demographic phenomenon, but most importantly an indicator of social and economic transformations; he draws a connection between urbanization and industrialization: "with industrialization, however, the transformation was striking. By 1801 nearly a tenth of the people of England and Wales were living in cities of 100,000 or larger. This proportion doubled in 40 years and doubled again in another 60 years. By 1900 Britain was an urbanized society. In general, the later each country became industrialized, the faster was its urbanization[4]" (Davis, 1965). One of the reasons why urbanization is closely linked with industrialization is that the rise of urban centres contributes to the concentration of economic activities in cities, which create opportunities for employment and attract rural populations, resulting in an acceleration of urbanization. This is also referenced in a study by Masika et al., *Urbanisation and Urban Poverty: A Gender Analysis* where urbanisation is linked to increased economic opportunities, better access to education and improved living standards (Masika, 1997). In general, it has been shown that urban areas serve as hubs of industrial and service-sector activity, attracting investments and fostering innovation (Tacoli, 2012). The agglomeration of economic activities in cities creates economies of scale, leading to increased productivity and income levels (Buchholz, 2023). Another important reason is the advancements in technology and infrastructure, as he points out that before the industrial revolution, urbanization was limited due to transportation constraints, which made moving people or goods from one city to another inconvenient and laborious. Moreover, Davis integrates urbanization into the framework of the demographic transition theory. He argues that urbanization follows a specific and predictable pattern as societies progress through the stages of demographic transition: the first stage (pre-industrial societies) where urbanization is minimal and the population is engaged in agriculture, the industrialization phase in which rapid urbanization occurs due to rural-to-urban migration. Finally, post-industrial societies, where urbanization becomes stable and the growth of the population slows down, as economies become mature.

---

[3] Davis defines urbanization as the proportion of the total population concentrated in urban settings. He explains that a common mistake is to consider urbanization as the growth of cities, which is incorrect as urban areas can still grow while their proportion of the total population remains constant.

[4] The indices used in the analysis reported in this article were the proportion of people living in places of 100,000 or more, and the proportion of population classed as "urban" in the official statistics of each country.

2.1.1 The Challenges of Urbanization

There are serious disparities in how urbanization rates have evolved between developed and developing countries: as developed countries have had a more gradual urbanization trend, accompanied by a relative economic growth, their cities had time to adjust to the increasing population. However, developing countries have experienced and are still experiencing rapid urbanization without the correlated industrialization or economic progress, which leads to urban issues such as slum formation, unemployment, and inadequate infrastructure. This particular topic has been addressed by a number of studies: in *The Transition to a Predominantly Urban World and its Underpinnings*, Satterthwaite reports that in low-income and middle-income nations "around a billion urban dwellers – a sixth of the planet's population – are homeless or live in crowded tenements, boarding houses or houses or shacks in informal/squatter settlements. […] Government investment in the basic infrastructure that all businesses and households need, […] falls far behind the growth in population and enterprises. So too does the provision of basic services." (Satterthwaite, 2007)
Tacoli, in her work, noted that urbanization in these contexts exacerbates existing vulnerabilities, especially for marginalized populations living in informal settlements.
She mentions that urban poverty[5], characterized by income insecurity and with limited access to basic services, also has a distinctly gendered dimension, as women disproportionately bear the burden of unpaid care work. In many urban contexts, women are compelled to balance paid employment with domestic responsibilities, resulting in *time poverty*[6] (Tacoli, 2012). Tacoli also highlights that with the intersection of gender and poverty in urban contexts, women in informal settlements face heightened risks of violence and exploitation. Again, Masika et al. argued that women-headed households are particularly vulnerable to urban poverty, as they often lack access to resources and support networks. The feminization of poverty is a critical issue in urban settings. This vulnerability is compounded by the lack of affordable childcare and inadequate social protections, which limit women's ability to engage in paid work.

[5] The term "urban poverty" denotes a condition in which individuals or families in cities do not have access to food, shelter, or healthcare due to a lack of financial resources. It is a complex issue generally present in megacities and it is known to affect millions of people.
[6] Time poverty refers to the lack of sufficient time to fulfill responsibilities, pursue hobbies or engage in activities that contribute to the person's well-being, because of heavy workloads, long commutes, or unpaid domestic work. Research found that it can lead to burnout or self-neglect and to worsening of mental health (Travers, 2024).

2.2 Urbanization and Gender Relations

At present, the percentage of women participating in paid employment is at its highest in history. Urbanization has deeply reshaped the relationship between men and women, especially in terms of economic roles, social dynamics, and labour distribution, by questioning the traditional gender roles and creating new opportunities and barriers for women, thus enhancing women's position in society. As cities expanded and populations shifted from rural to urban areas, women found greater access to formal employment[7] opportunities, which contributed to their empowerment; nevertheless, urbanization also exacerbated gender inequalities both within the labour market and in household structures (de Bruin & Liu, 2019).
A crucial aspect has been the increased access to higher education and professional development opportunities, as in urban areas women tend to have more exposure to higher-level educational institutions, training programs, as well as skill-development initiatives, all of which enhance their employability and career prospects.

Moreover, as mentioned by Buchholz in his work *Does Urbanization Increase Inequality?* in high-income countries researchers have found a strong positive relationship between the city's size and density and its average income level. This concept is known as "urban wage premium".
The urban wage premium is generally attributed to higher productivity in more urbanized locations, which is likely due to the possibility of workers to match with better jobs and learn more valuable skills and experience, resulting in higher wages (Buchholz, 2023).
The correlation between urban growth and higher wages was also previously theorized by Davis, who mentions that as productivity grows, the services and manufacturing industries can increase the manpower by paying higher wages (Davis, 1965).
However, despite urbanization providing more economic opportunities for women, gender disparities still persist in career advancements and wage levels. Although research linking urbanization to inequality is relatively limited, even with the increasing interest in geographies of economic inequality and how they vary across social categories like gender (Buchholz, 2023), the existing research on the topic still provides compelling points of discussion.
For instance, Buchholz recognizes that although women may face less inequality in more urbanized areas[8], which is possibly because they tend to possess higher-level cognitive and social skills than

[7] Formal employment is a type of employment created through agreements in the form of contracts between an incorporated company and an individual employee. It can be found in the economic sector, in banking, in the government and in industries such as the manufacturing, mechanic, and services, and differs from informal employment as the latter is not regulated and does not have the same benefits.
[8] Another researcher, Bacolod, finds that the female-male wage gap tends to be lower in larger cities in the U.S.

men, a gap still persists as they tend to get lower returns on those skills. He explains that there might be several well-established factors influencing this gap: education disparities, psychological factors, and occupational differences, as well as the ethnicity. Indeed, Buchholz finds that in the U.S. as cities get denser and denser in population, wage increases for white workers and men, whereas wage changes tend to be significantly lower for women, black workers, Asian Americans and other minorities (Buchholz, 2023).

There is also another aspect that increases the disparities between men and women: unpaid care work. While urbanization led to an increase in women's participation in paid employment, the persistence of unpaid care work remains a significant barrier to gender equality: the need to balance work and domestic responsibilities limits women's career progression and reinforces gender inequalities in the labor market.

According to the United Nations Development Programme, unpaid care work relates to "the provision of direct or indirect care without remuneration, carried out within the household. It includes, but is not limited to caring for children, elderly and sick individuals, washing, cooking, shopping, cleaning and helping other families with their chores" (*Unpaid Care Work – A Gateway for Gender Equality*). Indeed, Tacoli reports that unpaid care work is an essential issue which makes urban poverty a distinctive gendered dimension: women, despite making a crucial contribution to the prosperity of cities both through their paid and their unpaid labour, are still at a disadvantage, especially in urban areas of low- and middle-income countries. They are still compounded by the need to earn an income, sometimes with long-hours jobs which add to the time poverty when concurrent with the unpaid to be done while off of work (Tacoli, 2012). As mentioned in de Bruin and Liu work, as women go home from their paid job, they start their "second shift".

Unpaid work stems from the residues of the pre-capitalist world, when populations organized their daily life around the categories of age and gender: men worked outside in the fields or in farms, while women worked in the house and took care of the children, the elderly and the sick.

Industrialization and urbanization radically changed society, by having men and women participating in wage work and profoundly changing the organization and the meaning of daily life (Thorbek, 1988). While this should bring to an equally divided total burden of work (paid work plus unpaid work) between men and women, unpaid work is still mostly done by women, many times at the expense of sleep and leisure, which has detrimental effects on female mental health.

On a more positive note, studies have shown how better education and higher emancipation of women can drastically facilitate the lessening of the "second shift" and a better balance of the unpaid work within the family, contributing to better time-use and higher possibilities of better employment opportunities for women (de Bruin & Liu, 2019).

Another notable effect of urbanization is the increasing presence of women as employers and entrepreneurs: the possibility to engage in networking opportunities, benefiting from access to financial institutions and the provision of a supportive infrastructure made it so that women have more opportunities to establish and expand businesses, as well as advancing in their careers. There are still challenges to be faced for women entrepreneurs and employers, as studies show that gender bias is still rooted in organizations. Acker points out that organizations' structure is still not gender neutral and that most firms are almost always dominated by men; the "masculine ethic", which elevates the traits assumed to belong to men[9], is still the predominant image for managers (Acker, 1990).

Nonetheless, in recent years women found a better support ecosystem (for instance, the foundation of women's associations in varying areas), and urbanization has allowed women to leverage digital platforms and technology to scale their businesses and access new markets, thus creating job opportunities for other women.

2.3 Country Comparison on Urban Growth and Its Impacts on Women

All of what has been discussed in this chapter until now, has manifested in distinct ways across the regions in analysis: China, European Union, and India.

China has experienced an exceptionally rapid urbanization in recent years. Starting from a predominantly agrarian economy, industrialization, strong economic growth, and the relaxation of rural-urban migration regulations helped transitioning to an urban-industrial society. Müller et al. reports that in 1988, in order to foster urban development and growth "since development was prohibited on non-state land, the central government introduced a system of leasing long-term rights for the use of state-owned land, allowing most revenues to be retained by municipalities. The result was a massive redevelopment of the inner cities, and residential as well as industrial development in peri-urban areas." Government policies were focusing on industrial expansion, and cities rapidly became hubs of economic growth. The government approach to urbanization is characterised by top-down planning and large-scale infrastructure projects, which are furthering urban expansion and are

[9] Those traits might be: "a tough-minded approach to problems, analytic abilities to abstract and plan, the capacity to set aside personal, emotional considerations in the interests of task accomplishment, a cognitive superiority in problem-solving and decision making" (Acker, 1990)

increasingly leading to the rise of the so-called *mega-cities* (the two most important being Shanghai and Beijing).

Indeed, urbanisation in China was so fast, that in the late 1990s one of the government goals was to control the growth of the population in big cities, while promoting growth in smaller towns, and balancing the rural-urban disparity (Müller et al., 2019).

The Chinese government has frequently discussed on urbanisation, which is seen as a strong engine for the economy: in 2012, the Third Plenary Session of the 18th National CPC Central Committee[10] stated that a "new urbanisation process" was going to be installed to foster economic growth and increase productivity and life quality (*Third Plenary Session of the 18th CPC Central Committee*, 2012).

Many studies reflected on China's unique road to an urbanized society: as Davis had observed, developed countries generally tend to experience "parallel urbanization", meaning achieving high industrial growth with a parallel increase in the degree of urbanization; developing countries, instead, more often tend to deal with over or excessive urbanization. However, research on China is not cohesive on whether the urban growth's trend represents the one of a developing country: on one hand, studies theorise that China's economic growth outpaced urbanization (Chang & Brada, 2006), on the other hand, many believe the economic growth rate was not fast enough to actually go hand in hand with urbanization. For instance, a study by Chen et al., showed that, compared to other countries China's urban growth rate was indeed much faster than the growth rate of the economic development, thus falling in line with the trend Davis had observed on developing countries.

India's urbanization follows a different trajectory compared to China. A report by Dobbs and Sankhe on McKinsey Quarterly states that in 1950 India was a much more urban nation than China: India's urban growth rate then kept growing, although not at the same speed as the Chinese one. From 1950 onwards, the two largest nations (in terms of population) basically urbanized at the same time and somehow at the same pace (Dobbs & Sankhe, 2010).

However, despite the often-seen comparisons between the two countries, some studies show that the consequences are much different: while China invested ahead to increase the productivity and well-being of its residents in urban areas, India remained behind.

An example being that several studies suggest that in India the population living in urban areas might be severely undercounted; indeed, Xuafei Ren in its *Urbanization, Migration and Families in China*

[10] The CPC (Chinese Communist Party) meets once a year to discuss relevant topics about policy issues. The Third Plenary Session of the 18th National CPC Central Committee was conducted on November 12, 2012. More than 300 reform measures were unanimously adopted, many of them would be particularly relevant in the following years. (*Third Plenary Session of the 18th CPC Central Committee*, 2012)

and India, reports that the criteria used to define a city "urban" dates back to 1961 and classifies as urban only those cities which have "a municipal status or reach a threshold of population size, density and ratio of male employment in non-agricultural sectors", thus making the actual number of people living in urban areas possibly incorrect (Ren, 2022).

Another source also refers to this issue: Onda et al. report that studies on Indian economy use districts as units of analysis because of the uncertain boundaries that define metropolitan areas (which would be the more correct choice), thus reaching misleading conclusions on the real numbers (Onda et al., 2019).

India seems to have experienced the over-urbanization feature of developing countries, with all the issues that it entails: slums (otherwise called "informal settlements"[11]), housing problems, lack of water and sanitation supply and inadequate infrastructure; it is said that in Mumbai almost 50% of the population lives in slums, as the city has reached its saturation level.

Saturation has also been reached on employment, leading to many suffering from urban poverty as cities are not able to generate employment for those migrating from rural areas, thus creating an involution of society instead of an evolution, which fosters social and economic inequalities and social conflicts (Jaysawal & Saha, 2014).

On the other side, urbanization in the Europe Union follows the trajectory of developed countries: growing together with economic growth and industrialization, it is now well established. Since the beginning of the first industrial revolution, Europe transformed from a largely rural society to a primarily urban one, with its cities playing a crucial role in innovation, productivity, and advancements. As European Union is a union of countries, many policies and approaches to urbanization have been decentralized, with significant input from local and regional governments; however, the EU has still implemented policies on urban renewal and sustainable growth, in order to address challenges such as aging infrastructure, climate change, pollution and social integration of migrant population, such as the Leipzig Charter on Sustainable European Cities (Müller et al., 2019).

2.3.1 Women and Employment: Focus on China, India and European Union

The European Union is often vocal on gender equality, and has promoted initiatives, policies, and directives to reduce disparities (one example being the Gender Equality Strategy, which has directives

[11] In India, slums are considered informal settlements as the residents do not have legal ownership of the land they occupy

on gender balance in corporate boards, on the fight of violence against women and on the prevention of domestic violence (*Gender Equality Strategy - European Commission*)).

However, many studies report that disparities are still present, especially at the national level. An article published on *EuroCities* in 2024 reports that women are still underrepresented in politics, and according to Annelies Coennens, Gender and Diversity Officer at CEMR[12], one reason for the underrepresentation is the pervasive violence women still face on daily basis, which "acts as a barrier and even a discouragement" (*Eurocities*, 2024).

Moreover, despite all the initiatives and policies, women are still behind in high positions in companies or in decision-making seats in all the countries in analysis. Women are still more likely than men to engage in unpaid care work or low-income jobs, which severely impacts women's independence. Also, despite the significant progress, women remain the less active participant in the labour market and engage in more part-time employment than men (Christiansen et al., 2017) which results in lower wages and in *gender pay gap*.

However, studies have still shown that in urban areas women gained more bargaining power; one example is the report by de Bruin and Liu, in which they analysed the habits of Chinese people and the time dedicated to unpaid care work (although it is difficult to exactly calculate it), and found that despite Chinese traditional gender norms and its patriarchal tradition still continue to dictate the time-use patterns, women in urban areas devoted less time to unpaid work, especially in younger generations (de Bruin & Liu, 2019) (Zhang, 2024).

In India the pattern seems to be following a similar, albeit slower route. In their work on women empowerment and urbanization, Dhamija et al. reports that girls and women spend a much higher amount of time compared to boys and men on domestic activities and unpaid care for family members, which discourages participation in paid employment. As a matter of fact, it is said to be unclear whether urban growth actually enhanced the position of women in India: poor infrastructure, sanitization and the pervasive violence women are subjected to, pose a threat to female emancipation and bargaining power, thus reducing the opportunities for employment in higher-paying jobs (Dhamija et al., 2023).

2.4 Conclusive Remarks on Existing Literature

Studies on urbanization touch a broad range of topics: impacts on the environment, social and technological development, infrastructure, and economics. However, there are several gaps existing

[12] Council of European Municipalities and Regions

in the research on how urbanization impacted and continues to impact gender disparities, especially when related to employment.

A key gap is the lack of focus on the *type* of employment women engage in, rather than general labour force participation. Indeed, most studies provide an overarching view of gender employment trends (examples being Zhang, 2024; Buchholz, 2023; de Bruin & Liu, 2019; Masika, 1997) but fail to analyse the sectorial differences: there is almost no research on the impacts that urbanization has had on women in agriculture, industries or services.

There are also few mentions (an example being Christiansen et al., 2017) of part-time employment and its significance on the emancipation of women, and there is even fewer research on this topic related to China or India.

Another notable gap is the absence of comprehensive comparative studies that directly juxtapose the experiences of women in different countries, especially when comparing developing and developed countries. While individual studies explore gendered urbanization in individual regions (and sometimes comparing two countries as in Dobbs & Sankhe, 2010for India and China, or as in Müller et al., 2019 for EU and China), studies that highlight the systemic differences and similarities between various economies and societies are few.

Moreover, data on several indicators might be incomplete or incorrect, due to the differences in definitions and due to the varying census strategies between countries.

The lack of comprehensive research of this kind creates significant gaps in understanding women's real economic and social. This knowledge gap contributes to a lack of awareness among policymakers and stakeholders, ultimately leading to the development of inadequate and ineffective policies.

To address and complement these shortcomings, this thesis proposes a study structured as an analysis of chosen key employment indicators that reflect the employment situation of women, with a specific focus on China, India, and European Union. The research aims to explore the changes during the last 50 years, as the three countries live very different stages of development. It will provide the general evolution of urbanization and of urban and rural growth and will then analyse the specific topic of employment and its intersection with urbanization.

Therefore, the thesis will answer to these primary research questions:

RQ1) How does the level of urbanization relate to employment and unemployment patterns across China, India and European Union and is there a difference between the developing and developed countries?

RQ2) How has urbanization influenced the distribution of employment across the major employment sectors and do the effects differentiate among China, India and European Union?

RQ3) What are the trends across the three countries on precarious employment and informal economy, and to what extent does urbanization relate to women's transition from informal to formal employment?

These research questions address various important topics related to women employment and how they relate to the increasing urban growth. By answering these questions, this thesis aims to provide insights into how urban growth shapes employment and unemployment patterns in developing and developed countries, identifying trends and disparities on labour market structures. The comparative approach of this thesis will contribute to a deeper understanding of the global labour landscape by highlighting the challenges and opportunities presented by urbanization in different economic contexts.

CHAPTER 3. Methodological Framework

Answering to the previously outlined research questions requires a well-structured methodological approach that ensures an accurate collection, analysis, and interpretation of the data.
This chapter presents the methods adopted in this study, which contributed to the analysis of the trends in urbanization and employment patterns: the research will focus on trends of urbanization and urban growth, employment distribution across sectors, types of employment, as well as the contrasts and similarities across China, India and European Union, thus providing a comparative perspective between developing and developed economies.
The framework employed in this study follows a quantitative longitudinal research design, utilizing secondary data from globally recognized sources to ensure the validity and reliability of the findings. The methodological approach follows a sequential process: it begins with collecting the data, which is then cleaned and manipulated to serve the purpose of the research. Therefore, an exploratory data analysis (EDA) is used to understand the trends and patterns, followed by a correlation analysis to assess whether relationships among the variables are present and whether there are significant connections. Finally, a predictive analysis will be conducted on women's participation in various employment sectors using the ARIMA predictive model, enabling forecasts of future trends based on historical data. This will provide insights into potential changes in employment patterns and help assess the long-term dynamics in these sectors.
A critical component of the research methodology involved the visualization of the data, which plays a crucial role in uncovering meaningful patterns and trends: for this reason, different types of charts and graphs have been selected to present the findings effectively, with each visualization chosen based on its suitability for illustrating specific relationships and trends within the data. Indeed, for each visualization shown in this thesis, a brief reasoning as to why it has been used to best represent the data in analysis will be provided, in order to explain the underlying data visualization concepts and how they have been applied in this work.
Furthermore, the research leverages a range of computational tools and libraries to facilitate data processing, analysis and visualizations.
The programming language Python has been utilized for data analysis due to its versatility and extensive ecosystem of data science libraries. Libraries such as Pandas, NumPy and Scikit-Learn have been employed for data manipulation and statistical analysis, as they allow for efficient handling of large datasets, while the Statsmodels library has been used for the forecasting. Additionally, tools

such as Matplotlib, Seaborn and Altair[13] have been used to create informative and visually appealing representations.
Moreover, another data visualization software, Tableau, has been used to create some of the visualizations that will be reported in this work. As the number is limited, it will be specifically mentioned when a visualization has been created using Tableau.
The combined use of these tools has enabled a comprehensive and multi-faceted approach to analysing and presenting the data, ensuring clarity and accuracy in communicating the complex urbanization and employment trends.

3.1 Research Design

This section presents the research design which has been implemented in this thesis. Choosing the most accurate design plays a vital role in correctly addressing the research questions and in presenting the results. The selected design offers a systematic and flexible approach to the exploration of the intersections between urbanization and employment in the countries in analysis.

3.1.1 Longitudinal Research

After carefully formulating the research questions, it was essential to select an appropriate research methodology that effectively aligned with the objectives of this thesis. Given that the research questions focus on identifying trends, patterns and the interactions between various employment and urbanization-related variables, a longitudinal approach seemed the most suitable.
Therefore, this study adopts a correlational longitudinal research design, which is particularly effective in analysing data trends over extended periods.
Longitudinal studies, which are types of correlational research, involve the observation of the same populations over a period which can range from days to years. It is suited for this research as it provides a deeper understanding on how the population and employment-indicator values evolve over time, contrarily to cross-sectional studies, which offer only a snapshot of data at a single point in time. The temporal dimension of longitudinal studies is essential for understanding employment shifts, sectoral changes and gender-based transitions in response to urbanization.
Moreover, the longitudinal approach falls under the broader category of correlational research, which helps to establish associations between variables, thus providing insights into whether and how

[13] Vega-Altair (https://altair-viz.github.io/index.html) is a declarative visualization library for Python which is easy to use and user friendly, built on the foundations of the Vega-Lite grammar. It creates clean and aesthetically pleasing visualizations, also allowing for interactivity in the charts. Examples of plots that can be created through Altair can be found at the following website: https://altair-viz.github.io/gallery/index.html.

urbanization influences employment patterns. Various statistical techniques have been used in this study to measure the correlation among the variables, such as Spearman and Pearson correlation.

3.1.2 Variable Selection and Hypotheses

Once the research approach was established, the next critical step involved identifying the necessary indicators that would effectively address the research questions, aligning with the study's objectives. Therefore, a review of the research questions was conducted to determine the essential indicators required for a comprehensive analysis:

RQ1) How is the level of urbanization related to employment and unemployment patterns across China, India and European Union and is there a difference between the developing and developed countries?

RQ2) How has urbanization influenced the distribution of employment across the major employment sectors and do the effects differentiate among China, India and European Union?

RQ3) What are the trends across the three countries on precarious employment and informal economy, and to what extent does urbanization relate to women's transition from informal to formal employment?

Each specific question demanded indicators that allowed for trend analysis and cross-country comparisons: since the primary focus was to assess changes over time, it was fundamental to include a time-based variable, which helped to track employment and urbanization trends over multiple years. This typically involves a *Year* variable, as this type of effects can most successfully be studied across years, rather than across shorter timespans.

Furthermore, to enable cross-country comparisons, a variable identifying the country or the region hat to be present in the dataset.

As urbanization was the central theme of this research, it was vital to include indicators that measured the urban growth rate or the urban population growth. Other trends related to population could be studied and compared to urban growth, such as rural growth rate and the population in larger cities[14].

All research questions examined the relationship between urbanization and employment patterns, therefore indicators reflecting general employment and unemployment rates such as the labour force participation had to be considered. Finally, the data should have also included indicators representing

[14] For example, as reported in Chapter 2, Davis considered as indicators for urbanization both the urban growth rate and the population living in cities of more than 100,00 people, both variables that might be considered for this analysis.

the employment sectors (agriculture, services, industries, and other types of employment), together with those indicating precarious employment (vulnerable employment, family workers).

The selection of these variables ensured that the research comprehensively addressed the study objectives by covering the essential dimensions of employment, while the inclusion of temporal and geographic variables enabled meaningful comparisons and trend analysis. Additional indicators were incorporated to enhance the depth of the study and provide more nuanced insights, as other indicators could have contributed valuable context and enriched the interpretation of the findings.

Once the variables needed for the study were generally selected, a series of hypotheses were formulated as a guide and to provide a structured approach to the analysis. All hypotheses were grounded in existing literature and theoretical frameworks and aligned with the research questions.

Hypotheses on Research Question 1:

H1.1: There is a significant negative correlation between urban growth rate and unemployment rate and a positive correlation between urban growth rate and employment rates across all countries in analysis.

H1.2: There is an increase in women employers as urbanization increases and the two indicators have a positive correlation across all countries.

Hypothesis on Research Question 2:

H2.1: As urban growth rate increases, a reduction in agricultural employment and a corresponding increase in industrial and services employment is shown across all countries but is higher in developed countries.

Hypothesis on Research Question 3:

H3.1: Vulnerable employment is much higher in developing countries, especially in India.

H3.2: There is a general decrease of precarious employment rates as urbanization increases, and it is more pronounced in developed countries.

The formulation of these hypotheses was based on prior research findings on the transformative role of urbanization in employment trends and focused both on the disparities between men and women employment, and on the similarities or differences across the countries in analysis.

3.2 Data Collection

This thesis employed secondary data analysis as data sampling technique, which involves the use of existing datasets that have been previously collected, processed, and published by reputable organizations, rather than gathering primary data through direct means such as surveys or experiments.

3.2.1 Data Sources

The data utilized for this thesis was sourced from the World Development Indicators database, which is available for public access through the World Data Bank data catalog.[15]

The WDI database is a widely recognized source of data on global development indicators, providing a broad range of social, economic, and environmental statistics; it is updated multiple times each year to include the most recent developments and revisions, therefore it is important to specify that the version used for this analysis was updated in July 2024 at the latest. Consequently, any changes or additions made to the dataset after this date will not be considered in the analysis.

Before delving into the data analysis procedures, it is essential to provide a brief overview of the dataset and its source to establish the credibility and relevance of the data used in this study.

The World Bank, established in 1944 during the Bretton Woods Conference, is a global financial institution that provides financial and technical assistance to developing countries to foster economic development and poverty reduction. It operates as a conglomerate of five institutions: the International Bank for Reconstruction and Development (IBRD), the International Development Association (IDA), the International Finance Corporation (IFC), the Multilateral Investment Guarantee Agency (MIGA) and the International Centre for Settlement of Investment Disputes.

It is currently headquartered in Washington D.C., in the United States and comprises of 189 member countries. It is represented by a Board of Governors, "who are the ultimate policymakers at the World Bank. Generally, the governors are member countries' ministers of finance or ministers of development. They meet once a year at the Annual Meetings of the Boards of Governors of the World Bank Group and the International Monetary Fund. […] The World Bank operates day-to-day under the leadership and direction of the president, management and senior staff, and the vice presidents in charge of Global Practices, Cross-Cutting Solutions Areas, regions, and functions"[16] (World Bank, 2024). The World Bank provides leveraged loans to low and middle-income countries to further the economic development; its mission revolves around fostering sustainable development through

---

[15] https://datacatalog.worldbank.org/search/dataset/0037712

[16] *Who We Are:* https://www.worldbank.org/ext/en/who-we-are

projects aims at improving human capital, environmental sustainability, infrastructure development, industrial growth and governance structures (World Bank, 2024).[17]

The data used for this research originates from the World Bank Open Data Portal, an open-access data bank that provides users with the possibility to download and visualize data, graphs and reports covering a wide array of global economic and social statistics. The portal offers several key databases, including World Development Indicators (WDI), International Debt Statistics (IDS), Millennium Development Indicators (MDI), Poverty, Education and Gender Data.

For the purpose of the analyses in this thesis, the data has been specifically extracted from the World Development Indicators (WDI), which contains a comprehensive set of indicators designed to measure social progress, quality of life, economic and social development, environmental sustainability, and government performance. The dataset spans from 1960 to 2023. (*World Development Indicators (World Bank)*, n.d.).

The WDI database was chosen as data source for this research due to its reliability and credibility and its diverse indicator selection, as well as for its open-access nature, which allows for reproducibility and verification of findings.

To address the research questions, a subset of relevant indicators was carefully selected from the WDI database, as well as only the three regions considered in the analysis.

The selected indicators have been categorized in two major cluster: population statistics and employment statistics.

3.2.2 ILO Estimate

It is important to highlight that the indicators used in this study are modelled estimates provided by the International Labour Organization (ILO). This signifies that the data might not always reflect direct national reporting, but rather values derived from the ILO's statistical models. The estimates are generated to address potential gaps in data, ensuring consistency and comparability across countries. Indeed, due to resource constraints and varying statistical capacities, some countries are unable to conduct regular and systematic surveys, resulting in significant data gaps. To overcome this issue, the ILO developed and tested a series of econometric models that generate estimates for those countries whose data might not have been consistent or available every year. According to the ILO, the modelled estimates series provide "a complete set of internationally comparable labour statistics, including both nationally reported observations and imputed data for countries with missing data. The

[17] *What We Do:* https://www.worldbank.org/en/what-we-do

imputations are produced through a series of econometric models maintained by the ILO" (*ILO Modelled Estimates (ILOEST Database)*, n.d.).

Therefore, modelled estimates ensure data coverage and reliability, thus allowing for cross-country comparisons and trend analysis; however, it is worthy of note that estimates, while based on robust methodologies, might carry inherent limitations as they rely on assumptions and predictive modelling rather than direct empirical observations. On the other hand, empirical observations can also be subjected to inconsistencies due to methodological differences, variation in data collection practices and biases across countries.

Recognizing these limitations is crucial for interpreting the findings of this study with an appropriate level of caution, ensuring that the conclusions are considered with an awareness of above-mentioned data constraints.

3.3 Data Analysis

In this section, the data analysis procedures undertaken in this study will be outlined. These comprise of data cleaning, which is performed to structure the data appropriately for the research, exploratory data analysis (EDA), to gain insights and identify patterns or trends and distribution within the data, and then a correlation analysis to quantify the relationships among the indicators. Finally, a forecasting analysis on some employment indicators will be performed, to show the expected evolution in the following years. Each of these steps is described in detail in the following sections.

3.3.1 Data Cleaning

Data cleaning transforms the raw dataset into a structured format suitable for the analyses, ensuring easier manipulation of the data and better interpretation and accuracy of the results.

The cleaning process was conducted using both Python and Excel: upon drawing the raw dataset from the World Bank Data Bank, it was observed that the data was structured in a wide format, with the years presented as individual variables. However, this format posed challenges for the analyses and the visualizations, making it necessary to restructure the dataset.

Therefore, the first step involved converting the dataset from a wide to a long format, which allowed for all indicators to be consolidated under a single variable named *Indicator Name,* with a separate variable *Year*, capturing the corresponding time period. This restructuring improved the efficiency of the data processing by enabling easier filtering, aggregation, and visualization of trends over time.

Subsequently, to facilitate gender-based analysis, an additional variable was created. The variable *GENDER* was introduced using a custom Python function, which leveraged the *Indicator Code* variable to accurately classify the data into three categories: male, female, and total. The function was designed to recognize gender-specific indicators by decoding the standard naming conventions used in the dataset. Indeed, the *Indicator Code* variable has a crucial role in the World Bank's database design and follows a standardized coding system known as the Catalog of Economic Time Series (CETS). The coding system is structured hierarchically as follows:

- **Topic**: 2 digits representing broad categories (e.g. SL: Social Protection and Labour)
- **General Subject**: 3 digits specifying subcategories (e.g. EMP: Employment)
- **Specific Subject**: 4 digits providing more detailed information (e.g. MPYR: Employers)
- **Extensions**: 2-digits suffixes that denote specific attributes (e.g. FE: Female)

Based on this coding structure, an example indicator code such as SL.EMP.MPYR.FE.ZS represents the share of female employers as a percentage of the total female employed population (ZS is a second extension which signifies *Share*).

Therefore, after a careful review to ensure the data types were all correct, a search for missing values was conducted. This process highlighted the presence of a high number of missing values, particularly in employment-specific indicators, and especially before the year 1990.

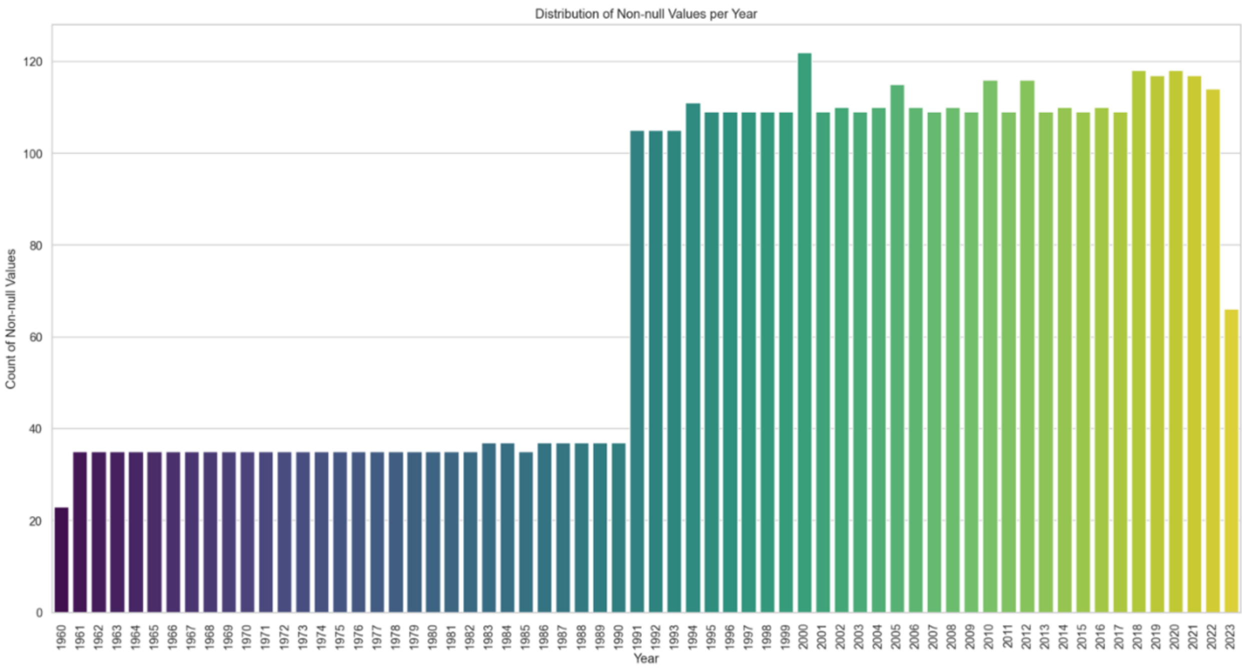

Figure 3.1 – Distribution of Non-Null Values Per Year

The *Figure* 3.1 shown above illustrates the distribution of non-null and non-missing values per year. In the early years the count is relatively low and stable, only increasing in 1991, which possibly suggests that data collection efforts were limited or inconsistent during the period between 1960 and

1990. After the noticeable increase in data from 1990, the number of available data points keeps rising and remains relatively high, albeit with some fluctuations. One noticeable aspect is the much lower availability of data in 2023 compared to the previous years: a possible reason might be the delays in reporting or the ongoing updates in the data collection processes by the data providers, as the data is not uploaded in the Data Bank simultaneously.
Given the limited availability of data before 1990, the analysis may focus more from the period of 1990 onwards, when data coverage is more reliable and consistent.

After the data cleaning process, the final dataset included the following key variables:

- **Year:** the year of data collection of the value for a specific indicator;
- **Country:** the geographical region (the three values are China, India and European Union);
- **Indicator Name:** the descriptive label for each urbanization or employment metric;
- **Indicator Value:** the numerical value associated with each indicator for a specific year;
- **Indicator Code:** the unique identifier of the indicator (based on CETS system);
- **GENDER:** a categorical variable indicating whether the data pertains to females, males, or the total population.

3.3.2 Exploratory Data Analysis

The Exploratory Data Analysis (EDA) step is a critical one in the data analysis process and focuses on examining and summarizing the main characteristics of the data using both statistical and visualization techniques. The primary goal of EDA is to uncover patterns, detect anomalies, test the hypotheses, and validate the assumptions, ultimately providing a deeper understanding of the data before conducting further analysis.
Through EDA, the researcher can gain insights on the distribution, the trends, relationships, and potential outliers within the dataset. This step is essential to ensure the quality of the data and to guide the selection of appropriate statistical models in later stages of the study.
A critical part of Exploratory Data Analysis are the visual representations, which help in communicating complex data patterns in an intuitive and often interpretable manner. Several visualization techniques have been employed to facilitate the analysis, each serving a specific purpose in understanding the underlying data patterns and in highlighting the different aspects of the data. The visualizations used include:

- *Line Graphs:* this type of graph is used to visualize trends over time and identify upward and downward trends, cyclical patterns, or periods of significant changes. Moreover, multiple lines, each representing a different class (the three different countries, for example) can be plotted together to show comparisons directly without the use of multiple graphs.

- *Histograms:* histograms are used to understand the distribution of numerical variables within the dataset, helping to identify patterns such as normality, skewness, and the presence of outliers
- *Box Plots:* these plots are effective in summarizing the spread and variability of the data, highlighting the median, quartiles, and potential outliers.
- *Bar Charts:* they a useful tool for comparing categorical data, offering clear visual comparisons across different categories. They can also be presented in a *stacked format* to incorporate an additional variable in the analysis, providing deeper insights into relationships within the data. However, careful attention must be given to their design, as improper stacking can reduce readability and interpretability.
- *Pyramid Chart:* used to visualize the distribution of indicators comparing female and male, provides a clear visualization of the differences in indicators across the values of the gender variable.

All these plots can be rendered interactive through the implementation of the Altair library, which provides easy-to-code functions and presents a pleasant user interface.

Interactivity plays an important role in enhancing the phase of exploratory data analysis by allowing the researcher to dynamically explore the dataset, adjust the views and gain deeper insights into the study. Unlike static visualization, interactive charts provide flexibility in analysing different aspects of the data, especially as they allow the user to filter, zoom and compare variables across regions and time periods.

In the study, interactivity was frequently used to reduce the number of graphs and explore the differences across countries or indicators, as well as time periods, as they allow for direct comparisons in one graph. The ability to customize visual elements, such as selecting specific indicators, or a specific country, helps to focus only on the areas of interest and target only the needed information.

Several types of interactivity tools were implemented to enhance the visualizations: tooltips, that provide detailed information when hovering over data points or data bars and showing the exact values, interactive filters (e.g. dropdowns selections) which allow the user to shift between indicators, gender or country, or dynamic legends where, as the user clicks on the values, the corresponding line or bar gets highlighted.

The exploratory data analysis phase was structured into two distinct sections, each focusing on different aspects of the dataset: *population data* and *employment-related data,* which allowed for a comprehensive examination of the dataset, ensuring that key insights were derived from both demographic and labour market perspectives.

Both sections were structured in the same way. They involved:

Descriptive Statistics: summary statistics were computed, including measures of central tendency, like mean or median, and dispersion, like standard deviation and range. This step is crucial to gain a general overview of the data.

Data Visualization: the charts reported above were used to explore the indicators and compare them. Different indicators were plotted against each other to visualize their relationships and the changes over time in terms of population growth and urban and rural growth rates, as well as employment data.

Distribution Analysis: the distribution of population, urbanization and growth rates was visualized through histograms and box plots, showcasing the differences across the countries, and detecting eventual variability and skewness in the data.

The findings of the exploratory phase provide a solid foundation for the next phase of correlation analysis, helping in understanding the meaningful relationships across the indicators.

3.3.3 Correlation Analysis

Correlation analysis is a fundamental statistical technique used to measure the strength and the direction of relationships between variables; in this thesis, correlation was conducted on population indicators and employment indicators. It was first analysed separately and then intersecting the indicators to gain a holistic understanding of their interdependencies.

The analysis was performed in various stages to improve the clarity and interpretability of the results: indeed, in order to facilitate effective visualization and interpretation, correlation matrices were plotted and displayed with varying indicators to avoid overcrowding and enhance graph readability.

Moreover, to provide a comprehensive understanding of the relationships within the indicators, the analysis was conducted at two levels: overall correlation, which analysed the combined data for all selected countries, and country-specific correlation, which examined the correlation within the indicators separately for each region to identify localized trends and differences.

Three widely used correlation techniques were applied to ensure robust and reliable results: Pearson's Correlation Coefficient (r), and Spearman's Rank Correlation (ρ).

Pearson's Correlation Coefficient (r) measures the linear relationship between two continuous indicators, meaning that a change in the magnitude of one indicator is associated to a change in the magnitude of the other, either positively (same direction) or negatively (opposite direction). Pearson's correlation assumes that data is continuous and normally distributed and provides values ranging from -1 (perfect negative correlation) to 1 (perfect positive correlation), with 0 indicating no correlation; a perfect correlation of -1 or +1 indicates that all the data points lie exactly on the straight line, although such results are very rare.

Spearman Rank Correlation (ρ), instead, is a non-parametric technique that measures the strength and direction of a monotonic relationship between two variables, and therefore does not assume a normal distribution, ranking the data before calculating the correlation. Analogous to Pearson coefficient, Spearman's results range from -1 to +1, with $\rho = -1$ or $+1$ describing perfect monotonic relationship, and $\rho = 0$ describing no association between indicators.

It is crucial to understand how to correctly interpret the results of correlation. Firstly, an observed correlation, or association, between two variables does not imply a causal relationship: while a causal relationship may exist, correlation analysis alone cannot justify such a conclusion. Moreover, correlation also does not measure the level of agreement between two variables: it is possible for the variables to exhibit a high degree of correlation while demonstrating substantial disagreement.

Correlation results were visualized in order to enhance their interpretability, mostly using heatmaps, which provide a summary of correlation coefficients and use colour gradients to differentiate from positive and negative correlations, allowing for quick identification of strong and weak relationships. Furthermore, in order to inspect relationships between key indicators and when the correlation results were unexpected, scatterplots and line graphs were also used to gain better insights into the relationship between the data points.

3.3.4 Forecasting Analysis

The final analysis in this study focuses on forecasting the participation of women in the three major employment sectors: agriculture, industries and services. These sectors represent key components of economic development, as understanding their trends provides valuable insights into sectoral shifts, gender-based labour market participation and economic transitions over time.

The method used for the forecasting is the AutoRegressive Integrated Moving Average (ARIMA).

ARIMA is a statistical forecasting method that models time series data based on past values and past forecast errors, and it is widely used for the predictions of economic and financial data due to its ability to capture underlying patterns and make accurate future predictions.

Stationarity of the series is crucial as it is an assumption of the ARIMA model: a time series is stationary when its mean, variance and autocorrelation remain constant over time, if a series is not stationary, the most common approach is to difference it, which means the previous value is subtracted from the current value.

To confirm whether a series is stationary, the Augmented Dickey-Fuller (ADF) test is applied, a statistical test that examines the presence of a unit root in the data. When computing the ADF, if the resulting p-value is lower than the significance level of 0.05, the null hypothesis of the non-stationarity of the series is rejected, thus confirming that the series is indeed stationary.

In cases where the null hypothesis cannot be rejected (p-value higher than 0.05), then differencing is applied, therefore removing the trends and making the series stable for forecasting.

The ARIMA model is characterized by three parameters: p (the order of the AutoRegressive AR term), q (the order of the Moving Average MA term) and d (the number of differencing required to make the time series stationary).

In other words, the parameter p represents the number of past values (or lags) which is used to predict future values, while q is the number of past forecast errors considered in the model. The selection of the parameters is guided by the plots of the Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF):

- The ACF shows how strongly previous observations influence future values. If the ACF decreases gradually, it suggests that the series is non-stationary, thus requiring differencing. Instead, if the ACF cuts off sharply after a certain lag, then it suggests a Moving Average (MA) process, which helps determine the q (the number of MA terms).
- The PACF measures the correlation between a time series and its past values while removing the influence of the intermediate lags. When it cuts off sharply after a certain lag, it suggests an AutoRegressive (AR) process which helps determine the p. Instead, if it has a gradual decline, it might indicate either a mix of AR and MA components or an over-differenced series.

To summarize, ACF helps choose the q parameter, while PACF helps choose the p; however, if both slowly decay, then further differencing is needed to make the series stationary, thus increasing the parameter d.

Once the parameters are selected to ensure an optimal fit, the ARIMA model is trained on a training set and validated using a test set to assess its accuracy. The model’s performance is evaluated using two error measures: the Mean Absolute Error (MAE) and the Root Mean Squared Error (RMSE), while the residuals are then analysed through a residual plot, where the residuals (the differences between the actual and the predicted values) should be randomly distributed around zero, indicating that the model captures all underlying patterns.

The MAE and the RMSE are two of the most commonly used evaluation metrics: the MAE measures the average absolute difference between the actual values and the predicted values, treating all errors equally, thus not differencing between large and small errors. It is generally an easy-to-interpret and to-compare measure as the results are in the same unit as the data. It is also robust against outliers.
The RMSE instead calculates the square root of the average squared differences between the actual and the predicted values, therefore penalizing large errors more than small errors, thus making it a useful measure for models where large forecasting errors are unacceptable.
Finally, once the model is ready and has demonstrated a good performance on the test set, it is used to forecast the indicators in the next ten years: the three indicators and the three countries will be treated separately, each with its own model, to ensure the best fitting of the data and therefore the best forecasting.

In conclusion, this chapter has provided a comprehensive overview of the methodological framework employed in this study, thus ensuring a structured and rigorous approach to addressing the research questions. This study adopts a longitudinal correlational research design, using secondary data to examine the trends in urbanization and women employment, by leveraging a combination of exploratory data analysis, correlation analysis and predictive modelling. The methodological approach implements the use of several visualization techniques to identify patterns and visualize key trends, ensuring clear and interpretable results.
In addition, the study employs a forecasting analysis using the ARIMA model to predict future trends and contributing valuable insights into the future of women in employment.
Overall, this methodological framework establishes a systematic, reproducible, and analytical approach, ensuring the study’s relevance in both academic and policy discussions.
The subsequent chapters will present the results of these analyses.

CHAPTER 4. Thematic Analysis

This chapter presents the thematic analysis of the study, focusing on key research themes and their interconnections across China, India and European Union. The analysis follows a structured approach, employing the methodologies outlined in the previous chapter to ensure a systematic and data-driven investigation of the research questions.

All the analysis have been computed in Python, with the code provided at the **following link**. Moreover, some visualizations have been implemented using Tableau, and can be accessed at this folder.

The chapter is divided into three main sections: *Population and Urbanization Dynamics, Employment Indicators Analysis* and *Exploring the Impact of Urbanization on Employment Dynamics*.

The first section will examine the evolution of urban growth across the three regions, highlighting key trends in growth rates, population shifts and rural-urban transitions; the analysis is supported by visualizations and statistical insights to illustrate regional differences and long-term patterns.

The second section will deep dive into the employment trends, exploring the sectoral employment distribution, types of employment and gender disparities, identifying shifts in labour market dynamics and providing a comparative perspective between developing and developed countries.

The final section will analyse how urbanization relates to the employment structures, focusing on correlation patterns, and thus providing a perspective on the socio-economic aspects of urbanization, comparing trends across the countries in analysis.

All findings will be interpreted within the broader context of economic development, highlighting both commonalities and divergences in how urbanization shapes employment structures in different regions.

4.1 Population and Urbanization Dynamics

This section will explore the dynamics of the interplay between population trends and urban development and growth by conducting an in-depth analysis of key population and urbanization metrics and their interrelationships. It will begin with the exploratory data analysis to examine and visualise the population and urbanization indicators, as it aims to uncover underlying patterns and structures in the data: it will provide an initial understanding of how population growth and urban expansion unfolded over time, allowing to identify variations and historical trends that may influence further analysis. The section will then transition to a correlation analysis to assess the strength and the direction of relationships between key indicators. This dual approach will allow to explore and establish meaningful connection between the indicators, providing a deeper understanding of the trends in urban expansion.

Table 8.1 of the Appendix in Chapter 8 reports a comprehensive list of all indicators analyzed and their respective definitions given by the World Bank.

4.1.1 Exploratory Data Analysis of Population Statistics

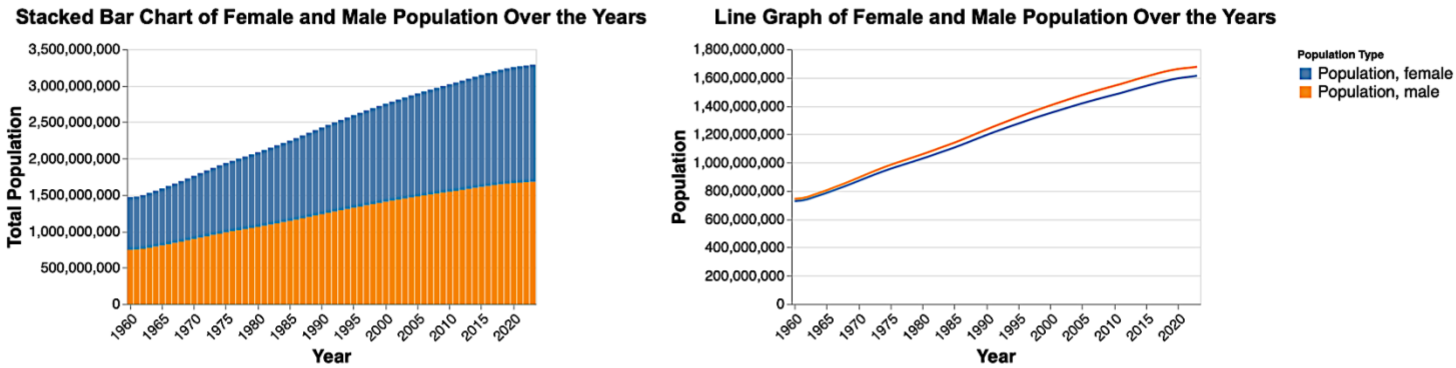

Figure 4.1 – Bar Chart (right) and Line Chart (left) of Female and Male Population Over the Years in China, India and the European Union

This first section starts with an overview of the total population trends in China, India and European union, analysing how their size evolved over time, followed by a more detailed examination of population growth rates, which focuses on how the rate of increase differs across the regions.

Starting from a general overview of the population, in *Figure 4.1* the two charts show the evolution of female and male population from 1960 to 2023, in China, India and EU. Two visualizations were chosen and printed one beside the other to facilitate comparisons: the stacked bar chart shows the total population structure, allowing a clear visual of the contributions of male and female; the line graph provides a more detailed trend analysis, making it easier to observe any fluctuations or deviations and explicating the growth.

From the charts it seems that the total population of China, India and EU shows a continuous upward trend, which indicates a significant demographic expansion. The population increase appears steady and consistent, with no fluctuations and declines: in the line graph it is shown that the male population (orange) consistently outnumbers the female population (blue) throughout the entire time span. However, both genders exhibit similar growth rates, maintaining a relatively stable gender ratio over time. Indeed, their lines follow nearly identical trajectories, with the male population widening the gap in recent years.

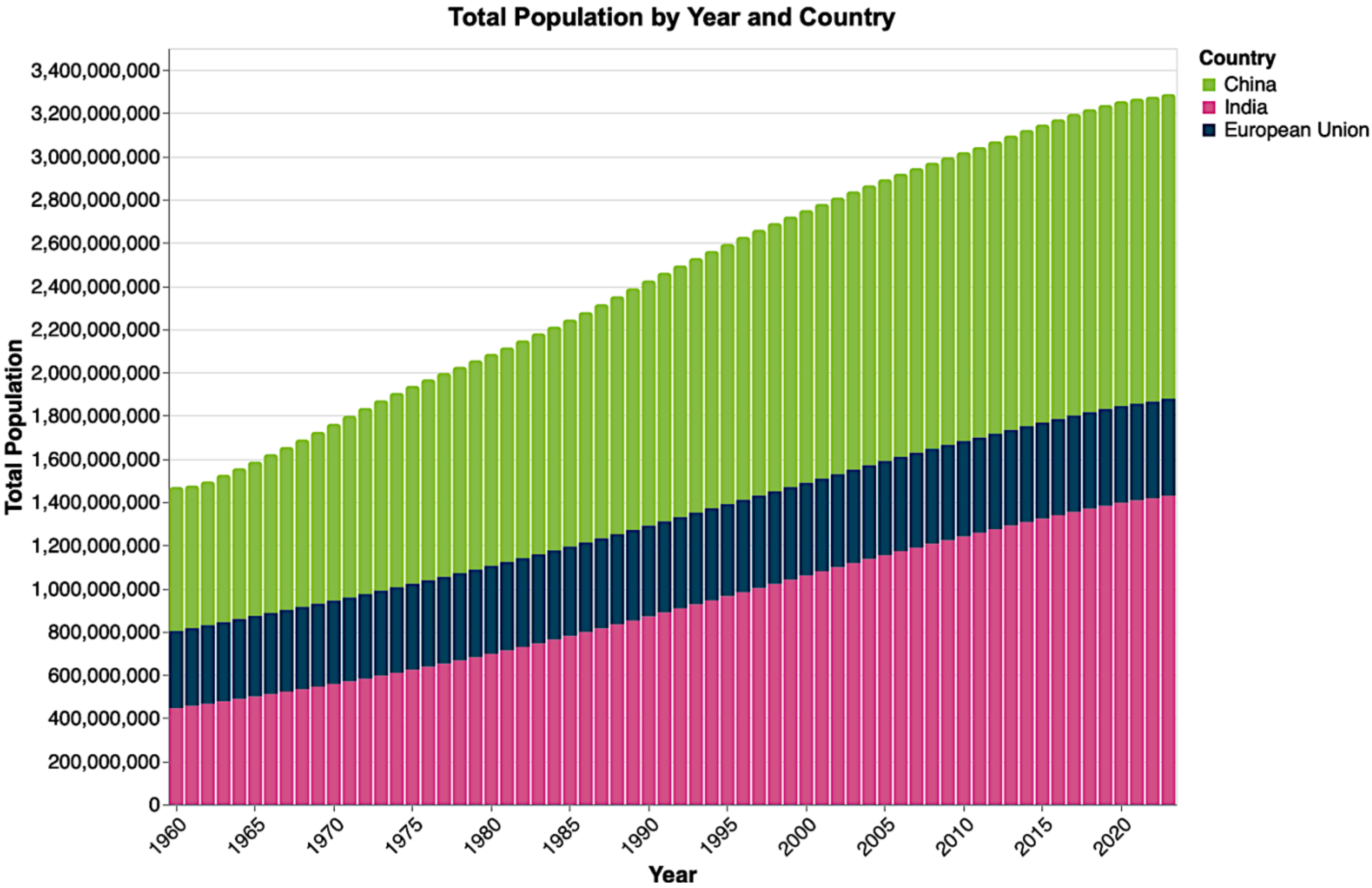

Figure 4.2 - Total Population by Country

The comparison now moves to the evolution of the population across the three countries. In *Figure 4.2* the stacked bar chart illustrates total population trends in the three countries from 1960 to 2023. This type of visualization has been chosen as it allows the cumulative population to be represented in a single view, enabling a direct visual comparison between the population sizes in the countries.

The overall population across the three regions has steadily increased over the decades, showing a consistent upward trajectory, with China and India as the primary drivers of population growth, while the European Union exhibits a more gradual increase.

It is notable that there is no period of decline, indicating that all three regions have experienced continued population expansion. In the earlier decades (1960-1980), China had a substantially larger population than both India and EU; however, India's population experienced a steeper growth by later years, from 2010 to 2020s, almost bridging the gap with China. This suggests that India is on track to surpass China's population, reflecting higher birth rates and continuous demographic expansions. The European Union's population growth, instead, is especially slower compared to the other two countries: EU has a more gradual increase, showing signs of stabilization, possibly due to lower birth rates and an increasingly aging demographic.

To gain a deeper understanding of population dynamics, the annual population growth rate was plotted for each country. This allows for a more detailed analysis of how population growth has evolved over time, highlighting the fluctuations and trends, as well as demographic patterns.

To enhance visualization and interactivity, an interactive line chart has been implemented: when selecting a country from the legend, the corresponding line remains highlighted while the others fade, making it easier to focus on a specific country while still allowing for comparisons across regions. This feature enables both detailed single-country analysis and cross-country comparisons, improving the clarity and interpretability of trends.

Figure 4.3 shows the annual population growth: generally, all three regions seem to have an overall decline in population growth rates, although at different paces and with different patterns. The European Union has consistently the lowest growth rate and has fluctuated around a much smaller range compared to the other two; China and India, as expected, have a much higher growth rate, with India being consistently at the top.

Focusing on China, represented in *Figure 4.4*, it is interesting to note that there is a spike in population growth in the very early years, reaching a peak in the middle of the 1960s, before undergoing a steep decline in the late 1970s and 1980s. China's growth continues to slow down, albeit with some fluctuations, such as an increase during the later years of the 1980s, followed by another decrease in

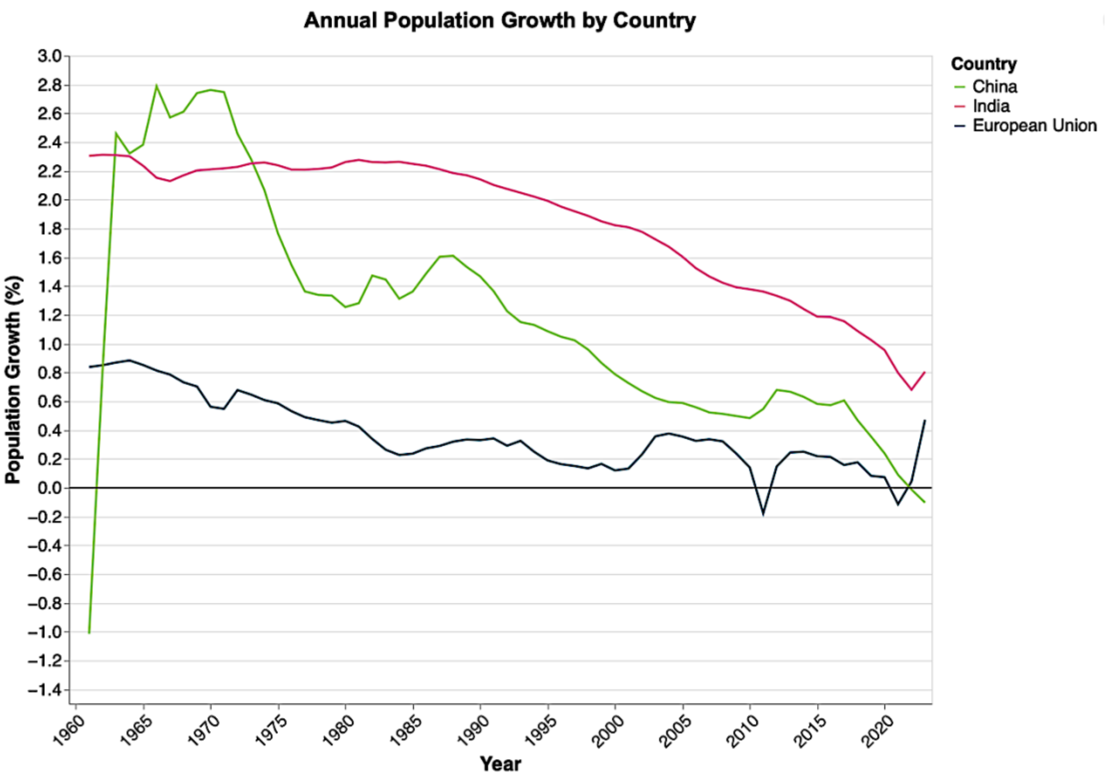

Figure 4.3 - Annual Population Growth by Country

the next decade. In recent years, after a slight increase, the rate eventually dropped below 0%, signalling population stagnation and potential decline.

On the other hand, India, shown in *Figure 4.5*, had a much more stable population, consistently remaining above 2% until the late 1980s; however, after the 1990s, a gradual decline is observed, which continues until recent years. It seems that, after reaching its lowest point in 2022, there might be another increase from 2023 onwards.

Figure 4.4 - Annual Population Growth - China

Finally, the European Union's population growth (*Figure 4.6*) is shown to have consistently remained relatively low, typically below 1%: after reaching its highest percentage during the 1960s, growth rate has declined up until the 1990s, where it has fluctuated around 0.2%-0.5%. It also showed a negative growth twice in recent years, which was typically followed by a

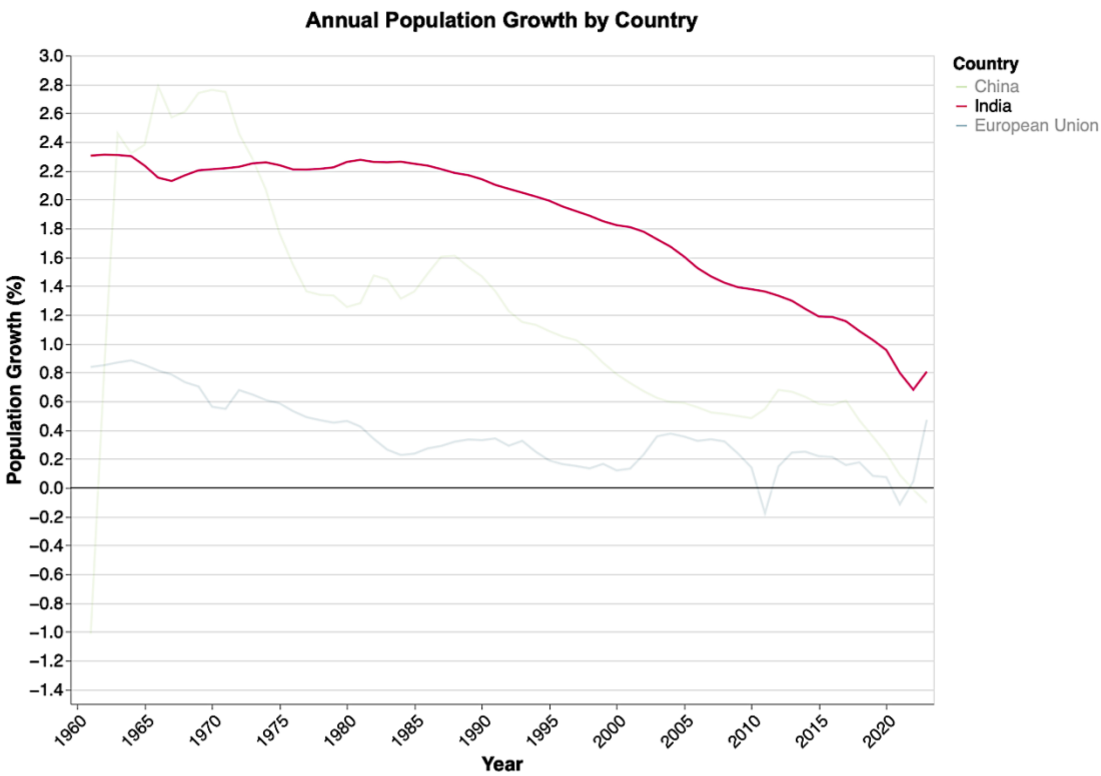

Figure 4.5 - Annual Population Growth - India

sharp increase. It seems that in the last two years the rate is increasing, reaching the highest growth since the 1990s.

Overall, while the total population continues to increase, the growth rate is steadily declining: India and China, despite both of their large populations, exhibit signs of demographic slowdowns, while the European Union's low and fluctuating growth signals a demographic stagnation.

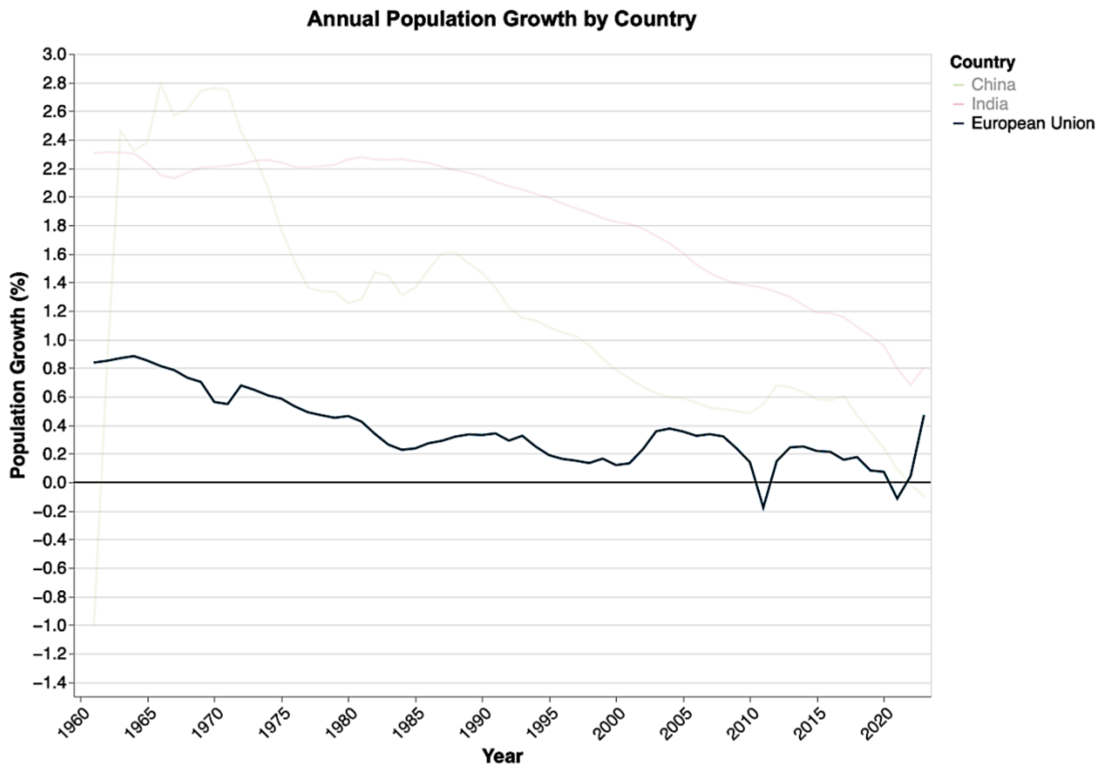

Figure 4.6 - Annual Population Growth - EU

Building on these findings on population growth, the focus now shifts to where the growth is occurring by analysing the urban and rural growth rates, illustrated in *Figure 4.7*. Again, two line graphs showing the growth rates are set one beside the other to enable direct and immediate comparisons between the two rates.

Both urban population growth (left) and rural population growth (right) data range from 1960 to 2023 for all countries in analysis.

Urbanization seems to be increasing in all three regions, as the growth rates never declines below 0%, while rural population growth is steadily declining.

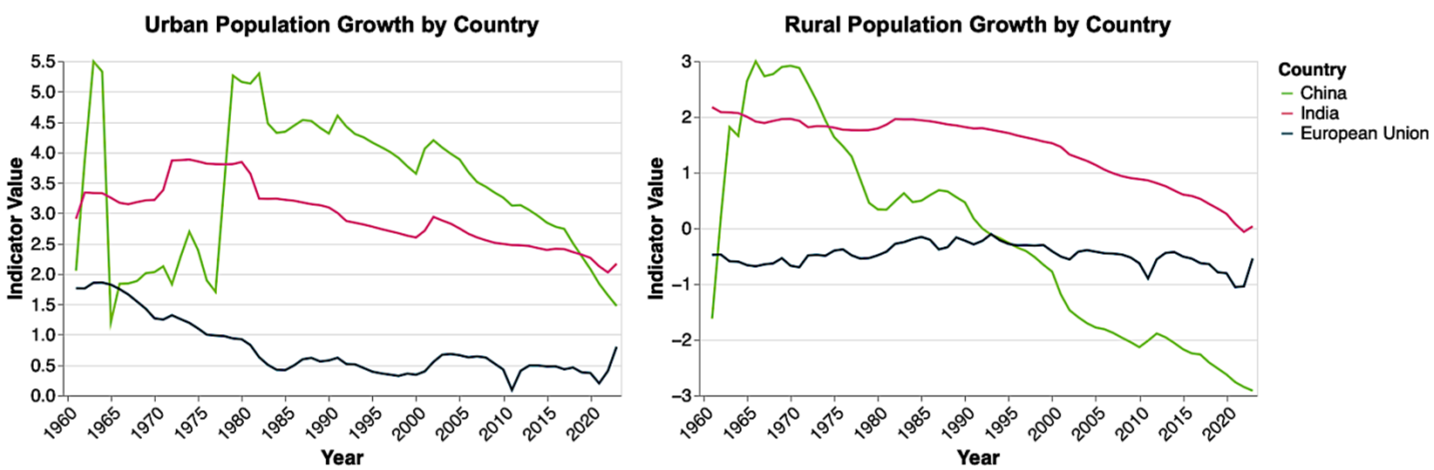

Figure 4.7 - Urban Population Growth (left) and Rural Population Growth (right)

China has experienced one of the most extreme cases of urbanization, with two distinct peaks in 1960s and in 1980s, followed by a gradual decline in urban growth rates after the 1990s and reaching one of its lowest values in 2023. It suggests that urban expansion in China is stabilizing as the country approaches high levels of urbanization. Conversely, rural population growth in China has sharply declined, turning negative after the 1990s and indicating a continuous rural-to-urban migration trend. Notably, from 1960 to 1980, urban growth rate and rural growth rate displayed opposite trends, suggesting a period of rural expansion and urban decline, with population moving from urban centres to rural areas.

In contrast, India's urban growth rate has been consistently high, fluctuating around 3%, despite still showing a gradual decline over time. Moreover, in recent years, there are indications of a slight upward trend, suggesting a potential resurgence in urban expansion. Similarly, rural population growth has been declining, though at a much slower rate compared to China, only recently reaching slightly negative values. This suggests that while India's population is gradually shifting toward urban areas, a significant portion continues to reside in rural regions.

Finally, the European Union exhibits the slowest urban growth rate, consistently remaining below 2% throughout the analyzed time period, which reflects the fact that many European countries were already urbanized before 1960, resulting in a more stable urban population. This trend is further reinforced by the rural growth rate, which has steadily remained negative, indicating a gradual rural population decline, a common characteristic of developed economies where urbanization reached maturity decades ago.

These trends are clearly in the urban population percentage, illustrated in *Figure 4.8.*

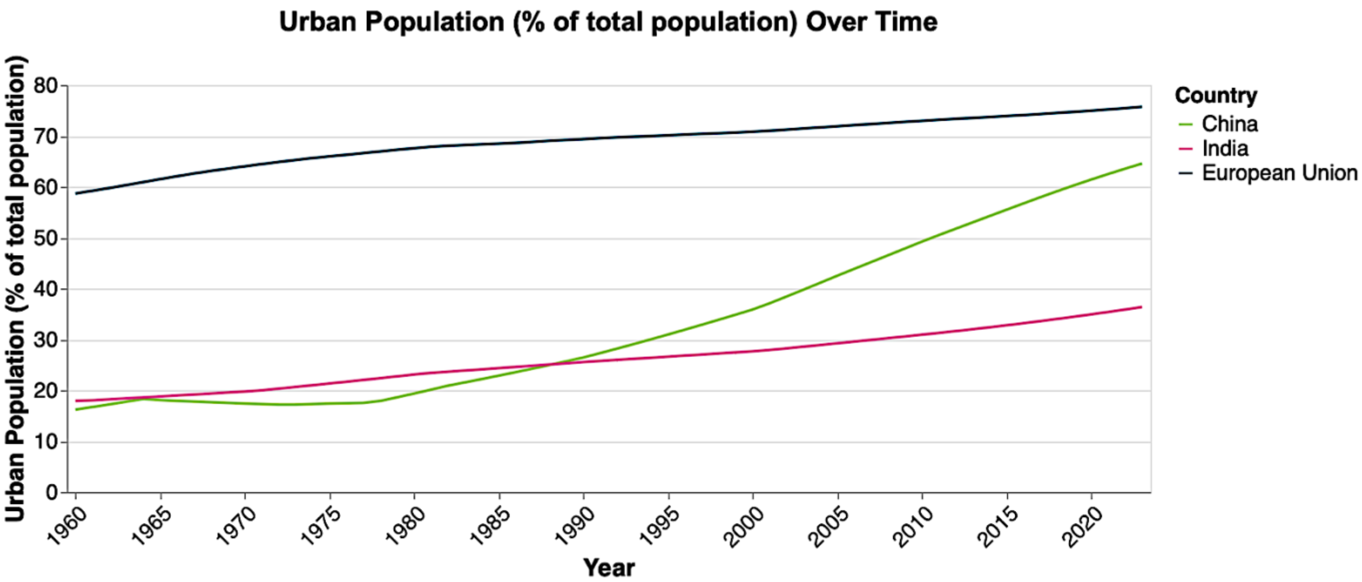

Figure 4.8 - Urban Population (% of Total Population) Over Time

As expected, the European Union has consistently maintained the highest percentage of urban population, starting at around 60% in 1960 and gradually increasing to nearly 80% in 2023: the slow and steady increase aligns with the patterns observed in *Figure 4.7,* confirming that most European countries were already highly urbanized before 1960. This gradual increase suggests mature urbanization, where growth is primarily driven by natural population increases and migration, rather than large-scale rural-to-urban population shifts.

On the contrary, China's urbanization has been far more dramatic: until the 1980s the urban population percentage remained relatively stable but has increased exponentially in the following years. Indeed, in 1960, only 16% of China's population lived in urban areas, but this figure has since surged to over 65% in recent years, which might suggest that China is still undergoing a significant urban transformation. From *Figure 4.8* it seems that China and India have had similar urban population percentages until 1988, when China urbanization surpassed India's significantly.
Indeed, India has experienced a much more gradual urbanization process and has the lowest percentage of urban population, starting at just under 20% in 1960 and rising to slightly above 35% in 2023: unlike China, India's urbanization process seems to be steady rather than abrupt, somehow mirroring the trend observed in the European Union, which is also consistent with India's urban growth rate. The slower pace of urbanization suggests that India's rural economy remains a dominant force in its demographic structure, with a significant portion of the population still residing in rural areas. These patterns underscore the different stages of urbanization in each region, with the European Union already mostly urbanized, China experiencing rapid urban expansion and India still undergoing a gradual transformation.

Another commonly used indicator to measure urbanization is the percentage of population living in urban agglomerations of more than 1 million people. *Figure 4.9* illustrates this metric, firstly focusing on the period from 1960 to 1981, offering insights into early concentration patterns in China, India, and the European Union. The trend of the indicator has been divided in three sections comprised all of the same number of years (section 1 in *Figure 4.9*: 1960-1981; section 2 in *Figure 4.10*: 1982-2003, section 3 in *Figure 4.11*: 2004-2023) to provide a more in-depth analysis of the indicator's evolution.
To effectively visualize these trends, a line chart has been used to depict the evolution over time, while a packed bubble chart, created with Tableau, provides an intuitive snapshot of the urban distribution. Although the bubble chart is interactive in Tableau, the printed version reflects the situation in 1960, as the general pattern remains relatively unchanged throughout the period 1960-

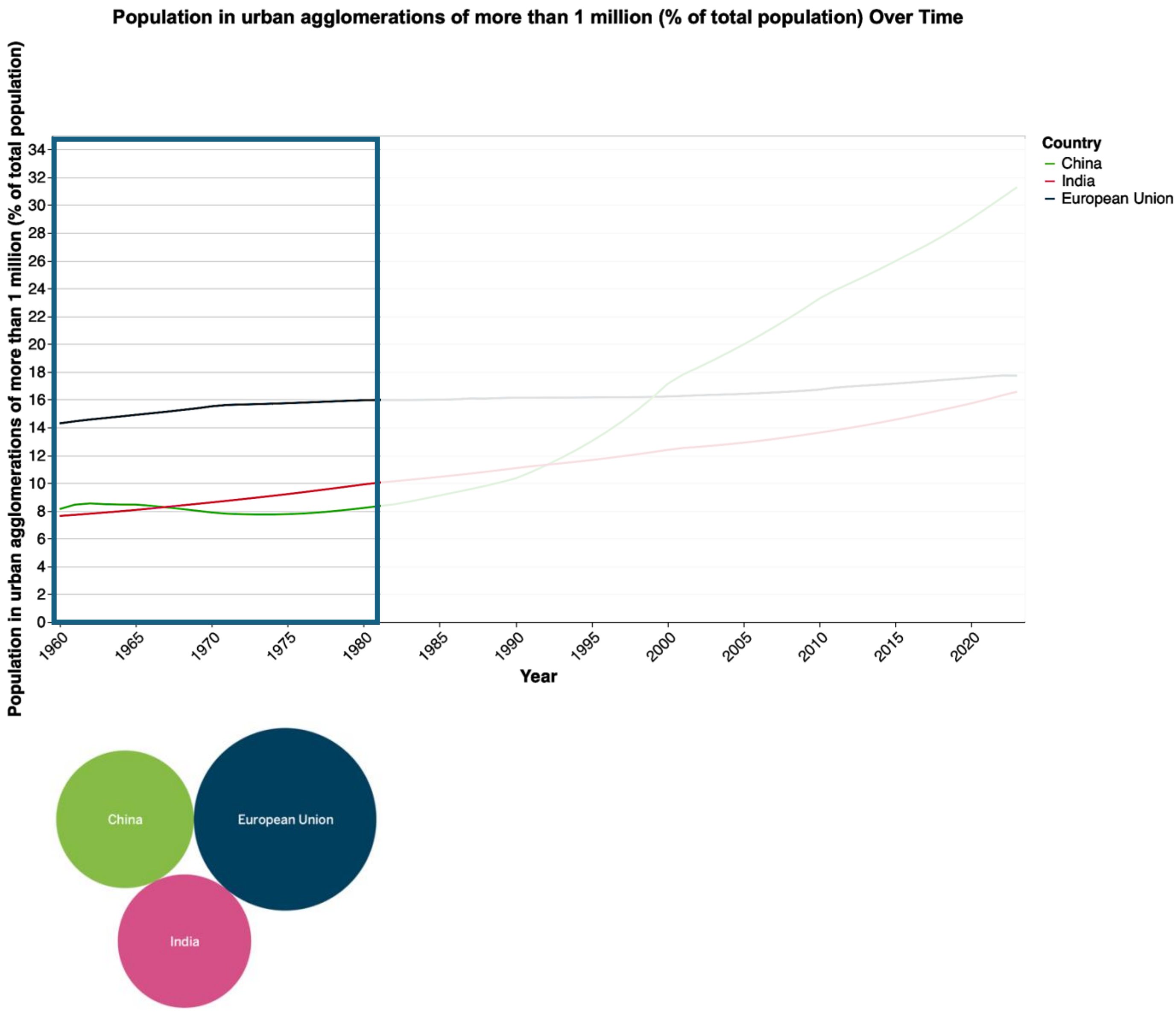

Figure 4.9 – Line Chart (above) and Bubble Chart (below) of Population in urban agglomerations of more than 1 million (% of total urban population) from 1960 to 1981

1981. The data reveals that the European Union consistently had the highest percentage of its urban population residing in large urban agglomerations, maintaining a steady trend around 15%-16%, and aligning with previous findings, indicating that the major European areas had already been well established before 1960. It can also be quickly deduced by the bubble chart, as the European Union bubble is significantly bigger than the other two. In contrast, China and India started with significantly lower shares, indicating a more dispersed urban population at the time; however, there is a gradual increase that can be observed for both countries, particularly in India. China's trend remained relatively stable during this period, albeit with some fluctuations.

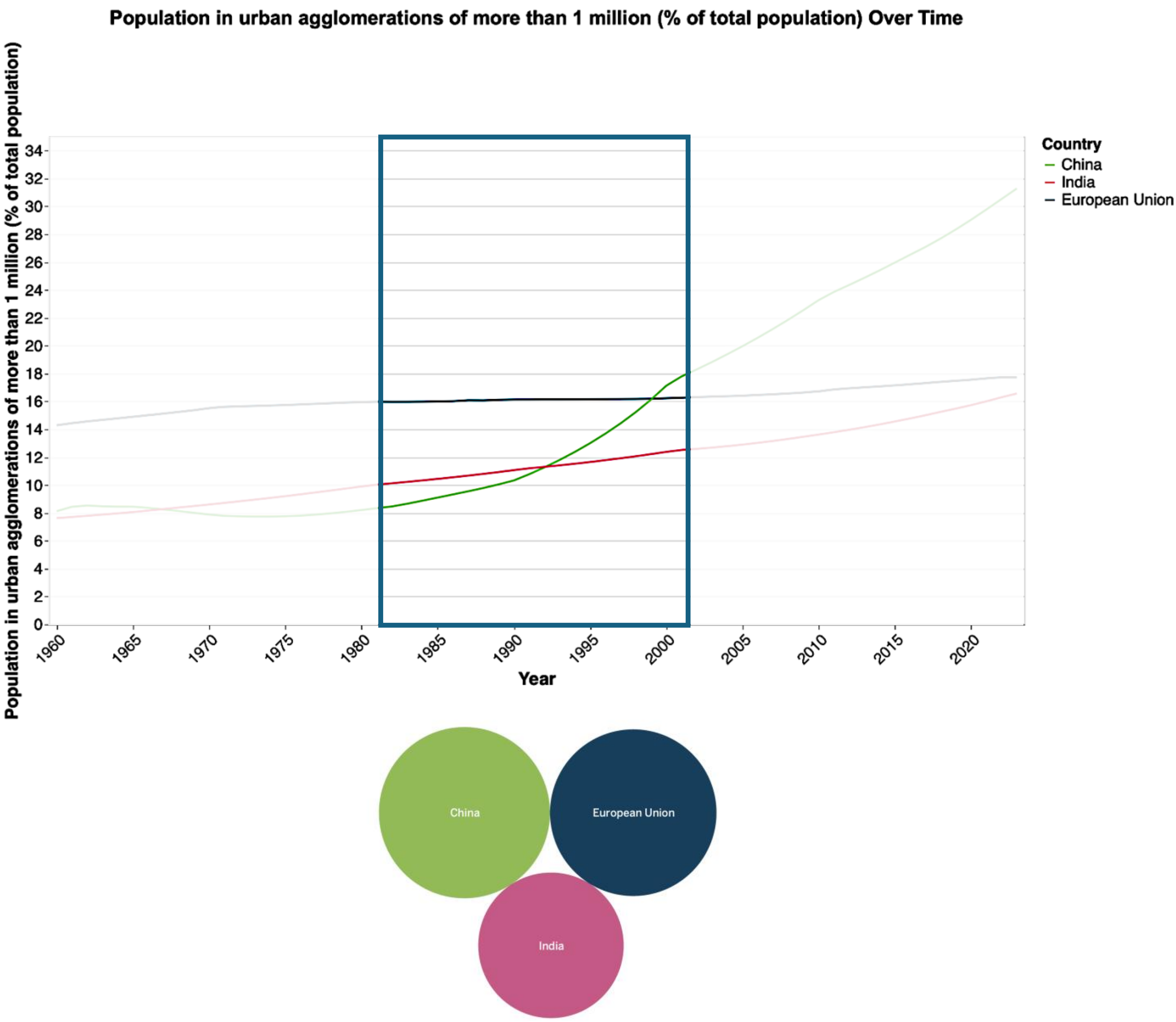

Figure 4.10 – Line Chart (above) and Bubble Chart (below) of Population in urban agglomerations of more than 1 million (% of total urban population) from 1982 to 2003

Over the next two decades, urban concentration patterns shifted significantly across the three economies, as illustrated in *Figure 4.10*. Once again, a bubble chart and a line chart are used in combination to effectively capture these changes over time.

Some key insights can be derived from the charts: the European Union continues to display an almost constant trend in the proportion of people living in large urban agglomerations, indicating that a plateau in urban growth has been reached in the region. On the contrary, India and China have an increasingly upward trajectory.

India shows a consistent rise, suggesting a gradual but steady urban growth, with increasingly more people living in major cities; however, the most dramatic transformation is shown in China, starting with the lowest percentage of people living in urban agglomerations, but then experiencing an exponential growth since the beginning of the 1990s, thus surpassing both China and European Union. Indeed, the bubble chart, set in the year 2000, visually reinforces this trend. Unlike in the earlier period, China's bubble has now surpassed that of the EU, and when comparing *Figure 4.9* and *Figure 4.10,* China's explosive urban concentration becomes even more pronounced.

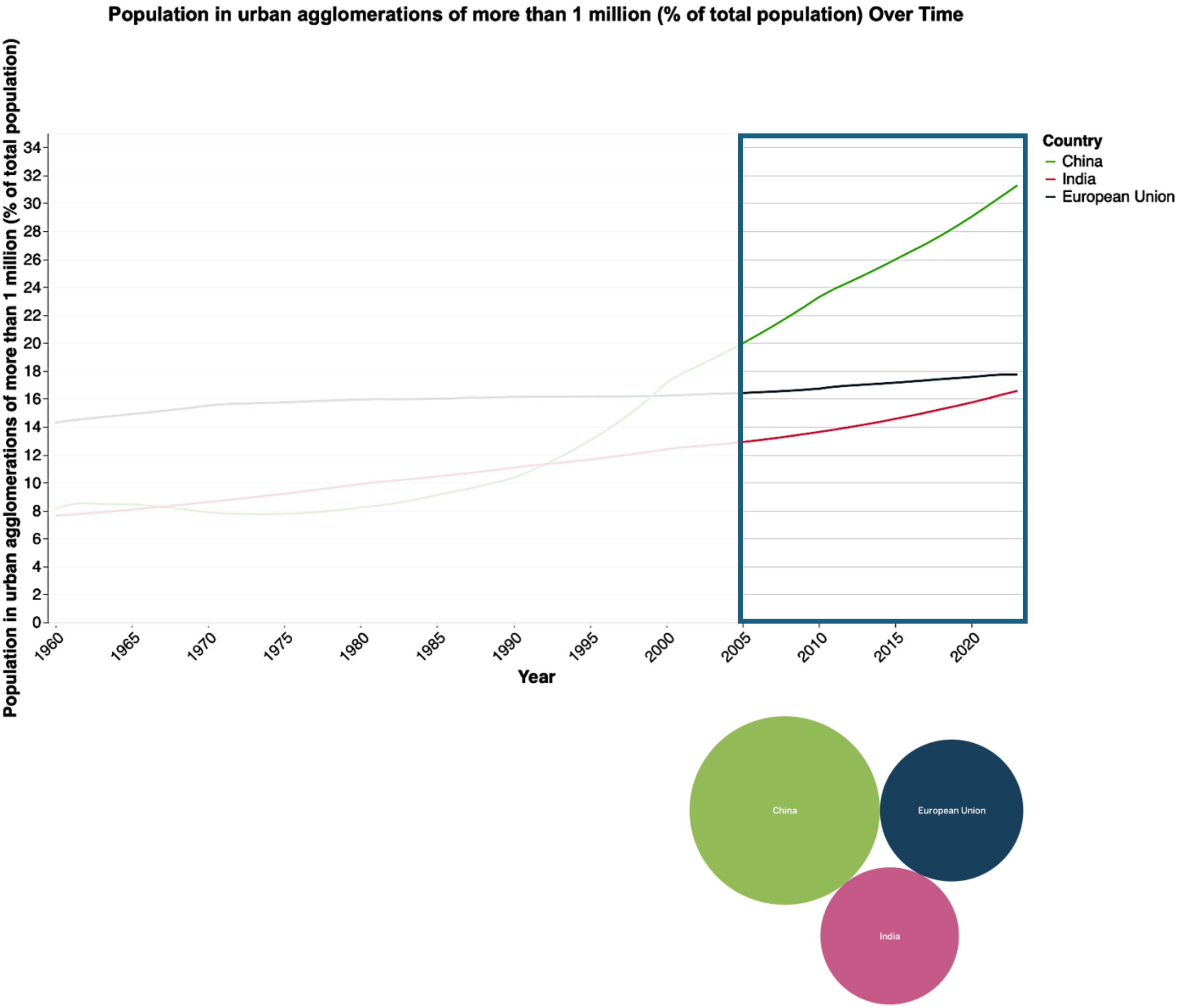

Figure 4.11 – Line Chart (above) and Bubble Chart (below) of Population in urban agglomerations of more than 1 million (% of total urban population) from 2004 to 2023

Finally, in the last two decades, China's dominance in urban concentration becomes even more pronounced: its shares of people living in urban areas continues to rise sharply, significantly outpacing both India and EU. This clearly shows China's ongoing urban expansion and the rapid development of its megacities. In *Figure 4.11*, it is reflected how significantly higher China's urban growth has been compared to the other two economies: India has followed a slower but steady upward trajectory, with a gradually increasing percentage of people living in urban centres and almost reach the European Union's level. The latter maintains a stable trend: there has been some minor growth, but the increase is relatively small compared to the other two regions.

The bubble chart, set in the year 2023, highlights the dramatic changes of urban growth in China: when comparing with the other two previous charts, the bubble is now much bigger than the other

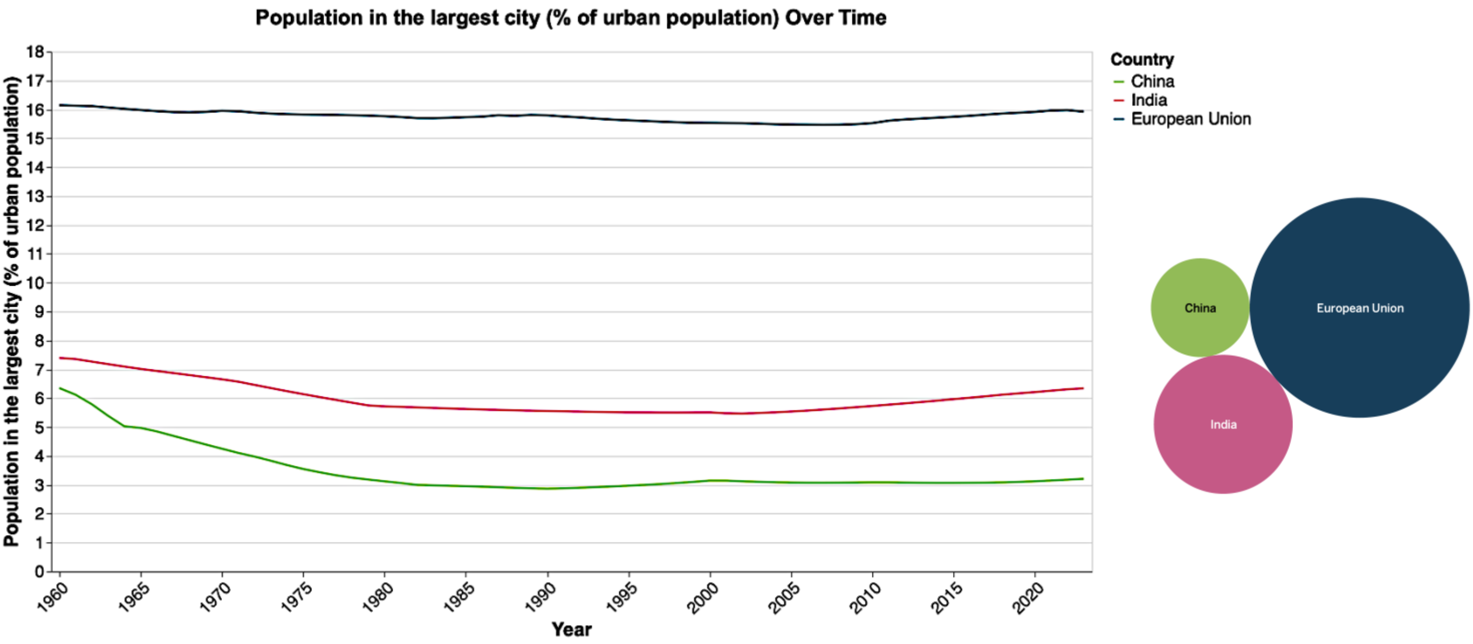

Figure 4.12 – Line Chart (right) and Bubble Chart (left) of Population in the largest city (% of urban population)

two, visually confirming its leading position.

Additional interesting statistics on urban population distribution are *Population living in the largest city (% of urban population)* and *Population density (people per sq. km of land area).*

The share of the urban population living in the largest metropolitan area, depicted in *Figure 4.12*, highlights distinct differences across the three regions: on one hand, both China and India exhibit a relatively low percentage of urban residents living in the largest metropolitan areas, with China showing the lowest value. This pattern suggests that in these countries, where largest cities are the increasingly common megacities, people generally prefer to settle in smaller metropolitan areas.

Conversely, in the European Union, where cities tend to be smaller and more evenly distributed, a greater share of urban population gravitates towards the largest cities. The bubble chart in *Figure 4.12*, set in the year 2023, demonstrates this trend, which has remained relatively stable over the 1960-2023 time period and indicates the European Union dominance in this aspect, with a significant higher proportion of urban residents concentrated in its largest metropolitan areas compared to India and China.

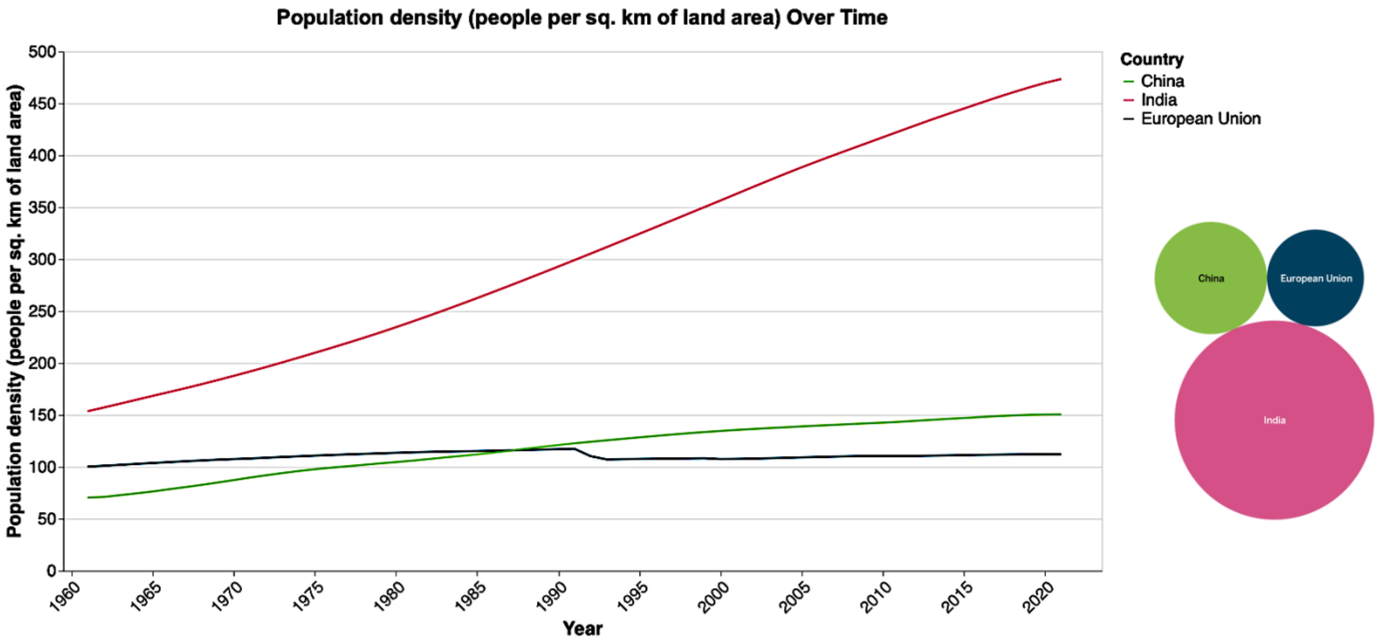

Figure 4.13 – Line Chart (left) and Bubble Chart (right) of Population density (people per sq. km of land area)

Finally, it is interesting to look at another indicator: the population density, which varies dramatically across the three regions. As illustrated in *Figure 4.13*, both in the line chart and in the bubble chart (set in 2023 to enhance the difference between India and the other two countries), India stands out with the highest population density, exhibiting an exponential growth, from 150 people in 1960 to over 450 people in 2023. China has also seen a steady increase, though at a much slower pace, while the European Union has experienced a decline. This divergence may be attributed to differences in population growth rates, as India and China's populations keep growing fast, whereas the European Union's steadier population growth has led to a declining density over time.

4.1.2 Correlation Analysis of Population Statistics

This section now presents a correlation analysis that will provide insights into the relationships across the indicators and how they move together, allowing to determine whether changes in one indicator are associated with changes in another. Indeed, understanding correlations is crucial for drawing meaningful conclusions from population data.

As anticipated in the Research Methodology chapter, two correlation techniques have been used in this study: Pearson's Correlation Coefficient and Spearman's Rank Correlation. In this section, only the latter has been used as method to calculate the correlation, as it was deemed more suitable to the data.

Spearman's Rank Correlation measures the strength of the monotonic relationship, meaning that when calculating the correlation between two variables, the correlation is positive if the values of one variable increase with the values of the other variable, while it is negative if as the values of one variable increase, the ones of the other variable decrease.

The correlation has been calculated using a dataset with all countries in analysis, as the purpose was to explore the general patterns and relationships globally, rather than focusing on individual regions, thus allowing to identify trends that might be present universally.

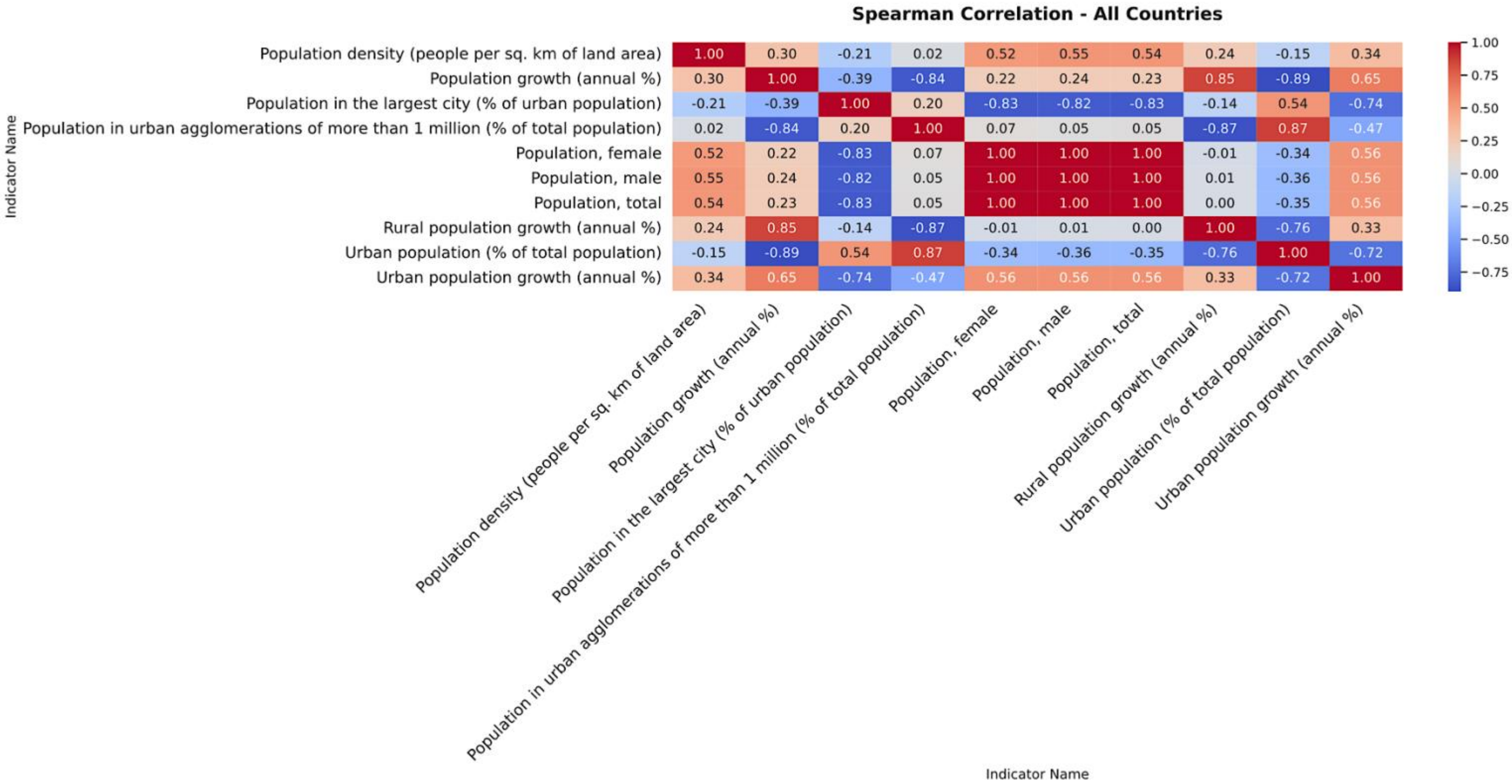

Figure 4.14 - Spearman Correlation of Population Statistics

Moreover, focusing on a specific subset of the data might introduce bias and might result in missing data trends, meaning that a larger dataset might be best for more robust correlations.

Figure 4.14 illustrates the correlation of the population statistics computed through the Spearman's Rank Correlation method. It is important to note that the indicators *Population, total*, *Population, male* and *Population, female* are perfectly correlated with one another, which was to be expected as the female population and male population follow the same trend and together represent the total of the population.

Starting from those indicators most positively correlated to one another, the indicator *Population in urban agglomerations of more than 1 million (% of total population)* has a high positive correlation (0.87) with *Urban population (% of total population)*: it is coherent with what has been found until now, as it indicates that the increase in urban population parallels the increase in larger cities, and explains why both indicators have been used to describe urbanization.

The indicator *Population in urban agglomerations of more than 1 million (% of total population)* has a high negative correlation (-0.87) with the *Rural population growth (annual %)* indicator, possibly representing the population migrating from rural to urban areas since rural growth decreases as the population moves to urban centres.

Another interesting aspect is that *Population growth (annual %)* has a high positive correlation with *Rural population growth (annual %)* (0.85) and a moderately high positive correlation with *Urban population growth (annual %)* (0.65), but has a high negative correlation with *Urban population (% of total population)* (-0.89). These relationships are interesting because they reflect what has been previously shown in this chapter: as population growth decreases, so do the urban population growth and rural population growth, but *Urban population (% of total population)* keeps increasing. These findings do not establish a direct cause-and-effect relationship between the indicators; rather, they highlight the presence of correlations or similar trends. This means that while the indicators may move in the same direction or exhibit related patterns over time, it does not necessarily imply that one variable directly influences or determines changes in the other. External factors, such as economic policies, social dynamics, and labor market conditions, may simultaneously impact both indicators, leading to observed similarities without a direct causal link.

Moreover, although it was to be expected that *Rural population growth (annual %)* has a high negative correlation with *Urban population (% of total population)* (-0.76), it is also important to note that rural growth has a slightly positive correlation with *Urban population growth (annual %)* (0.33): it was, indeed, illustrated in *Figure 4.7* that the two indicators do show a similar decreasing pattern, thus allowing for a moderately high monotonic relationship.

In addition, *Population in the largest city (% of urban population)* is highly negatively correlated with all the population indicators (*Population, total*: -0.83; *Population, male:* -0.82; *Population, female*: -0.83), suggesting that although the population is constantly increasing, those people living in urban areas prefer to remain in smaller cities.

4.2 Employment Indicators Analysis

The following sections will provide a comprehensive analysis of key employment indicators using a structured approach: it will first being with an exploratory data analysis, aimed at identifying trends and patterns in employment across different sectors and demographics, thus, it will include a correlation analysis, which will examine the relationships between the indicators to understand the how they behave and influence one another. Finally, a forecasting analysis will be conducted to predict future trends in employment distribution, offering insights into potential shifts into the employment sectors. All the indicators and their respective descriptions are outlined in *Table 8.2* of the Appendix in Chapter 8.

Understanding the employment patterns is essential for assessing the structure and evolution of a country's labour market: the indicators shed light on the country's workforce participation, the job distribution and the gender engagement and disparities, across China, India and European Union.

4.2.1 Exploratory Data Analysis of Employment Indicators

The first aspect to be analysed is women's labour force participation, meaning the percentage of employed women of 15+ years of age in each year, as it serves a crucial indicator in terms of the evolution of women's employment dynamics across the different economies. A line graph, illustrated in *Figure 4.15* has been plotted to illustrate how the indicator has evolved over time.

The trends observed in the graph highlight significant regional differences in the extent to which

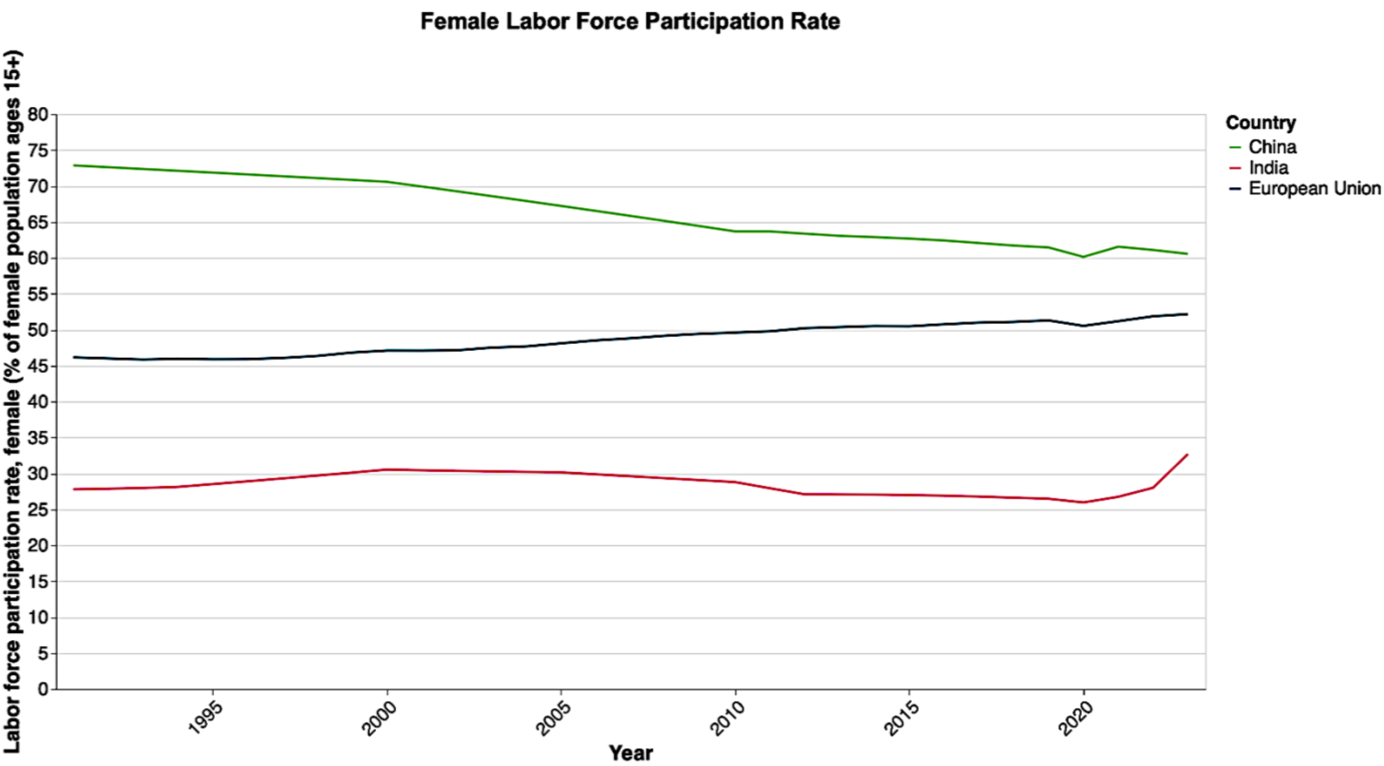

Figure 4.15 - Women Labour Force Participation Rate Over Time

women engage in labour market: China, for example, exhibits the highest share of female labour force participation, remaining consistently above 70%. However, the trend has been gradually decreasing, and in the last few years it seems to have reached plateau, fluctuating around 60%. The EU, on the contrary, shows a relatively stable and slightly increasing trend, hovering around 45%-50%. Finally, India demonstrates the lowest female labour force participation, stagnating around 30% and showing no notable increasing trend.

On the contrary, the trend in unemployment for women follows a different trajectory.

Figure 4.16 illustrates the unemployment as a percentage of the female labour force in China. India and EU from 1991 to 2023: although both China and India have slightly similar trajectories, in the European Union unemployment has been fluctuating significantly in the last years.

In China and India, the rate increases until approximately the beginning of the 2000s, then reaches a

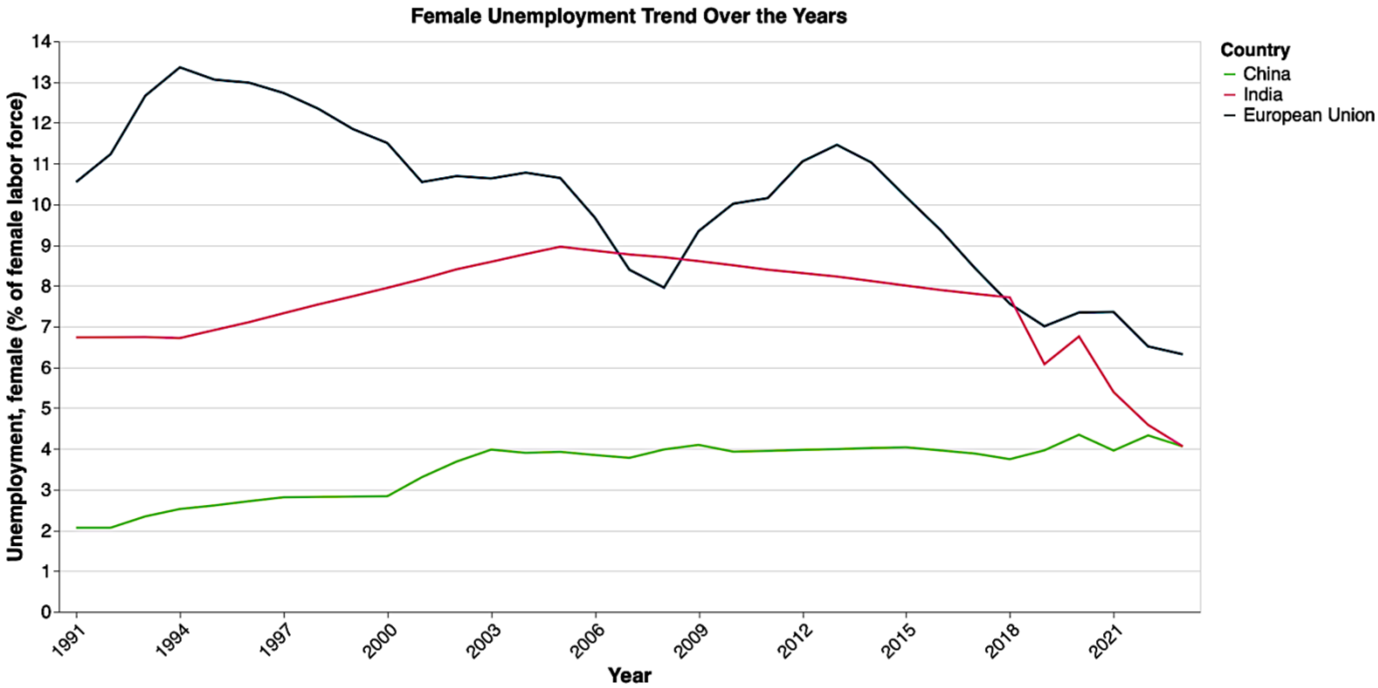

Figure 4.16 - Unemployment in Women Over Time

plateau (in China's case), or decreases (in India's case); in particular, China has the lowest unemployment out of the three economies, and it seems to be staying consistently below 5%. Instead, India has a slightly higher percentage, but in the most recent years there seems to have been a steep decline, as India reached China's level.

Conversely, the European Union exhibits the highest female unemployment rates for most of the observed period, peaking at above 13.5% at the end of the 1990s; it then declined until 2008, where it sharply increased again until 2014. In the last few years, unemployment seems to be declining, reaching its lowest level (slightly above 6%) since at least 1990.

After the general overview on female participation in labour force and unemployment, to provide a clear comparison of the employment indicators and their gender differences a pyramid chart has been implemented: in this visualization, the gender-specific components of each indicator have been stripped to facilitated direct comparison between male and female employment statistics. The chart also includes an interactive element, allowing the selection of a country from a dropdown menu at the bottom, enabling easier cross-economy comparisons.

The values in the charts represent the mean of each employment indicator, offering a concise and clear snapshot of the overall employment landscape and, in order to enhance readability, female values are displayed as negative, thus creating a mirrored effect with the male values. This symmetrical representation allows for a more intuitive comparison of gender disparities across various dimensions.

Figure 4.17 illustrates the values related to China. On average, the female labour force participation rate (around 67%) is slightly lower than the male participation rate (around 80%). Notably, men exhibit a higher employment rate in agriculture, reflecting a stronger dependence on rural labour opportunities, while women are more employed in industry and services. Additionally, both men and women show similar levels of vulnerable employment, which the world bank as the sum of own-

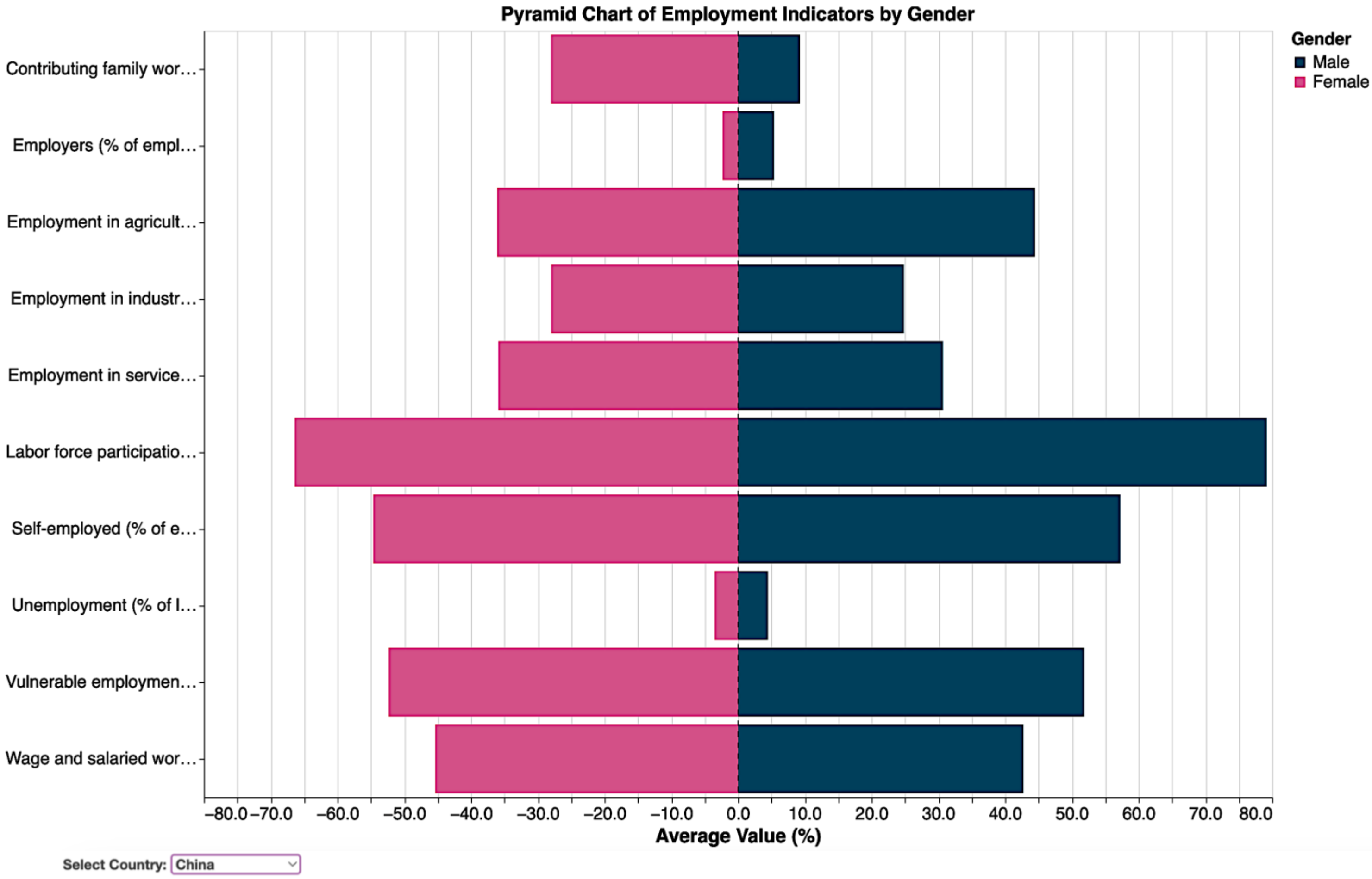

Figure 4.17 - Pyramid Chart of China Employment Indicators

account workers (self-employed workers without employees) and contributing family workers.

Contributing family workers are, as defined by the World Bank, “those workers who hold ‘self-employment jobs’ as own-account workers in a market-oriented establishment operated by a related person living in the same household” and are also known as unpaid family workers. Indeed, these workers generally provide vital labour inputs but do not receive a regular wage and are not considered when making decisions on the family business they are working for. In China, on average, the percentage of female contributing family workers (around 38%) is much higher than that of men (around 8%), indicating that women are more likely to engage in informal family-based labour.

Interestingly, from *Figure 4.17*, it seems that on average more employed women receive wages or salaries compared to employed men, but men are more likely to be self-employed compared to women.

Indeed, out of the total female labour force, only 3% of women are classified as employers, contrarily to the over 5% of employed men, which shows a notable gender gap in leadership and entrepreneurial roles.

Conversely, the employment landscape in the European Union (EU) presents differences to that of China. However, one similarity remains: men continue to dominate labour force participation rates (around 65% compared to that of women of 48%), as shown in *Figure 4.18*.

One of the most striking differences between the EU and China is the significantly lower percentage of agricultural workers in the EU, for both men and women, although male agricultural employment remains quite higher in comparison to that of women. However, the service sector is heavily dominated by women, with approximately 79% of employed women working in services compared

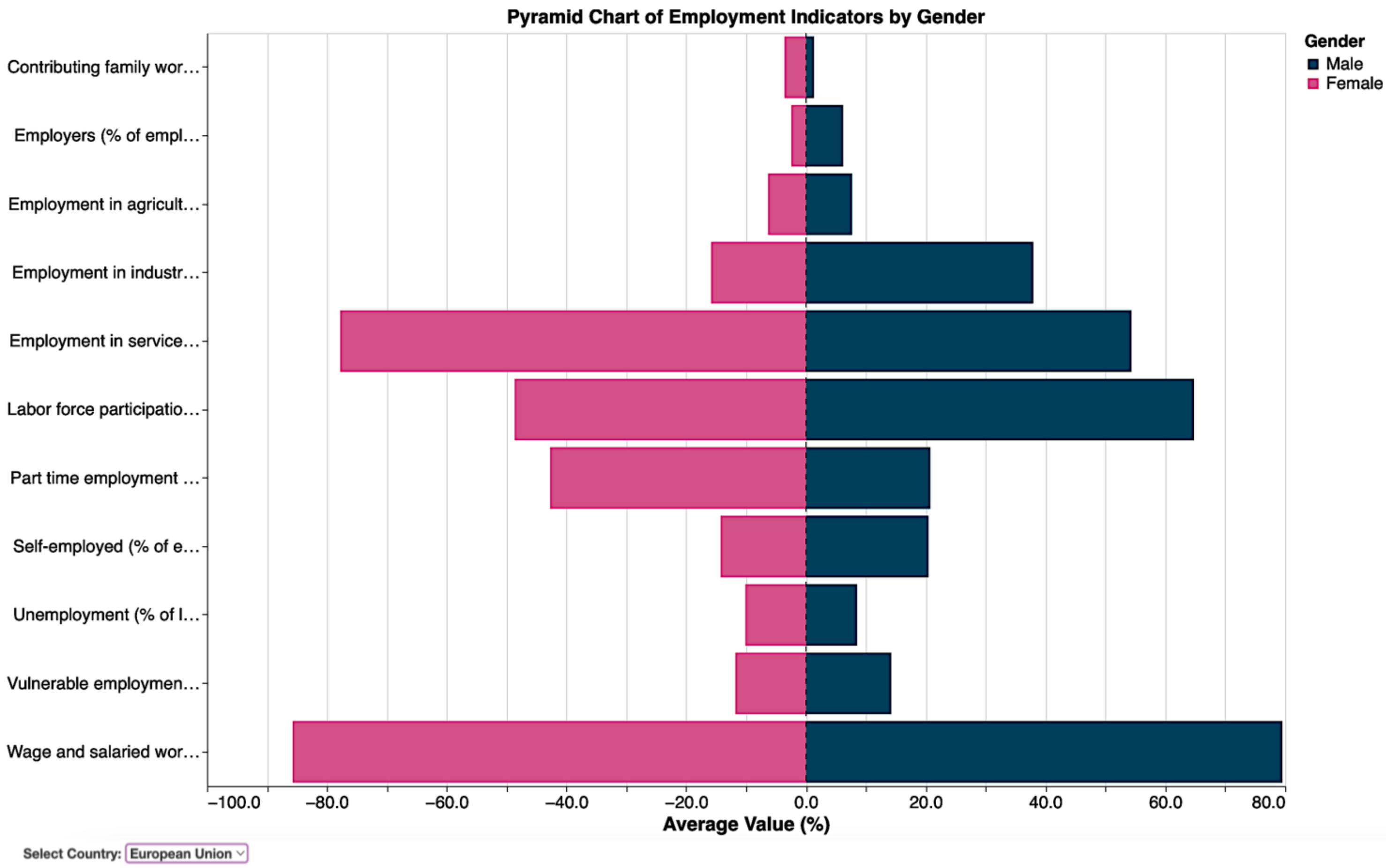

Figure 4.18 - Pyramid Chart of European Union Employment Indicators

to 53% of men; on the other hand, male employers in the industrial sector significantly outnumber female employers, with 38% of male employers compared to just 15% of female.

Another notable trend in the EU is the higher prevalence of part-time employment among women compared to those of men. However, the values for this indicator are absent for China, thus rendering it impossible to provide a comparison.

Furthermore, although the number of contributing family workers is much lower than in China, women in the EU still, on average, have been more likely than men to be in such roles, and, from the indicator *Employers* it also seems that men remain more likely than women to be in an employer position (6% of men compared to 2% of women).

Finally, although the percentage of contributing family workers is much lower than in China, women in the EU are still more likely than men to be in such roles, however, since it seems that on average male workers are more likely to have a vulnerable employment position, it might be more common for men to be own-account workers (as written before, vulnerable employment refers to the percentage of both contributing family workers and own-account workers).

Men are, as for China, more likely to be employers compared to women, but are also on average more likely to have vulnerable employment.

In contrast, India’s labour market follows a very different trajectory from both China and the European Union. The most pronounced disparity is again in labour force participation rates, where only 28% of women have participated on average compared to 78% of men, which highlights the significant barriers to women’s economic engagement during the years.

Moreover, unlike in the EU, where services dominate, or in China, where there is almost a balance across the three employment sectors, Indian women (and men as well) are predominantly employed in agriculture, with much lower engagement in industry or services. Notably, men have been, on average, much less involved in agriculture and much more in industries and services, providing a different trend compared to both the other economies.

Furthermore, self-employment is quite widespread both for men and women, and so is vulnerable employment: indeed, waged and salaried workers are at a much lower percentage compared to China and EU, with male and female salaried workers making up less than 20%. In addition, it is clear that women have worked quite significantly more than men as contributing family workers, possibly indicating that, as vulnerable employment seems to have been equally distributed, men might have more consistently worked as own-account workers. Contributing family workers in India greatly surpass the levels of China and EU, underscoring the challenges that Indian women have faced in accessing formal employment opportunities.

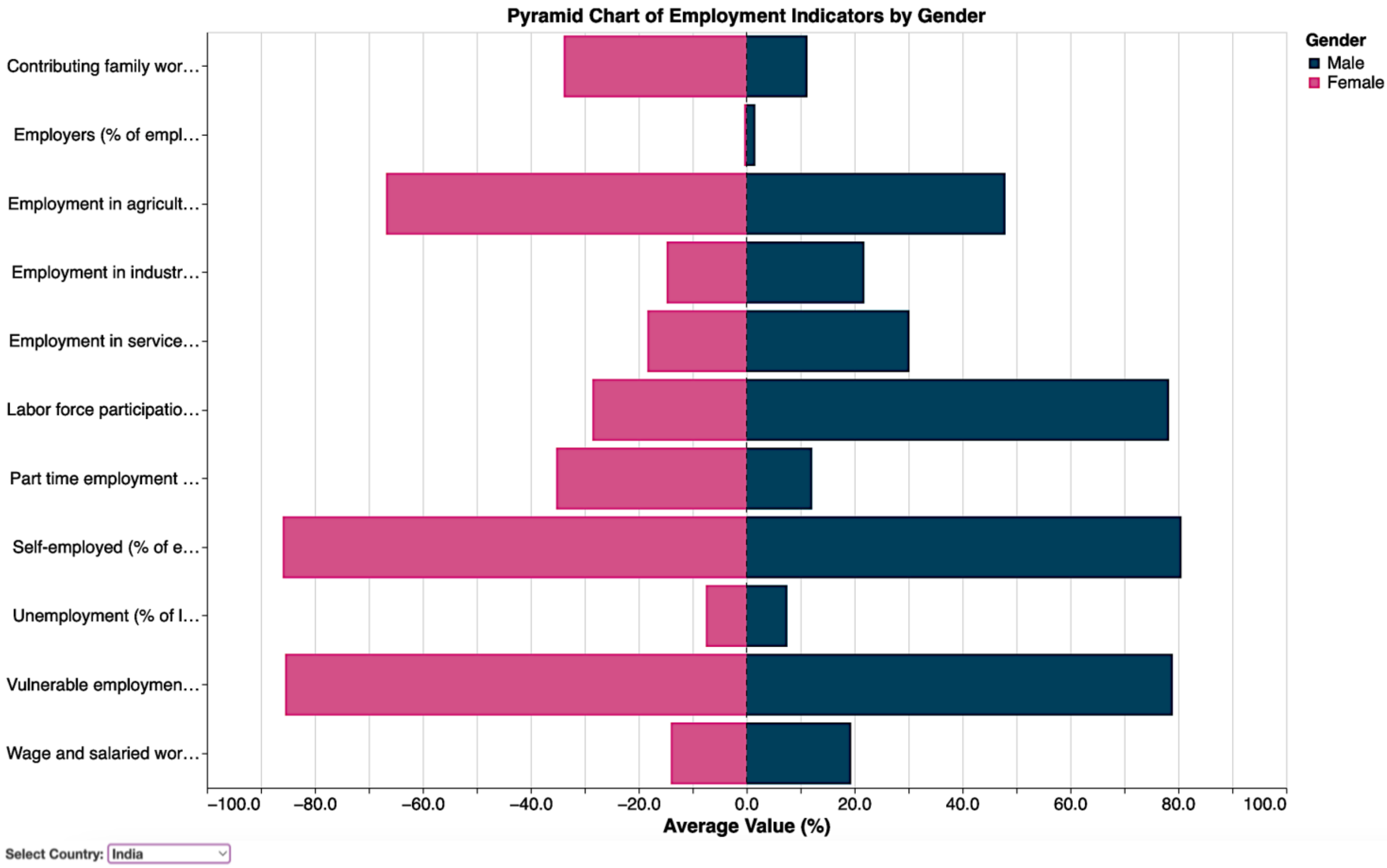

Figure 4.19 - Pyramid Chart of India Employment Indicators

On this note, India also has a much higher part time employment percentage compared to men, even though slightly lower than the one of the European Union.

Finally, the percentage of employers in India is exceptionally low, and among women it is almost non-existent, suggesting that entrepreneurial roles for women have remained extremely rare, possibly due to a combination of socio-economic, cultural, and institutional barriers.

Notably, unemployment rates in the three countries are relatively balanced between men and women, with China and the European Union showing slightly higher unemployment rates for women.

To gain deeper insights into the distribution of employment indicators across the three economies, a box plot analysis has been conducted. This visualization allows for a cleaner understanding of the spread, central tendencies, and potential outliers within the data, offering a more nuanced perfective on the disparities across countries. It is important to note that the time period in analysis in this case ranges from 1991 to 2022, as there was no data before 1990.

Figure 4.20 illustrates the distribution of women's employment across the three primary sectors, agriculture, industry, and services, reinforcing the trends observed above in the pyramid charts.

Among the three economies, India has the highest percentage of women employed in agriculture: the

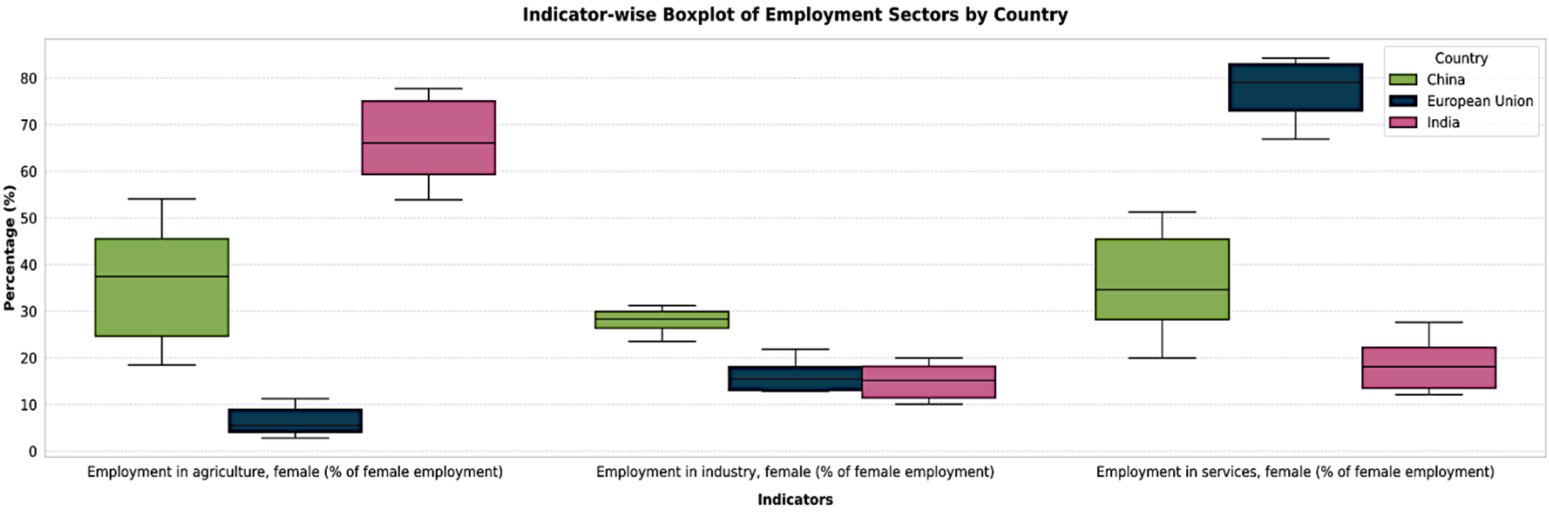

Figure 4.20 - Distribution of Employment Sectors Across Countries

minimum value is slightly above 50%, while the maximum approaches 80% with no outliers, and the median lies beyond 60%, indicating that at least half of the values are concentrated above this threshold. The distribution is positively skewed, meaning that most data points fall within the lower range.

From the comparison of the three boxplots, the contrast between the three economies is very visible: while India's boxplot is on relatively compact and positively skewed, with values ranging primarily between 60% to 75%, the European Union boxplot is exceptionally short, with its values never exceeding 10%, emphasizing that agriculture plays a minimal role in European women's employment. China's distribution falls between the two extremes, with a broader range of values compared to India and EU. These findings confirm again that women in India seem to be much more engaged in agricultural employment, whereas in the European Union, female employment in agriculture is rare.

Conversely, women in the European Union are far more engaged in the services sector: the EU box plot is short and negatively skewed, with values clustering on the higher end of the 70% range. China exhibits a similar pattern to its agricultural sector, although this time the distribution is positively skewed, implying that fewer values exist at the higher end. India, on the other hand, has the lowest representation of women in services, with a short and low-positioned boxplot, which has confirmed what has already been illustrated in *Figure 4.19.*

Interestingly, all three economies show a similar pattern for female employment in industry: although China reports slightly higher female employment in the industrial sector, none of the three countries exceed 35%, and India and the EU even exhibit almost identical distributions. These findings show that both in developed and developing countries, women's employment in industry remains very rare.

Figure 4.21 provides a more detailed view of the sectoral distribution of female employment, reinforcing the trends previously identified in the boxplots and pyramid charts. The line charts have been used to offer a deeper analysis of how women's employment in agriculture, industry and services has evolved over time, allowing for a clearer understanding of long-term shifts; moreover, by

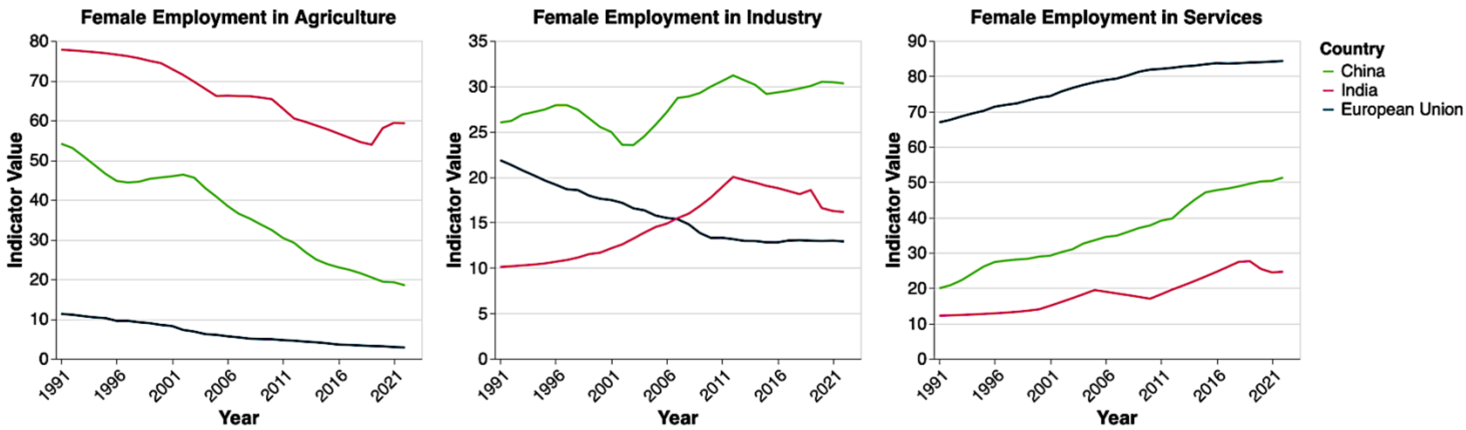

Figure 4.21 - Female Employment in Sectors Through the Years

presenting all three sectors side by side, comparability across sectors and countries is enhanced, making it easier to observe structural changes and assess how each country has transitioned.

Starting from the agricultural sector, as already assessed, India has the highest percentage of women working in agriculture; however, the trend has steadily declined over time, only picking up in the most recent years. This declining trend, combined with the increasing percentages of employment in industry and in services, indicates a gradual transition away from the primary sector and towards the secondary and tertiary, probably due to the advancements in country development.

Notably, the increasing trend of employment in industry and services seems to be decreasing in recent years, which, combined with the slight increase of employment in agriculture that is observed approximately during the same time period, might mean another shift of women labour towards the primary sector.

A similar trend is observed in the other developing nation, China. China has the highest percentage of women employed in industry, with a relatively stable but increasing trend, although the values seem to have reached a plateau after an important increase during the decade from 2001 to 2011. China also has had a steep decrease in women working in agriculture, which was mirrored by an increasing number of women working in services, possibly reflecting the shift to a more developed society.

The European Union continues to lead in women employment in services, and the trend keeps increasing, albeit slowly; in fact, it seems that in recent years the values might have reached a plateau, after surpassing 80%. The share of female employment in agriculture decreases, mirroring the increase in services, and maintains the lowest levels across the three economies, while women working in the industry sector have rapidly decreased until approximately the 2010s, when the values reached a stable level, aligning with the plateau seen in women working in services.

Now focusing on another critical aspect of employment, *Figure 4.22* presents the boxplots for three key indicators related to vulnerable and informal employment across China, EU and India: *Contributing family workers, Vulnerable employment and Self-employed.*

India exhibits the highest values of vulnerable employment and self-employed women, with values

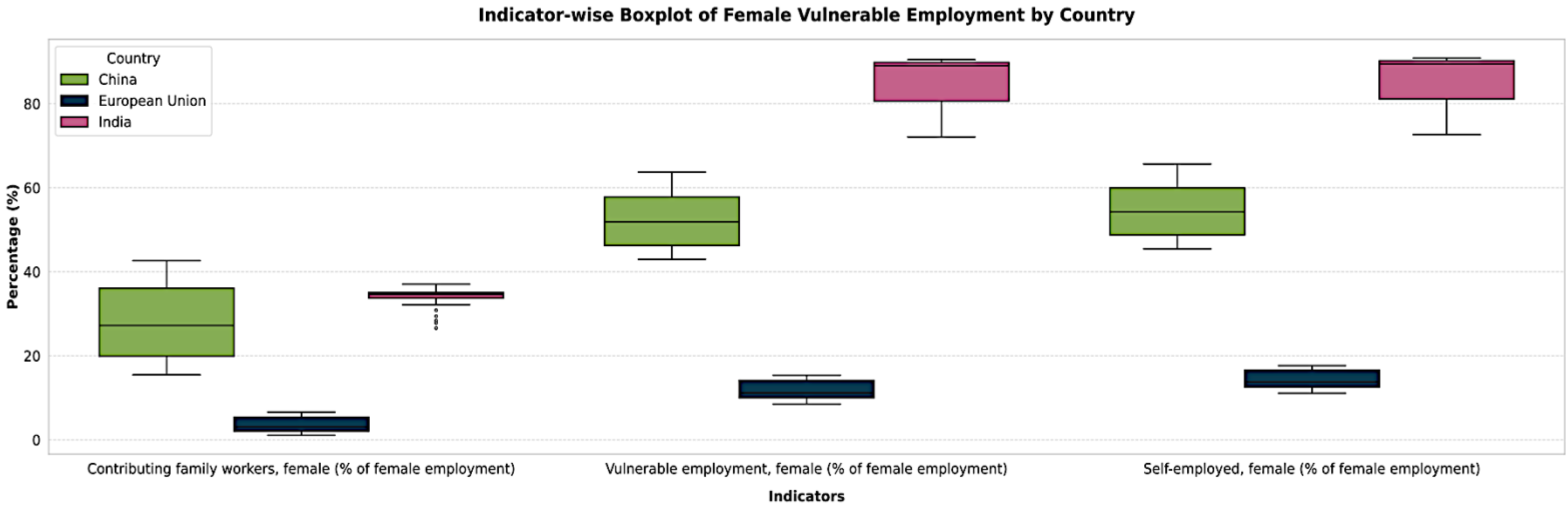

Figure 4.22 - Distribution of Female Vulnerable Employment Across Countries

ranging from approximately 70% to nearly 90%, with no outliers. Both box plots are negatively skewed, suggesting that most values are concentrated on the higher end of the range. Indeed, in all countries, the two indicators seem to follow an exceptionally similar distribution. China has a more spread range of vulnerable employment values, as the median value is around 50%, with a distribution extending slightly above 60%, indicating higher variability during the years.

Contrarily, the European Union has the lowest values of vulnerable employment and self-employment among women, with both indicators remaining below 20% and having a particularly short boxplot suggesting little variation over time and reflecting the greater job stability and formal employment opportunities for women in the EU.

Finally, women as contributing family workers seem to be slightly more common in China with respect to the other two economies: China's boxplot is longer, with values extending up to 40%. However, the the distribution is slightly positively skewed, suggesting a concentration of the values on the lower end. In contrast, India's box plot is particularly short, and its values seem to have a much shorter range, indicating that there have been no major changes to the percentage of women working as contributing family workers over the years, with only minor fluctuations, as it also has a few outliers on the lower end.

The EU, as expected, has consistently low values with a narrow range, indicatingO that no major changes have been reported throughout the years.

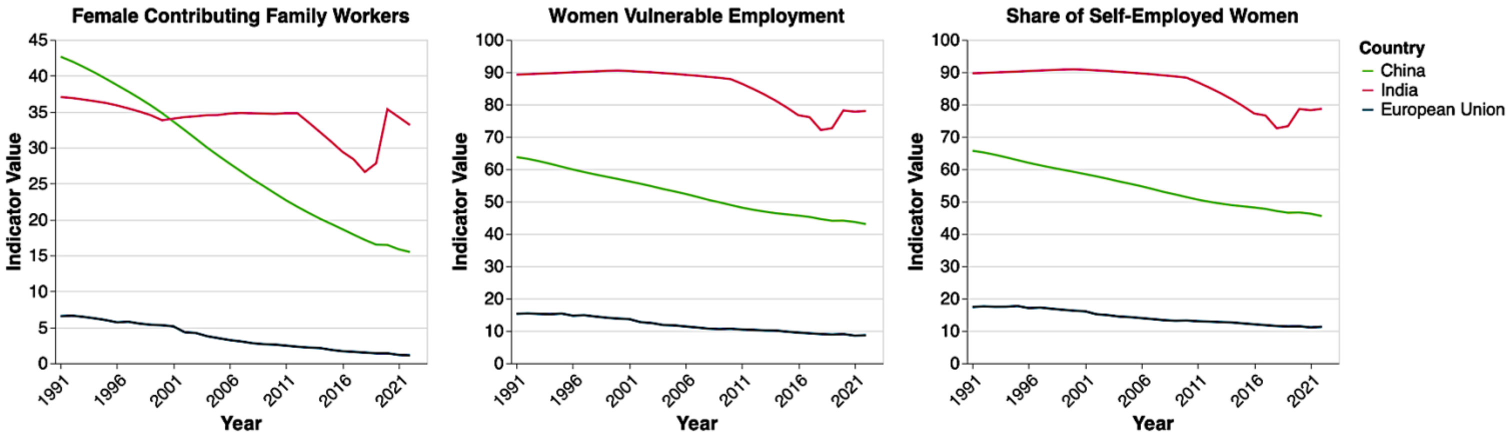

Figure 4.23 - Female Contributing Family Workers (left), Vulnerable Employment (centre) and Self-Employment (right) Across Countries

Building on the previous boxplot analysis, *Figure 4.23* again presents line charts tracking the evolution of these three indicators of vulnerable and informal employment, in order to allow for a more detailed temporal analysis.

India and China started 1990 with a significant proportion of women working as contributing family workers, compared to the European Union; all three economies show a decline over the years, in particular, China presents an exceptionally steep decrease in the share, possibly reflecting the country's rapid economic transformation which increased formal employment opportunities for women. India, on the other hand, although still presenting a slight decrease, has a more fluctuating trend, with periods of rebounds.

It is important to highlight that these statistics are estimates of the ILO statistical models, which, as explained in Chapter 3, are estimates generated through a combination of ILO statistical methodologies. These models are designed to fill gaps in missing data by imputing values based on available labor statistics and trends. As a result, they do not necessarily reflect real-world employment figures with absolute accuracy but instead provide an approximation based on modeled estimations: this is especially visible when comparing the trends of vulnerable employment and self-employment, which despite having slightly different values, somehow have the exact curves in all countries. Because the indicator of vulnerable employment is comprised of both own-account workers (self-employed workers without employees) and contributing family workers, it is entirely possible that the results of the ILO model would have reflected one of the two indicators.

Nevertheless, the trend reflects closely what the box plots in *Figure 4.20* showed: India has the highest percentage of women in a position of vulnerable employment and self-employment, a trend which started as a plateau until approximately 2011, when it began to decrease. It seems to have picked up again in very recent years, while both India and the European Union show a steady decreasing trend. Another interesting aspect that was illustrated in the pyramid chart overview was the disparity of the shared of waged and salaried workers across the three economies, and the comparison between male and female workers. In *Figure 4.22* the evolution of the indicator *Wage and salaried workers* is shown. Again, the data spans the time period from 1990 and 2023 and highlights the key differences

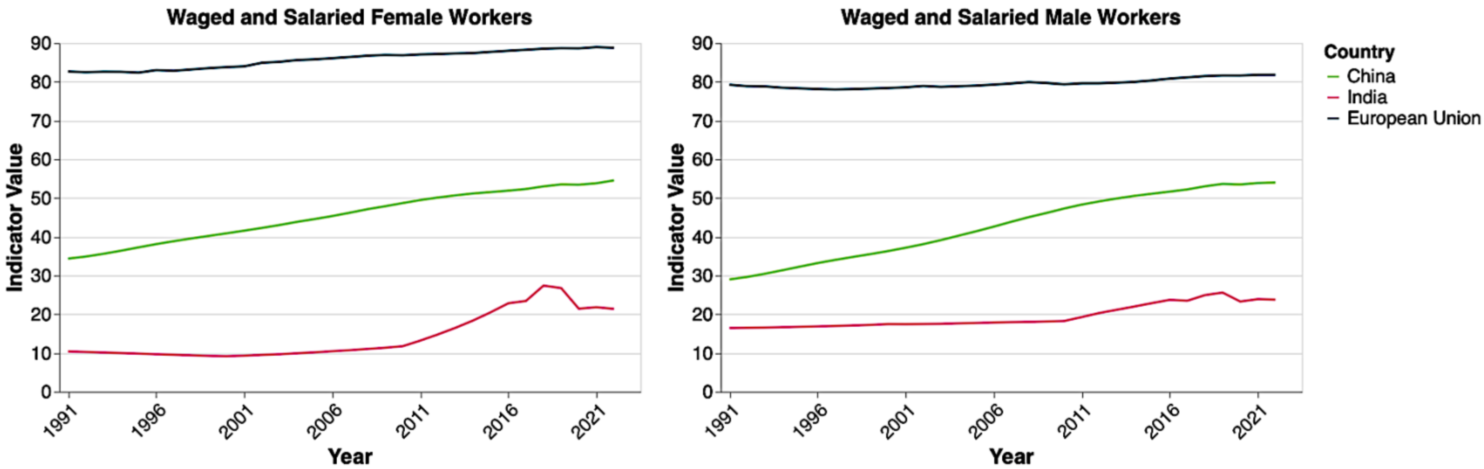

Figure 4.24 - Female Waged and Salaried Workers (left) and Male Waged and Salaried Workers (right)

already shown between the countries: the European Union has steadily had the highest percentage of salaried workers, both in the case of female workers and male workers, however, while there is a slight increase in the share of female workers, no such pattern is visible in male workers' case. Nonetheless, with a percentage of salaried workers that fluctuates around 80%, and women's share being higher than that of men, the European Union shows high employment stability compared to China and India.

China's share has increased for both men and women, reaching almost 60% in both cases, with men's share having increased slightly more than that of women.

Figure 4.25 - Female Employers (left) and Men Employers (right)

Lastly, India exhibits the lowest share of salaried workers, starting at slightly more than 10% in 1991 and reaching approximately around 22% in 2023. While male salaried workers have shown a slow but gradual growth, female salaried employed remained on a plateau around 10% until the 2010s, when it showed a quick increase and reached a peak of almost 30%. However, the trend seems to be decreasing in recent years.

There were other indicators where notable differences between men and women were shown: *Employers* and *Part time employment.*

Figure 4.25 illustrates the evolution of female and male employers as a percentage of total employment in the three economies from 1991 to 2023. The data reveals both gender and country disparities. China and European Union display similar trends regarding women employers: female employers have steadily increased from approximately 2% to more than 2.5%, with the trend in the European Union slightly decreasing in the most recent years. On the other hand, the share of male employers is much higher in comparison, being at its highest in the EU, but exhibits a decreasing curve both in EU and China. India, instead, follows a different trajectory: the share of women employers in India is much lower, and fluctuating around 0.5%, with a slight increase in recent years; male employers are slightly higher in number (starting at approximately 1%), but instead of decreasing as in China and EU, the trend has shown a rapid increase from 2011 onwards.

An important indicator that would have provided valuable insights is part-time employment. However, the availability of data on this aspect is very limited: China lacks any recorded values, and India only has data for the most recent years, making it difficult to analyze long-term trends. In contrast, the European Union has a more comprehensive dataset, allowing for a more detailed examination of part-time employment dynamics between men and women. Unfortunately, due to these data limitations, a meaningful comparison across the three economies was not feasible.

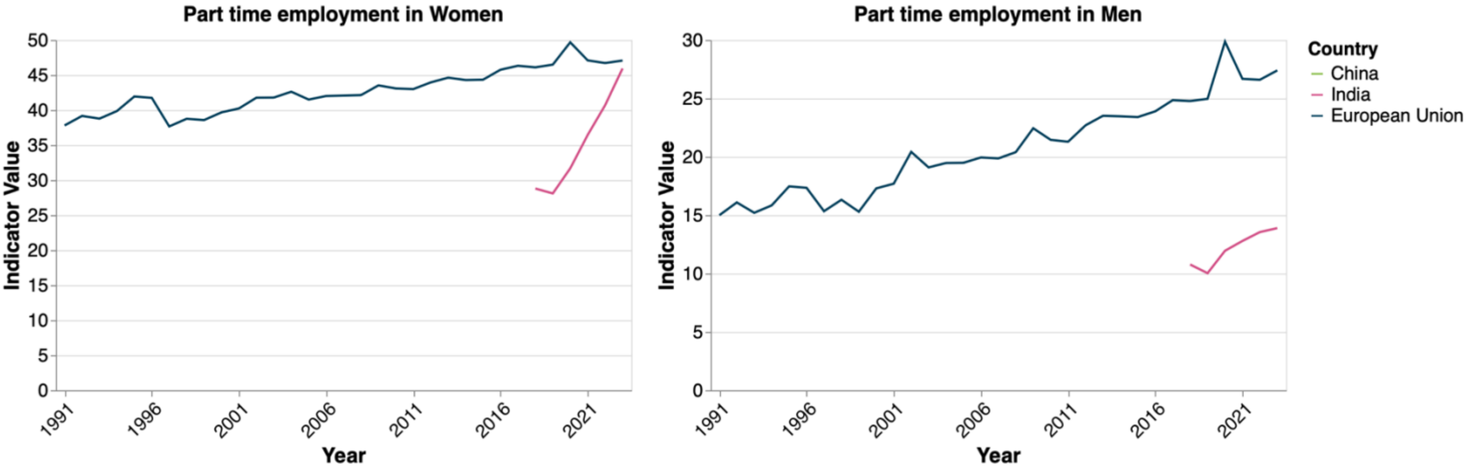

Figure 4.26 - Female Part-Time Employment (left) and Male Part-Time Employment (right)

However, *Figure 4.25* illustrates the available data: on one hand, the EU exhibits a consistently high percentage of women engaged in part-time employment, which has increased to nearly 5% in recent years, while men's highest share of engagement in part time was around 2020, when it peaked at 30%. Both trends seem to be increasing, as does the trend in India, which, despite having a limited availability of data, shows that both the share of women and men in part time is quickly getting higher, especially in women.

4.2.2 Correlation Analysis of Employment Statistics

Understanding the interrelationships between the employment indicators is crucial for identifying patterns and dependencies in labour market dynamics, and correlation analysis provides valuable insights into how different aspects of employment relate and interact with each other.

This section aims to uncover systematic relationships between employment indicators, highlighting key trends that define the labour markets of China, India and the European Union.

For this analysis, the correlation will be conducted using the Pearson Correlation Coefficient, which measures the strength and direction of linear relationships between the indicators, and it will be performed separately for each country to identify country-specific trends and continue the comparison of the employment structures across different economic contexts.

Figure 4.27 illustrates a heatmap which displays the Pearson Correlation among the female employment indicators in China, with the darker red tones indicating positive correlations and the blue tones indicating negative correlations.

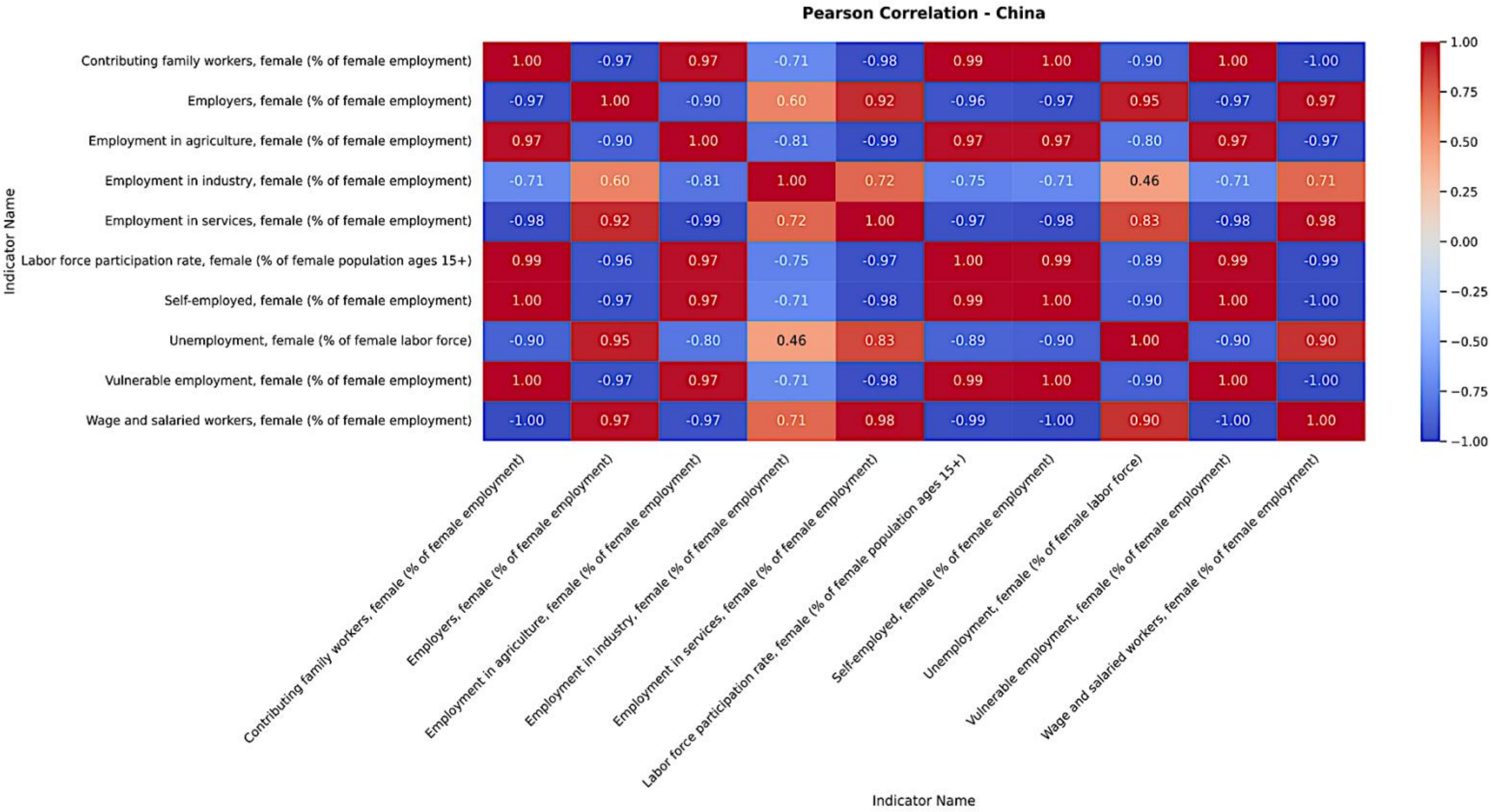

Figure 4.27 – Pearson Correlation of Employment Indicators in China

Regarding the negative correlations, there are some perfectly negative correlations (coefficient = -1) between the indicators *Wage and salaried workers* and *Contributing family workers*, *Self-employed*, and *Vulnerable employment.* Indeed, when plotting the scatterplot of these indicators and the *Wage and salaried workers* indicator, it shows a perfect linear relationship, as illustrated in *Figure 4.28*, where the scatterplot of *Wage and salaried workers* and *Vulnerable employed* is shown.

Perfectly negative or positive correlations are exceptionally rare in real-world data, therefore, the presence of such strong relationships in the dataset may be influenced by the fact that the data originates from ILO model estimates, rather than direct national reporting. Because these models are designed to ensure consistency and comparability, they might introduce artificially strong correlations due to the way missing data is imputed and trends are smoothed.

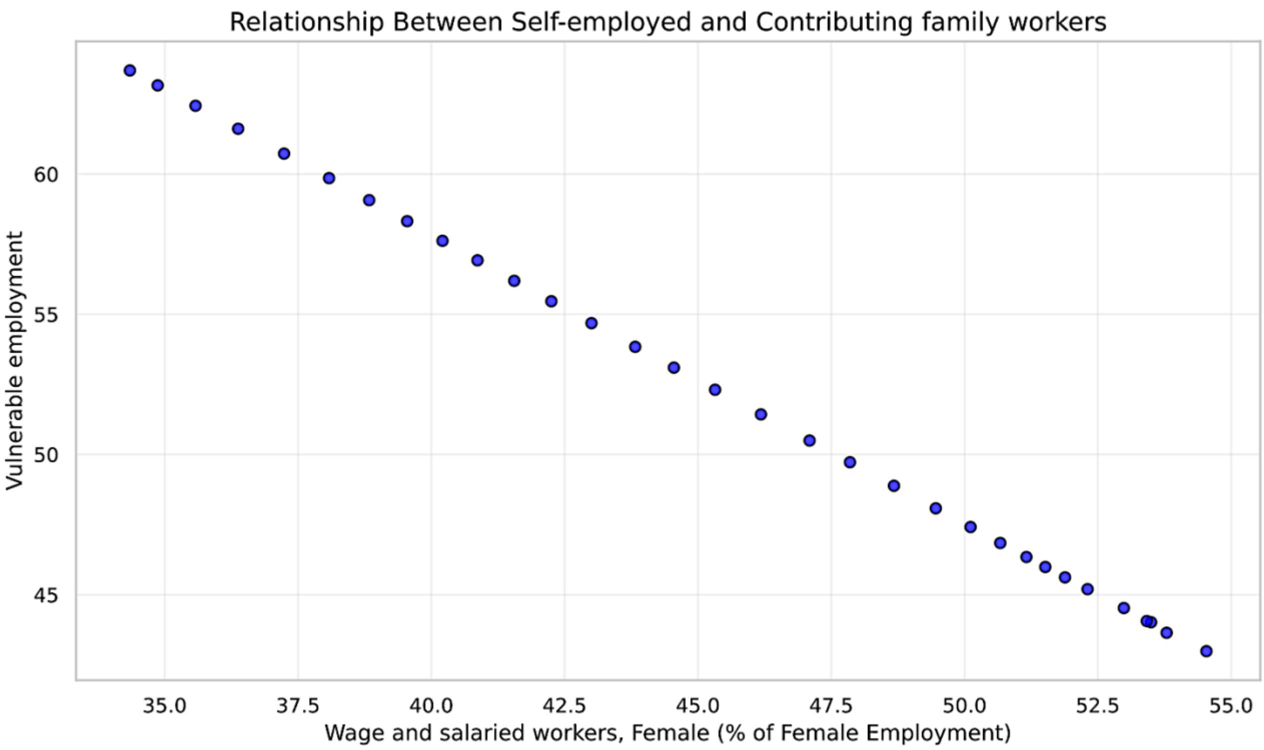

Figure 4.28 – Scatter Plot of Vulnerable Employment and Wage and Salaried Workers

Figure 4.28 is only an example, but there are other cases in which the correlation is perfectly negative or perfectly positive: for instance, *Vulnerable employment* is also perfectly positively correlated with *Contributing family workers* and with *Self-employed.*

However, there are also other particularly high correlations shown: some interesting ones include *Self-employed* being positively correlated to *Employment in agriculture* (0.97) but being negatively correlated to *Employers* (-0.97) and *Employment in services* (-0.98). This indicates that the share of self-employers and that of employment in agriculture increase together, but the trend is reversed when compared to the share of employers and of women employed in services. The negative correlation between self-employed women and women employers can be explained through the structural differences in these employment categories, as self-employment means that an individual works independently, while being an employer requires hiring employees, thus resulting in being negatively correlated with each other.

Looking at some other indicators, *Waged and salaried workers* has some interesting correlations with employment in three sectors: it is highly negatively correlated with the share of employment in agriculture (-0.97), but highly correlated with employment in industries (0.71) and, most importantly, in services (0.98), suggesting that when employees shift from the agricultural sector to the others, the percentage of salaried workers also increases. This is possibly due to increasing levels of formal employment in the secondary and tertiary sectors, compared to the primary sector.

The indicator of *Unemployment* also has interesting correlations with the employment sectors, as it is highly negatively correlated with the share of women working in agriculture (-0.80), while it is positively correlated with the share in services (0.83). It is also negatively correlated with *Vulnerable employment* (-0.90), meaning that as the levels of unemployment increase, the levels of vulnerable

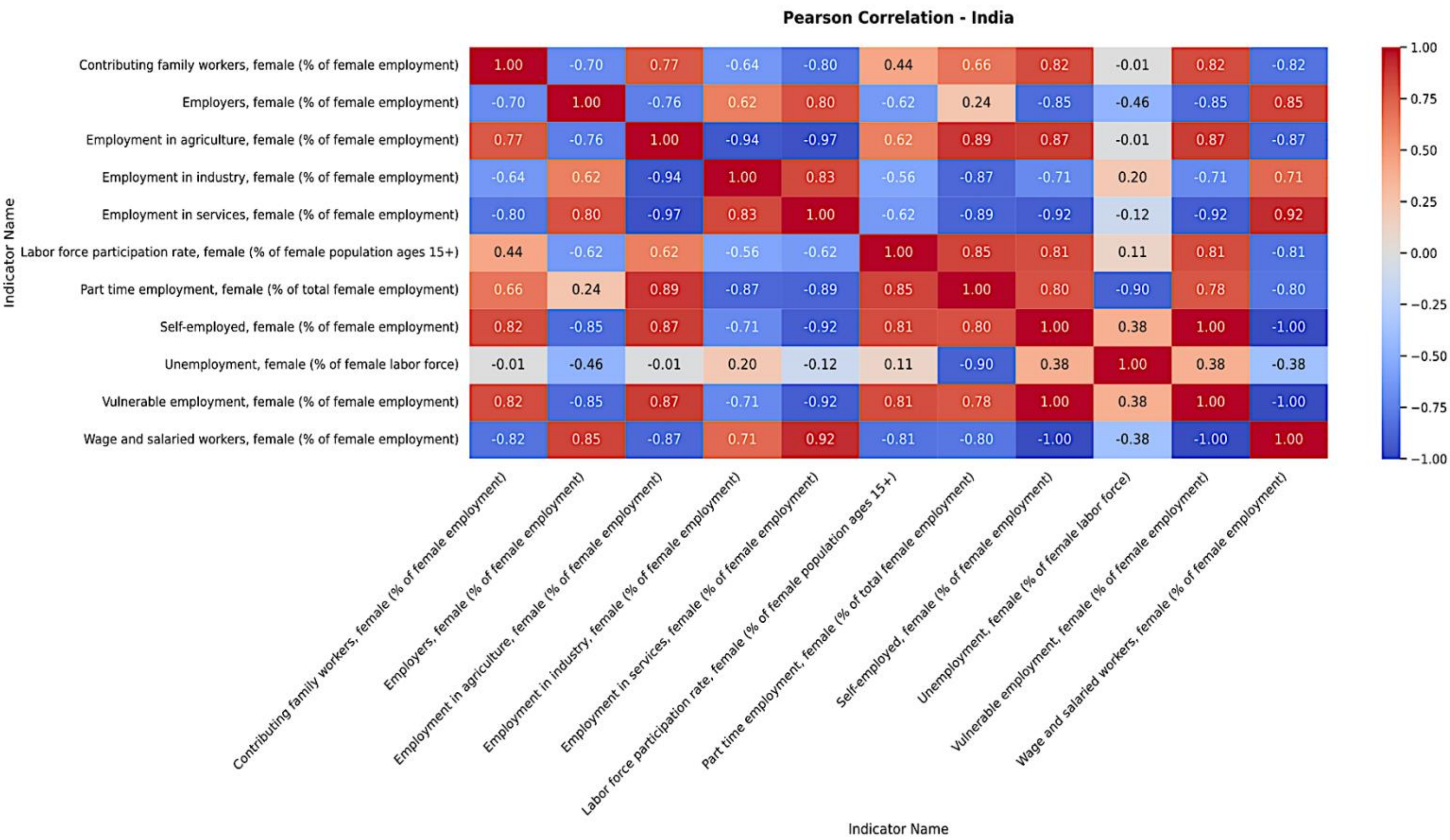

Figure 4.29 - Pearson Correlation of Employment Indicators in India

employment decrease, possibly due to the desperate need of people to work for a living, which brings them to look for uncertain employment when formal employment is not available.

Following the correlation analysis for China, *Figure 4.29* presents the Pearson Correlation matrix for India. Just by looking at the different colours, it shows that the correlations are quite different from those of China: to start, there are only two indicators which have a perfectly positive correlation, and those are *Vulnerable employment* and *Self-employed.* Again, it makes sense that those indicators have such a high correlation, considering that self-employment is considered to be part of the vulnerable employment indicator.

Nonetheless, despite having slightly lower correlations, many relationships seem to have remained similar to China. For example, *Vulnerable employment* maintains the same types of correlations with the three employment sectors, exhibiting a high positive correlation with the share of employment in agriculture, and a high negative correlation in the two other sectors. Moreover, the *Employment in agriculture* indicator has a high negative correlation with both the two other employment sectors, suggesting that people mostly tend to shift from the first sector to the other two, or the contrary. The three sectors also maintain the same correlation as China with the share of self-employed women, while *Unemployment,* this time, shows very little correlation with most indicators.

Finally, *Figure 4.30* shows the correlation analysis for the European Union. Different trends compared to the other two economies are shown in the figure: the share of salaried

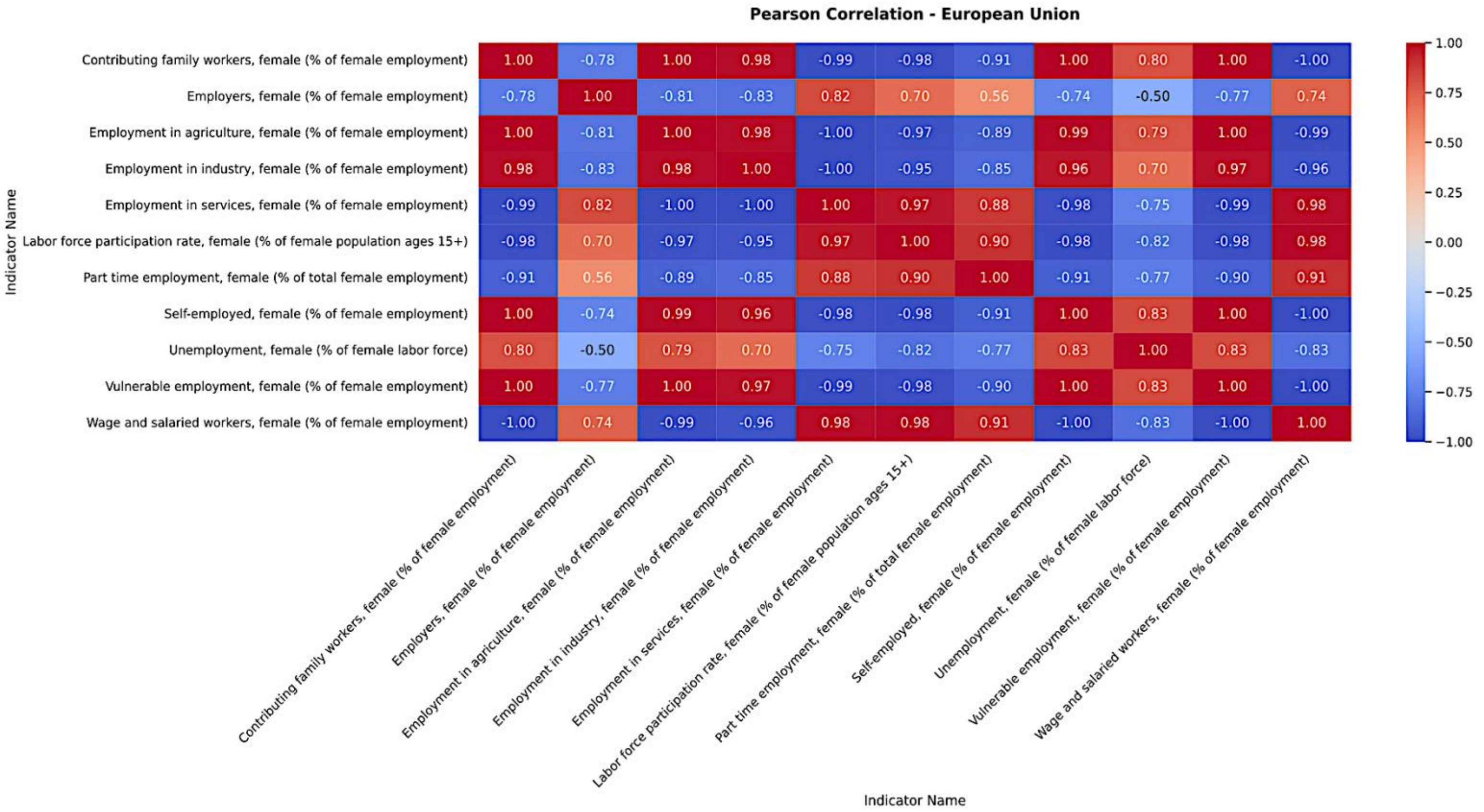

Figure 4.30 – Pearson Correlation of Employment Indicators - EU

workers has a perfectly negative correlation to the share of contributing family workers, which is explained by the fact that contributing family work is most often considered unpaid work. Moreover, *Contributing family workers* has a perfect positive correlation with *Vulnerable employment*, but also with the share of employment in agriculture and with self-employed. Furthermore, *Wage and salaried workers* has a highly negative correlation with the employment in industry that is quite different from the correlation in the China and India, which was positive. Another difference is that unemployment is highly correlated with the employment in agriculture and industry but is negatively correlated with employment in services. This slightly differs from China, as the correlation with agricultural employment was highly negative, and with services highly positive, but it also differs from India, where it seemed there was no correlation between unemployment and these indicators.

4.2.3 The Future of Women in the Employment Sectors

Understanding past and present trends is crucial, but forecasting future patterns provides valuable insights for policymakers, businesses and researchers. In this section, a time-series forecasting is presented using the ARIMA method to predict the evolution of female employment in the different sectors across China, India and European Union, by analysing historical data and extending the projections for the next 10 years. These forecasts offer a data-driven perspective on the expected trajectory of women's participation in the labour market: given the structural differences in employment patterns across the three economies, the forecasts are divided by country, allowing for a nuanced comparison of potential shifts in the labour dynamics.

By applying the ARIMA forecasting model, the analysis will help answer key questions about the future of female employment labour participation, and potential disparities between developed and developing countries.

To ensure the accuracy and reliability of the employment forecasts, a systematic approach has been followed to determine the optimal parameters for the ARIMA model. This process has been applied consistently across all employment indicators and countries; however, for the sake of simplicity and readability, the detailed step-by-step explanation will only be provided for the first forecasted model, with subsequent models following the same methodology.

Starting from the future of women employed in agriculture in China, the first step is the stationarity check, using the Augmented Dickey-Fuller (ADF) test. The ADF test gives a test statistic and a p-value: the test statistic (ADF Statistic in *Table 4.1*) is compared to the critical values at different significance levels (usually at 1%, 5% and 10%), and, if it is more negative than the critical value then the null hypothesis that the data is non-stationary can be rejected, otherwise it cannot. At the same time, it also prints the p-value, which, if lower than the significance level of 0.05, then allows to reject the null hypothesis.

ADF Statistic	-0.79
p-value	0.82
Critical Value (1%)	-3.67
Critical Value (5%)	-2.96
Critical Value (10%)	-2.62

Table 4.1 - Results of ADF for Agriculture in China

When tested, the results provided are shown in *Table 4.1*. The results indicate that the null hypothesis of non-stationarity cannot be rejected. The ADF statistic is -0.797, which is greater than the critical values (-3.670, -2.96, -2.62), indicating that it fails to reject the null hypothesis that the data is non-stationary. The same is confirmed with the p-value, much higher than the significance level of 0.05.

This confirms that differencing is needed before fitting an ARIMA model. To difference a time series, as explained in Chapter 3, means to transform non-stationary data (data where there is a trend in the variance of the data over time) to stationary (the variance is constant over a period of time). In general, the first order differencing is applied to the data (formula of which can be found at the link: Equation_1), then, if the data has not reached non-stationarity yet (again tested through the ADF test),

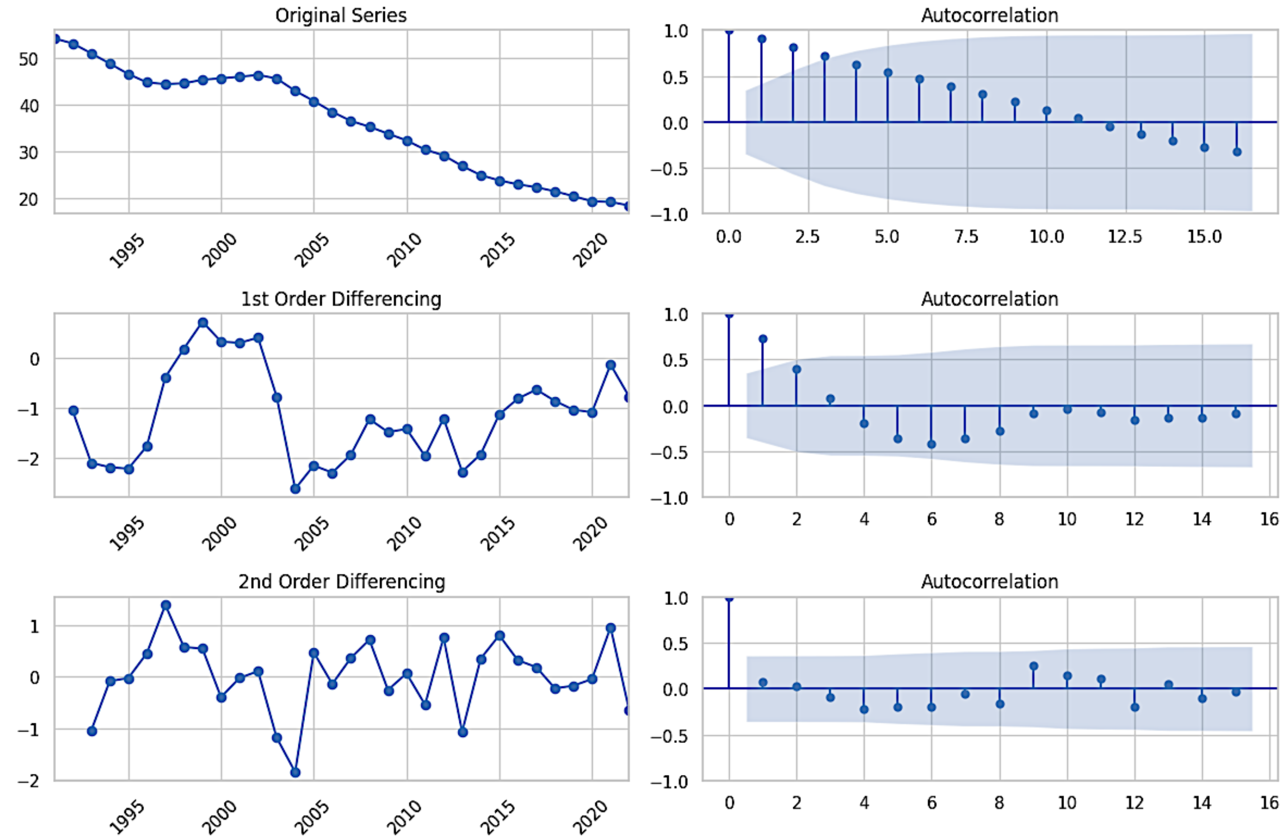

Figure 4.31 - Differencing of Employment in Agriculture in China

a second order differencing is applied (complete formula at the link: Equation_2). Differencing the data is needed because non-stationary data does not have a predictable behavior, thus resulting in incorrect predictions since the past observations may not be representative of future ones. *Figure 4.31* shows the 1st and 2nd order differencing and the autocorrelation plots derived from the *Employment in agriculture* in China. It seems that the time series reaches for stationarity with two orders of differencing, as it still shows a pattern in its autocorrelation after the first order. Moreover, the high autocorrelation at multiple lags suggests strong dependence of the data over time, indeed, autocorrelation represents the relationship between an indicator's present value and its past ones, and in stationary time series it is essentially constant. In this case, after the first differencing, the autocorrelation still has a slow decrease, indicating that the data is not stationary yet.

ADF Statistic	-2.08
p-value	0.25
Critical Value (1%)	-3.67
Critical Value (5%)	-2.96
Critical Value (10%)	-2.62

Table 4.2 – ADF Results After 1st Order Differencing

A second order differencing is thus computed. It is clear in the graph that no trend has remained, and the autocorrelation shows that all lags now have very low autocorrelation values. Although the second differencing has improved the situation, it might not be as necessary; indeed, the first-order differencing series was very close to stationarity, suggesting that the d parameter should be d = 1 instead of d = 2. However,

ADF Statistic	-4.95
p-value	2.68e-05
Critical Value (1%)	-3.67
Critical Value (5%)	-2.96
Critical Value (10%)	-2.62

Table 4.3 – ADF Results After 2nd Order Differencing

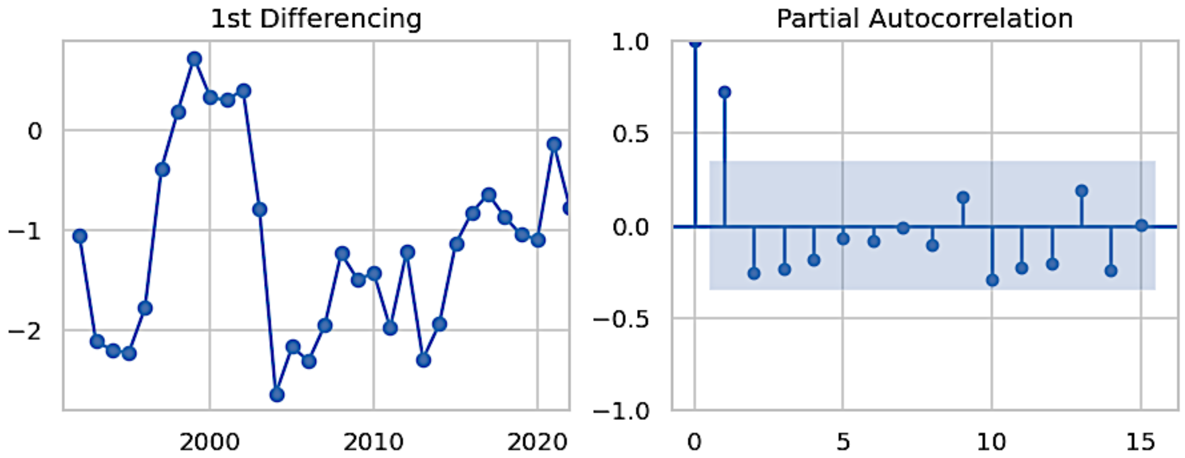

Figure 4.32 - Partial Autocorrelation Plot of Employment in Agriculture in China

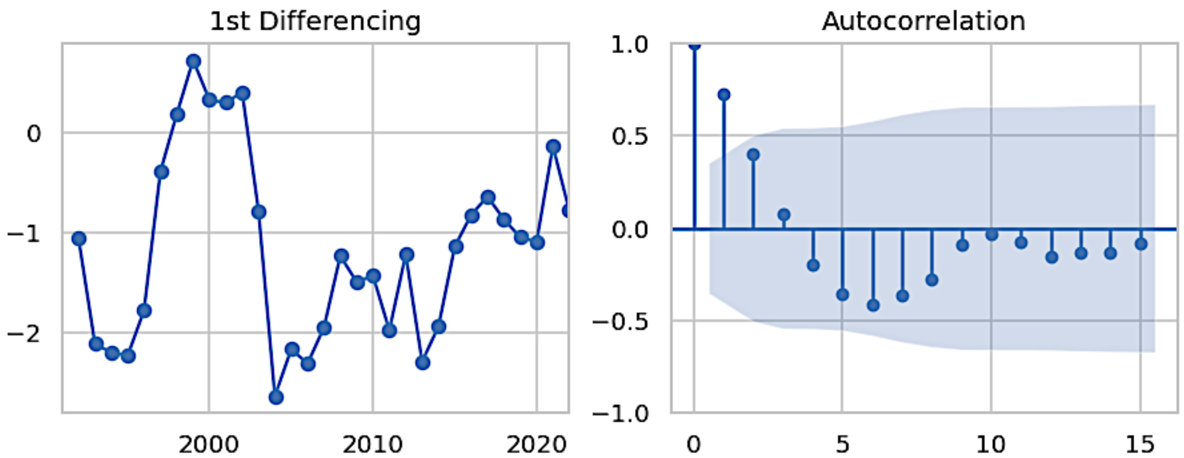

Figure 4.33 - Autocorrelation Plot of Employment in Agriculture in China

in order to confirm it, another ADF test will be computed on the first and second order of differencing. The results of the second ADF test after the 1st and 2nd order differencing are shown in *Table 4.2* and *Table 4.3*.
It seems that the second differencing was, indeed, needed, since the p-value was not lower than the significance value after the 1st differencing. Moreover, the ADF statistics is much lower than the critical values. It is then confirmed that the second differencing is needed and that $d = 2$.

The model now requires the order of the AR term (p).
The partial autocorrelation plot of the differenced series is therefore shown in *Figure 4.32*.
It is observed from the plot that two lags are above the significance line, although the p will be set at $p = 1$ at the beginning in order to be slightly more conservative.
Finally, as with the partial autocorrelation, the q (MA term) parameter can be derived from the autocorrelation plot of the differenced series, which is shown in *Figure 4.33.*
Again, the first two lags are shown to be above the limits, but the first lag is evidently much higher, therefore setting $q = 1$.

Therefore, the final *ARIMA(p,d,q)* model will have the following parameters: *ARIMA(1,2,1)*.
After training the model on the training set, from the results statistics reported in *Figure 4.34*, it seems that all autoregressive (AR) and moving average (MA) terms have high p-values, meaning they are not statistically significant. This suggests that this model might not be the best fit as it is probable that

```
                               SARIMAX Results
==============================================================================
Dep. Variable:        Indicator Value   No. Observations:                   32
Model:                 ARIMA(1, 2, 1)   Log Likelihood                 -31.241
Date:                Tue, 11 Feb 2025   AIC                             68.481
Time:                        14:33:45   BIC                             72.685
Sample:                             0   HQIC                            69.826
                                 - 32
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1          0.1515      3.682      0.041      0.967      -7.064       7.367
ma.L1         -0.0708      3.692     -0.019      0.985      -7.306       7.164
sigma2         0.4698      0.139      3.388      0.001       0.198       0.742
===================================================================================
Ljung-Box (L1) (Q):                   0.00   Jarque-Bera (JB):                 1.55
Prob(Q):                              0.97   Prob(JB):                         0.46
Heteroskedasticity (H):               0.97   Skew:                            -0.55
Prob(H) (two-sided):                  0.96   Kurtosis:                         3.11
===================================================================================
```

Figure 4.34 - Results of ARIMA(1,2,1)

```
                               SARIMAX Results
==============================================================================
Dep. Variable:        Indicator Value   No. Observations:                   32
Model:                 ARIMA(1, 1, 0)   Log Likelihood                 -31.993
Date:                Tue, 11 Feb 2025   AIC                             67.986
Time:                        14:33:49   BIC                             70.854
Sample:                             0   HQIC                            68.921
                                 - 32
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1          0.8785      0.107      8.245      0.000       0.670       1.087
sigma2         0.4398      0.096      4.557      0.000       0.251       0.629
===================================================================================
Ljung-Box (L1) (Q):                   0.55   Jarque-Bera (JB):                 2.04
Prob(Q):                              0.46   Prob(JB):                         0.36
Heteroskedasticity (H):               0.80   Skew:                            -0.57
Prob(H) (two-sided):                  0.73   Kurtosis:                         3.55
===================================================================================
```

Figure 4.35 - Results of ARIMA(1,1,0)

the AR/MA terms might be unnecessary for the model. Therefore, the next step is to fine tune the parameters and try a simpler model such as ARIMA(1,1,0)

In this second model whose results are shown in *Figure 4.35*, the p-values are definitely below the significance level, indicating that the AR term is significant; moreover, the AIC and BIC are also slightly lower than in the first model, suggesting that this model might be better than the previous one, since it is simpler and has no loss of predictive power.

Therefore, the predictions are computed using ARIMA(1,1,0) and are evaluated with the Mean Absolute Error, whose result is 4.32 and Root Mean Squared Error, whose result is 4.37; a residuals

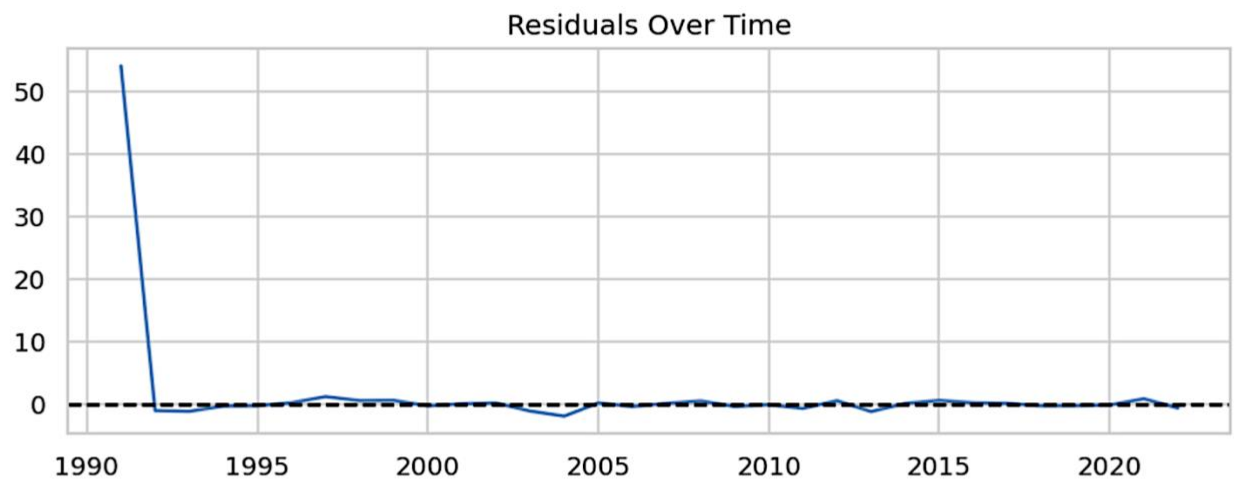

Figure 4.36 - Residuals from ARIMA(1,1,0)

plot is also shown in *Figure 4.36*. It seems that the model’s predictions differ, on average, about 4% from the actual data points, which is further confirmed through the RMSE. Moreover, looking at the residuals plot, it illustrates how the model struggled to fit those first few data points approximately around 1990, although later years show that there are no clear patterns and that the residuals are close to zero, indicating that the model has fitted the later data well.

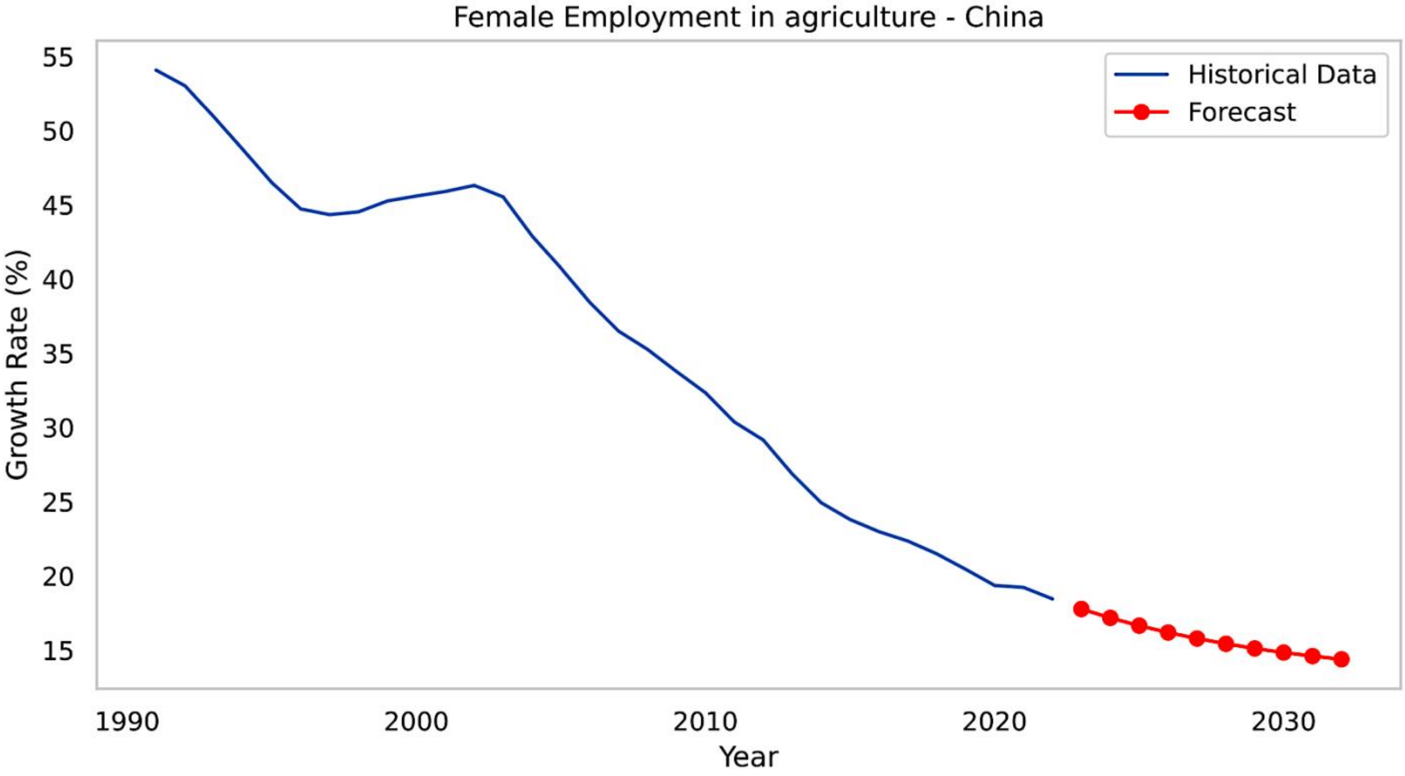

Figure 4.37 - Future Employment of Women in Agriculture in China

Finally, this model is used for the predictions of women's share of employment in agriculture during the next 10 years: in *Figure 4.37*, it shows that the historical data has a consistent downward trend which has been picked up by the ARIMA, which, indeed, has predicted a continuous decline over the next 10 years.

It is important to emphasize that this model generates predictions based solely on historical data trends and does not account for potential future disruptions or unforeseen events that could alter the trajectory of the sector. External factors such as policy changes, technological advancements, economic crises, or environmental challenges could significantly impact employment trends, leading to deviations from the forecasted path.

Figure 4.38 - Future Employment of Women in Industry in China

Further analysis examines the projected evolution of female employment in the industry and services sectors, presenting only the final results to maintain conciseness and avoid unnecessary repetition, although the complete process and results can be found at the linked code.

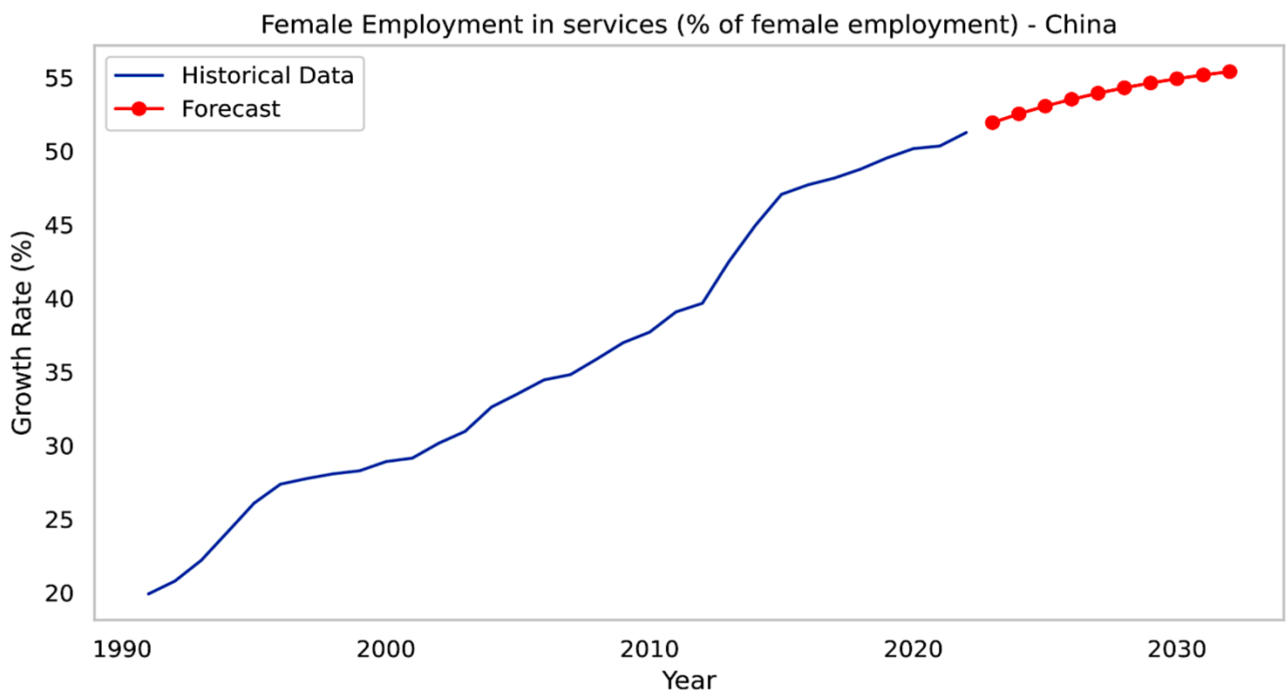

Figure 4.39 - Future Employment of Women in Services in China

Figure 4.40 - Future Employment of Women in Agriculture in India

In *Figure 4.38* and *Figure 4.39*, the forecasted trends of women working in industry and working in services are illustrated: the ARIMA models used are *ARIMA(1,1,1)* for the predictions of employment in services and *ARIMA(1,1,0)* for the predictions of employment in industry.

In *Figure 4.38*, the historical data trend shows substantial growth in female employment in the industrial sector from the early 2000, peaking at approximately 2010; nonetheless, after reaching its highest point, the trend appears to decrease, thus increasing again. The forecast picks up from then

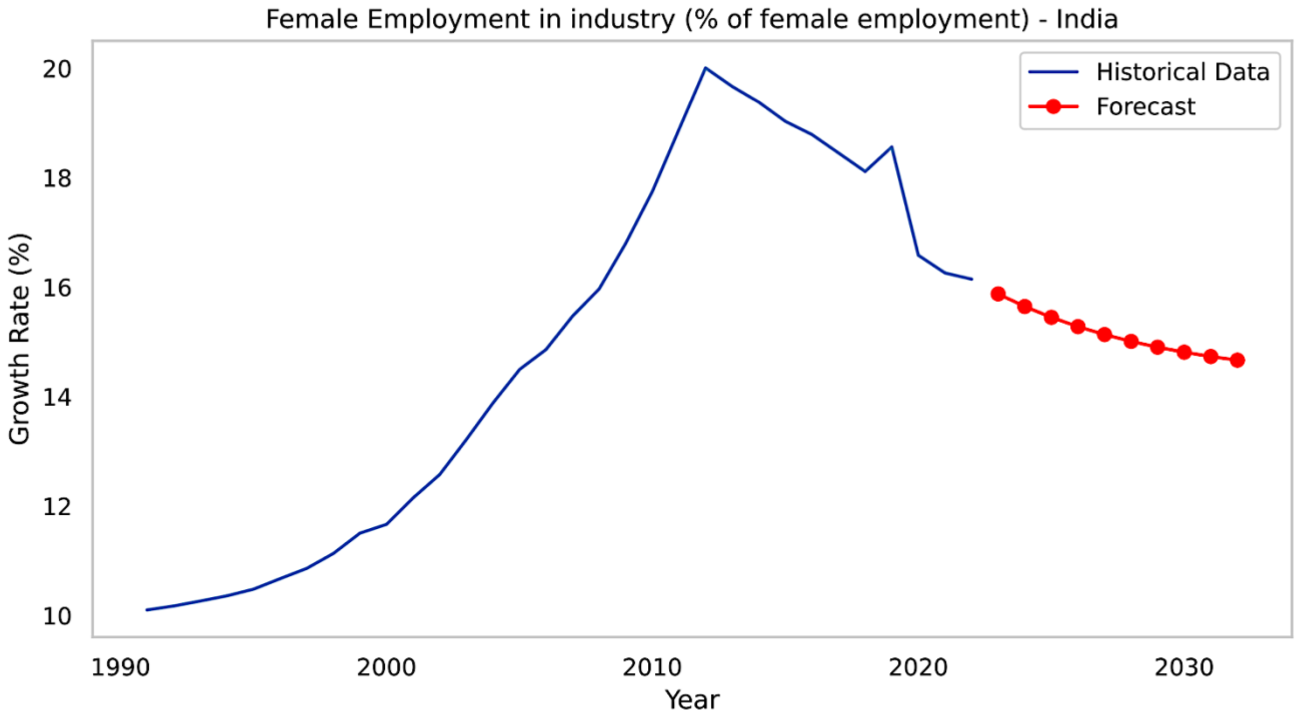

Figure 4.41 - Future Employment of Women in Industry in India

the trend seems to be directing towards stability and suggests that a plateau might be coming in the following years. These predictions might have also been influenced by the fluctuations observed in the historical data, which may have introduced variability in the forecast.

In the service sector (*Figure 4.39)*, instead, the growth is shown to be steady and continuous over the years, with a particularly sharp rise since the 2000s. The forecasted values continue the upward trajectory, albeit at a slower pace, indicating a good fit of the model on the data.

The next country in analysis is India: again, only the results of the forecasts will be presented to maintain conciseness and provide insights into possible future trends.

India's forecasts are presented in *Figure 4.40, Figure 4.41* and *Figure 4.42.*

For what regards women in agriculture (*ARIMA(1,1,1)*), the historical data shows a clear declining trend in female employment in the sector over the years, with an increase in the late 2010s: the percentage of women working in agriculture declined from over 75% to approximately 55% in 2020, but then increased again in the most recent years, which is where the forecast picks up. The predictions suggest that the trend will stabilize in the short term, rather than increase: indeed, the data does show that a stable period might be starting. However, it seems unlikely that the actual trend will respect the predictions as they are too stable to mirror reality and do not really represent what is shown to be the historical behavior of the data. It is probable that, while some stability might indeed happen in the next few years, it will fluctuate again in the short term.

Figure 4.42 - Future Employment of Women in Services in India

On the other hand, *Figure 4.41* depicts a fluctuating but increasing historical trend in female employment in the industrial sector, which peaked at approximately 20% in the early 2010s and then started quickly decreasing.

The predictions, based on *ARIMA(1,1,0)*, indicate a continuation of the decreasing trend, but also show a very light increase at the beginning of 2030; it seems that the model had a better fit for this indicator than the previous one, and therefore managed to predict a more realistic trend.

Finally, in *Figure 4.42*, the most dynamic sector of female employment is depicted. Historically, the percentage of women in services has grown significantly, reaching over 26% in recent years; due to the recent decrease, the forecasts started at a time of uncertainty on whether the trend will increase again or stabilize. As what was shown for the share of employment in agriculture, the predictions (*ARIMA(1,1,1)*) present a picture of stability in the following years, which is probably unlikely to happen as the data shows frequent fluctuations.

Following the predictions for China and India, this next paragraph presents the ARIMA-based forecasts for the engagement of women in the three sectors in the European Union.

Figure 4.43 shows a continuous decline in the share of women working in agriculture: it is pretty straightforward that the forecast, predicted with a model *ARIMA(1,1,1)* will show a continuation of this trend, reaching less than 2% of women working in agriculture after 2030. Again, it is debatable

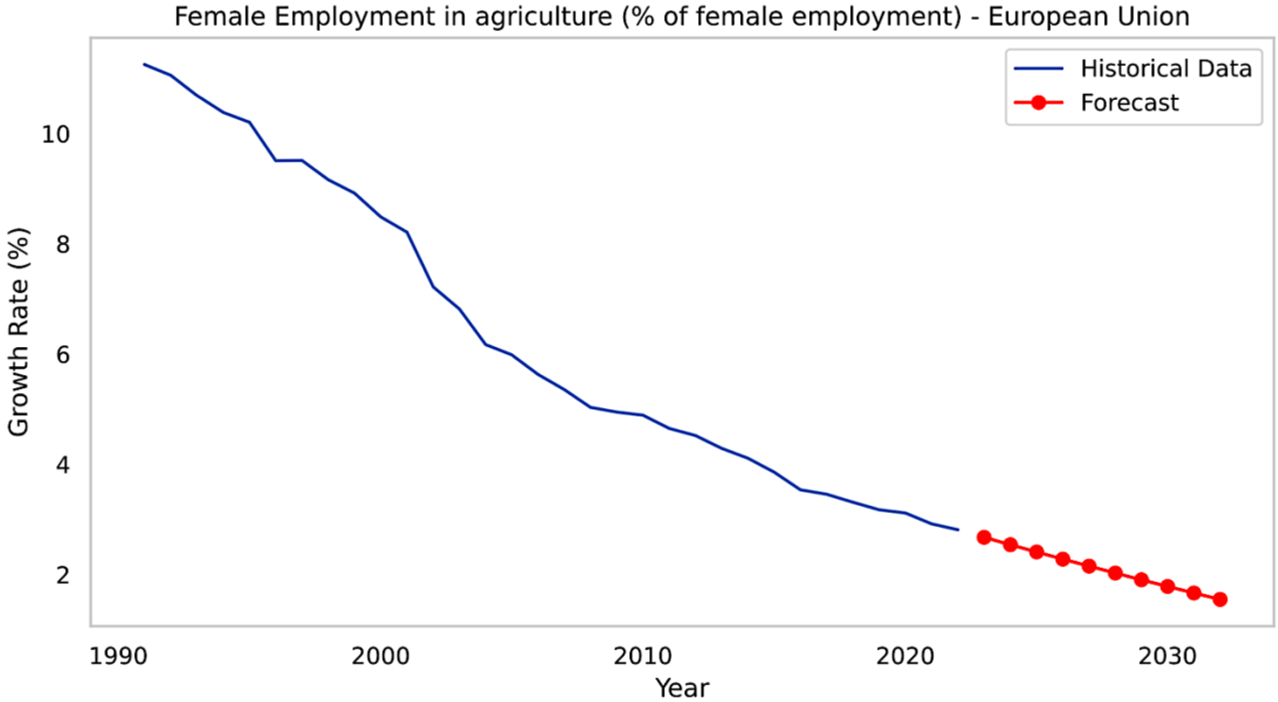

Figure 4.43 - Future Employment of Women in Agriculture in the EU

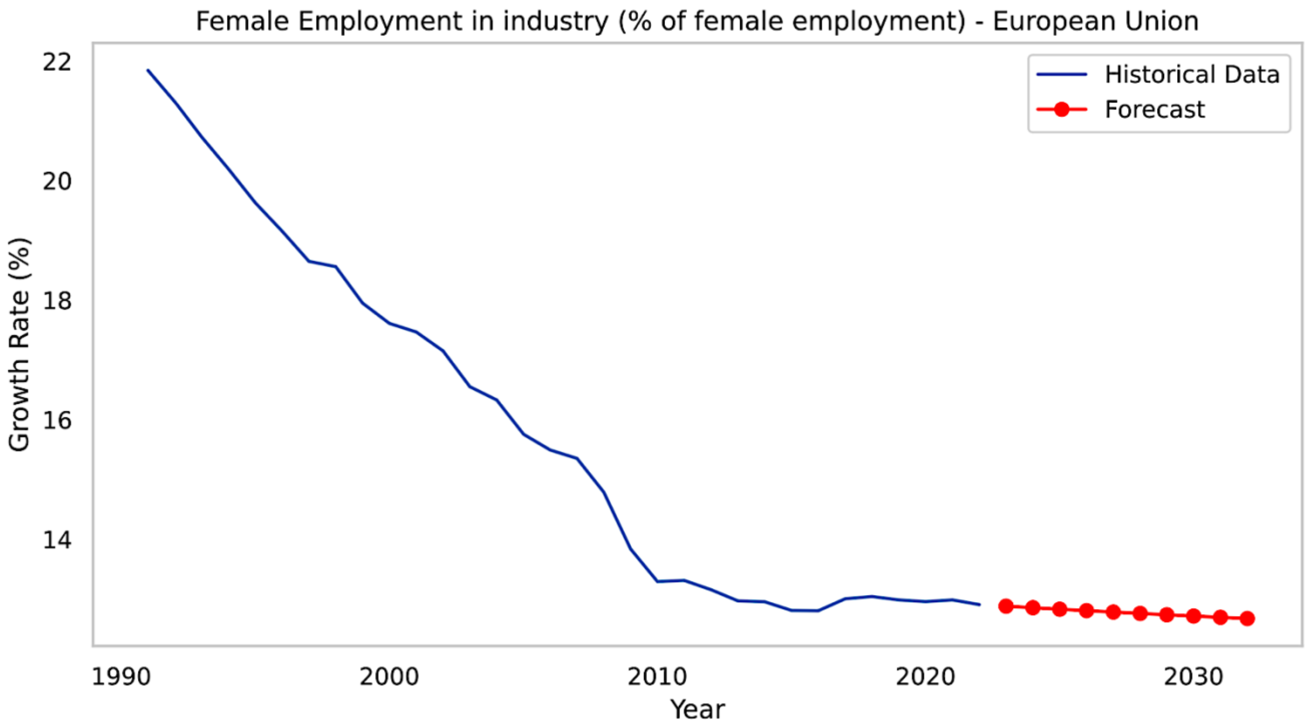

Figure 4.44 - Future Employment of Women in Industry in the EU

whether it could really happen, as socio-economic changes or policy interventions and shifts in labor market dynamics might significantly alter the trajectory.

On the contrary, it seems that the share of women working in the employment sector has declined quickly until 2010, but has remained stable in recent years, and the forecast confirms the plateau that the data might have reached. The model is again an *ARIMA(1,1,1)* and, though slightly declining over

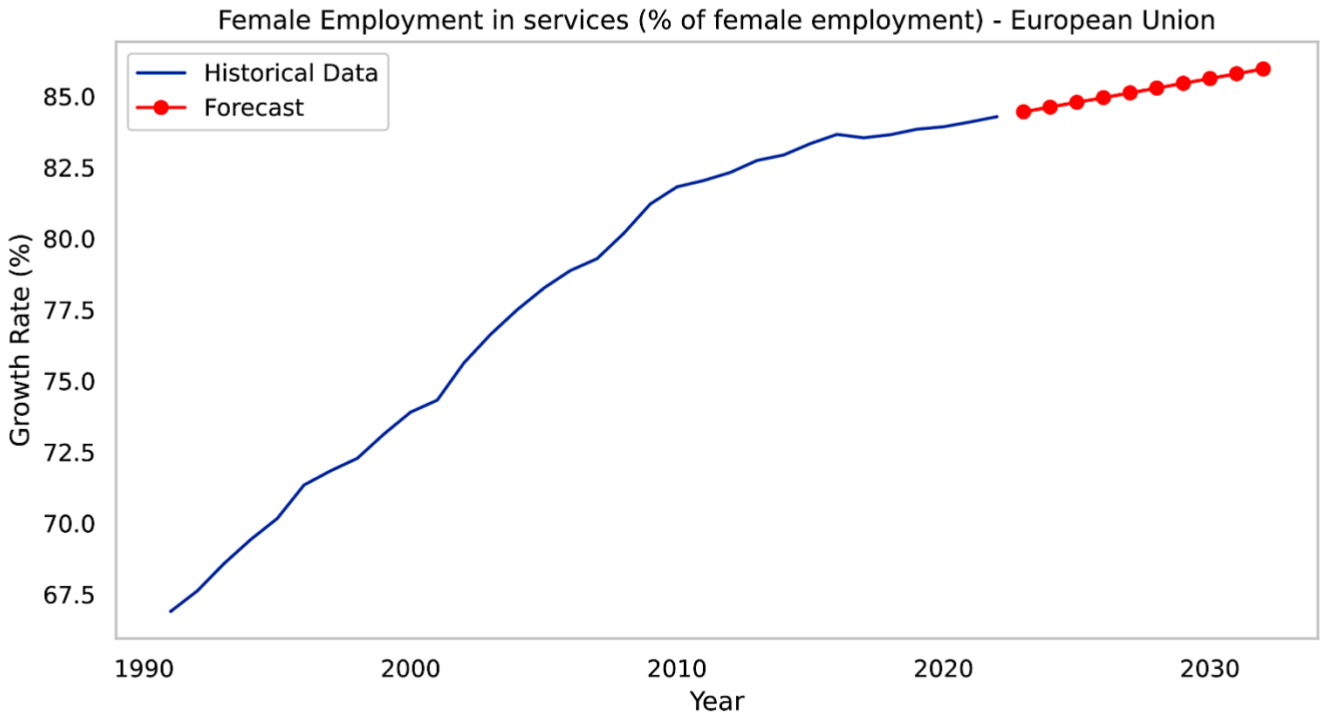

Figure 4.45 - Future Employment of Women in Services in the EU

the years, it suggests that the share will possibly remain at approximately 5% for the foreseeable future.

A much different is observable in the percentage of women working in services, since the historical data shows an increasing trend with no significant fluctuations, which is again confirmed by the 10-years forecast (*ARIMA(1,2,1)*), that shows a continuous increasing trend, reaching over 85% after 2030.

These forecasts provide a valuable perspective on the future trends of female employment across the three economies. However, while these projections help illustrate the expected trajectory based on historical patterns, it is important to acknowledge the inherent limitations of the model: ARIMA relies solely on historical data and does not incorporate external that could significantly impact employment trends. This limitation means that while the model might capture past trends and extrapolate them into the future, it does not account for unexpected disruptions or policy interventions that could alter the course of employment distribution.

Additionally, the availability and consistency of data vary across the three economies, which may introduce inconsistencies in the accuracy of the forecasts. Despite these constraints, the model still provides a general overview of potential future trends and serves as a useful tool for comparing employment trajectories among the three regions.

The findings reaffirm the ongoing structural shifts in female employment, with a continued decline in agriculture, stagnation in industry, and a growing concentration in the services sector—particularly in the European Union; China and India, while following similar trends, exhibit more dynamic changes. This comparison underscores the varying stages of economic transformation in each region and highlights the importance of context when interpreting employment forecasts.

4.3 Exploring the Relationship of Urbanization and Employment Dynamics

This section aims to combine all the previous analyses on urbanization and women employment patterns in order to establish whether there is a connection between the evolution of urban growth and that of the ever-changing dynamics of women's labor situation. It will first provide an overview of the Pearson's correlation between the two most important urbanization indicators, *Urban population (% of total population),* and *Population in urban agglomerations of more than 1 million (% of total population),* and the employment indicators of the three countries. Thus, it will try to answer to the questions and prove the hypotheses that have shaped this thesis so far.

Figure 4.46 combines the correlations of urbanization with the employment indicators, focusing only on the urbanization indicators, without again proposing the full correlation heatmap.

In the European Union, the labor force participation rate exhibits an exceptionally high positive correlation with the urbanization indicators, reaching 0.99 for the *Urban population* and 0.96 for the *Population in urban agglomerations of more than 1 million*; in contrast, India displays a low but negative correlation of approximately -0.34, while China's labor force participation presents a highly

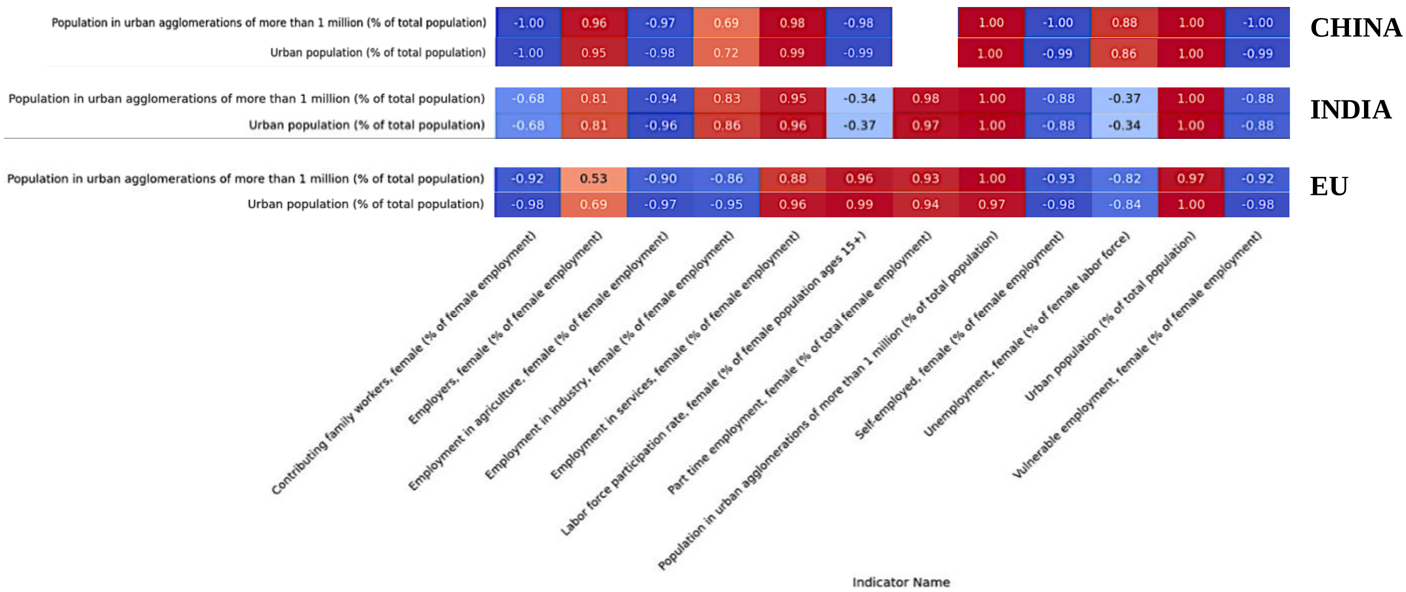

Figure 4.46 - Correlation of Urbanization Indicators and Employment Indicators by Country

negative correlation with both the urbanization indicators.

It is interesting to note that the *Unemployment* indicator follows the opposite pattern: in China, urbanization correlates positively with unemployment, while in the European Union the correlation is strongly negative. India, instead, maintains a consistent weak negative correlation.

Regarding the employment sectors, there are also several differences between developing and developed countries: the share of women in agriculture shows a strong negative correlation with urbanization across all three economies, indicating that as urbanization progresses, female employment in agriculture declines. Employment in industry, instead, presents diverging trends: in the European Union the correlation is strongly negative, whereas in China and India the correlation is highly positive. On the other hand, employment in services is positively correlated with urbanization across all three economies.

Another interesting aspect to analyze is the correlation between urbanization and share of female employers. While it is positive in all three economies the correlation is positive, it seems that the strength of this correlation varies between developed and developing countries: the European Union exhibits a weaker correlation than China and India, and among the developing countries, China's correlation is stronger than India's.

Finally, the correlation between urbanization and precarious employment (as measured by *Contributing family workers, Self-employed workers* and *Vulnerable employment*) is strongly negative across all three economies, although India exhibits a slightly weaker correlation.

To conclude, in all three regions analyzed, urban growth is strongly associated with a decline in agricultural employment among women and a simultaneous increase in service-sector employment. However, the effects of urbanization vary considerably depending on the economic context: in the European Union, a developed economy, urbanization is closely linked to higher female labor force participation and lower unemployment rates. In contrast, in developing economies such as China and India, the relationship between urbanization and women's employment presents a more complex picture.

Furthermore, the analysis suggests that urbanization is associated to reducing precarious employment for women across all three countries; however, the extent of this effect varies, with more pronounced improvements observed in developed economies compared to developing ones.

The next chapter will present a comprehensive discussion of the results, synthesizing the key findings from the analysis and directly addressing the research questions outlined in this study. By examining the interplay between urbanization and women's employment dynamics, the discussion will provide deeper insights into the broader implications of these trends across China, India, and the European Union. Furthermore, this section will evaluate the extent to which urbanization is associated with labor force participation, employment distribution across sectors, and the transition from informal to formal employment. Through this discussion, the research questions will be systematically answered, highlighting both the commonalities and divergences between developed and developing economies in the context of urban growth and gendered labor market changes.

CHAPTER 5. Discussion of the Results

The analysis presented in this thesis has sought to explore the relationship between urbanization and employment dynamics, with a particular focus on gender disparities in China, India, and the European Union. By investigating a range of employment indicators, including labor force participation, sectoral employment distribution, and precarious work, this research has aimed to provide a comprehensive understanding of how urbanization relates to women's role in the labor market.

This section will discuss the key findings in relation to the three primary research questions that have guided this study. The first research question (RQ1) examined how urbanization is associated with employment and unemployment patterns across different economic contexts, particularly distinguishing between developing (China and India) and developed (European Union) economies. The second research question (RQ2) explored the extent to which urbanization is related to the distribution of employment across the three major sectors (agriculture, industry, and services) highlighting variations among the three regions. Finally, the third research question (RQ3) analyzed trends in precarious employment and the informal economy, assessing whether urbanization has facilitated women's transition from informal to formal employment.
By synthesizing the results obtained from exploratory data analysis, correlation studies, and forecasting models, this discussion will reflect on both the expected and unexpected trends that emerged. It will also consider potential explanations for observed variations between countries, the limitations of the data, and the broader implications of these findings. Through this comparative approach, the discussion aims to contribute to a deeper understanding of how urbanization continues to shape gendered labor market structures, identifying both opportunities and persistent challenges for women in the workforce.

5.1 The Evolution of Urbanization

The findings in this study confirm that urbanization has been a persistent and accelerating trend across China, India and the European Union, albeit at different rates and with varying demographic implications.

The results indicate that China and India have experienced significant population growth over the past six decades, while the European Union has maintained a slower and more stable population increase: these findings align with the literature on demographic transitions, which suggest that developing countries tend to experience rapid population growth during early urbanization phases, whereas developed countries undergo a slower and more controlled demographic expansion.

China’s population growth, however, has slowed considerably in recent decades, largely due to policies such as the One-Child Policy and broader economic transformations that, according to research, have reduced fertility rates due to factors such as increased female workforce participation and changing socio-economic priorities.

On the other hand, India continues to experience a moderated to high population growth, reflecting a later stage in demographic transition compared to China; the literature suggests that India is following a similar path to China, only with a delay of a few decades.

The European Union, as expected, has the slowest population growth rate, reflecting a mature demographic transition, especially considering the increasing aging population and the lowered birth rates, which reflects the trends of highly urbanized and industrialized economies that typically experience stagnating or even declining populations.

An important observation from the findings is the clear divergence in urban and rural population growth trends across the three economies. China has exhibited the most extreme case of urbanization, with rapid urban growth occurring at the expense of rural population decline; indeed, since the 1990s, rural population has become negative, as more people abandon rural areas in favor of cities or urban centers. India, however, while still experiencing a rise in urbanization, still shows a slower rural decline: the findings indicate that rural population growth remains relatively stable albeit slightly declining in recent years. This suggests that while urban migration is occurring, a significant portion of India’s population continues to reside rural areas: literature suggests that this is due to India’s continuous reliance on agriculture and a slower industrial expansion.

A different picture is depicted in the European Union, a region that was already highly urbanized before 1960, and that now shows a stable but slow urban growth, with its rural population steadily declining. Indeed, as the urbanization is highly advanced, there is limited internal migration from rural to urban areas, and urban expansion often occurs through suburbanization and metropolitan growth, rather than from migration.

On this note, the findings show that there is an increasing population density in urban centers, particularly in China and India, where urban agglomerations (cities with more than 1 million inhabitants) have expanded rapidly: this trend clearly shows what had already been anticipated in the literature, in the sense that developing countries are more prone to experience a surge in megacities as urbanization accelerates, often driven by a too quick population growth and job concentration in metropolitan areas. This increasing population in urban areas has often resulted in the formation of slums, when migrating population from rural areas to cities has outpaced the availability of space and opportunities for employment. For example, in China, findings show that the urban population (as the percentage of the total population) has spiked from 16% in 1960 to over 65% in recent years, which led, according to the research, to the rapid expansion of megacities such as Shanghai, Beijing or Shenzhen. The exasperation of the increase in population density, however, is particularly visible in India: the findings show that the country has the highest population density compared to China and the EU, although the urban population, despite having increased, remains much lower than China. Both the findings and the literature suggest that this particular pattern of the population density in India is driven by the overpopulation and its fast growth which is making it harder for the government to provide enough space in its major cities, thus resulting in a high level of population density. From these findings it is also pictured how China's urbanization process has been more aggressive than India's, and the literature confirms that it is largely due to state-led policies.

The European Union presents a much different urbanization pattern, as the majority of the urban centers were established before 1960, thus resulting in a slower urban growth and more stable statistics on the topic. The findings show that the EU has the highest share of people residing in its largest city, although the population living in agglomerations of more than 1 million has remained relatively stable and much lower than China. It stands to notice that urban areas in Europe are much smaller than cities in China or India, especially because Europe's urban growth has had a gradual increase over time, contributing to a more balanced urban and infrastructure expansion of cities and thus avoiding the advent of megacities. For this reason, most people in the EU live in its biggest urban areas, but the share is not remotely close to the extent of China.

5.2 Urbanization and Employment Dynamics in Developed and Developing Economies (RQ1)

After a brief overview of the findings on the urbanization and population patterns across the three economies, the focus now shifts to answering the three research questions that guided this study.

The first research question asked how the level of urbanization relates to the employment and unemployment patterns across China, India, and the European Union, and whether there is a difference between developing and developed countries. The analyses suggest that urbanization is highly related to the shaping of labor force participation and unemployment, and that there are significant variations between developed and developing economies.

To answer the question, two hypotheses were formulated for guidance: H1.1 hypothesized that urbanization would have a negative correlation with unemployment and a positive correlation with employment across all regions.

The findings, however, present a different picture. They indicate that while this hypothesis is valid in the context of the European Union, where urbanization is associated with increased labor force participation and reduced employment, the findings for China and India suggest a more complex relationship.

In the European Union, the correlation between urban population growth and female labor force participation rate is highly positive, which reinforces the notion that urbanization might contribute to better employment opportunities, as theorized in the literature. As cities expand, access to formal employment, educational opportunities and professional networks improves, leading to women increasingly engaging in the job market, as well as a reduction in unemployment rate. This aligns with the economic theories that suggest that developed urban economies offer a more structured labor market with fewer barriers to entry for women.

Conversely, in China and India, findings show that urbanization is even negatively correlated with the participation of women in labor force. The results indicate a highly positive correlation between urbanization and unemployment in China, suggesting that rapid urban growth might be outpacing job creation, which could be attributed to rural-to-urban migration exceeding labor market absorption capacity, leading to rising unemployment despite the economic expansion. India, on the other hand, shows a weak negative correlation between urbanization and both unemployment and labor force participation rate. When looking at the other findings, indeed, it seems that while urbanization is on track with employment generation, its effect is less pronounced than in European Union.

Therefore, it can be said that H1.1 is *partially confirmed*, as the expected trend holds only for the European Union, and thus for the developed economy, while China and India demonstrate more nuanced patterns, where urbanization and the share of employed women are not on the same track. Because correlation does not necessarily imply causation, it is important to note that this does not stand to assess that urbanization necessarily influences the employment dynamics. It is possible that while urban population growth contributes to the share of employed women or to the level of unemployment, external factors such as socio-political changes, culture and government policies also play a part in shaping the afore-mentioned dynamics.

The second hypothesis proposed that as urbanization increased, the share of women employers would rise across all countries, resulting in a positive correlation between these indicators. The findings strongly encourage this hypothesis, although the magnitude of the correlation varies by country.
From the visualizations of the evolution of the indicators, a general increasing trend for both the urban population and the share of women employers is shown.
Moreover, China exhibits a strong correlation between urbanization and the rise of women employers, indicating that as cities expand and the country becomes more urbanized, more women might be able to enter entrepreneurial roles, which is linked to what has been proposed in the literature on the rising emancipation of women in urban areas. This might be also attributed to government policies promoting female-led businesses and increased access to financial capital in urban areas. However, despite the positive trend, it is illustrated in the findings how China still lags behind the EU in the overall percentage of female employers, suggesting that structural barriers still remain.
India also demonstrates a positive correlation between urbanization and women employers, though to a lesser extent that China: the slower pace of urban economic transformation in India may indeed contribute to fewer opportunities for female entrepreneurs, in addition to gender disparities in financial inclusion and cultural constraints that may limit the full potential of women.
Finally, in the European Union urbanization is positively associated with female entrepreneurship, although the increase is not as dramatic as in developing economies. The higher baseline of women employers in developed economies suggests that urbanization might act as an additional enabler, providing resources such as financial services or professional networks.
Thus, H1.2 is *confirmed* as all three economies exhibit a positive correlation between urbanization and the percentage of women employers, as well as showing increasing trends in both indicators.

From the findings, it can be assessed that urbanization does seem relate to the changing employment dynamics, but its contribution differs between developed and developing economies. While in developed economies urbanization probably enhances and enables the increasing participation of women in labor force as well as the decrease of unemployment, in developing economies such as China and India, the rapid urbanization has not necessarily resulted in increasing engagement of women in the job market. Urban growth’s negative correlation to the share of employed women in China and India, in addition to the theories provided in the literature, suggests that, while urbanization has certainly shifted the dynamics, other factors such as the employment capacity in urban areas, economic policies and the structural composition of labor markets, as well as the intrinsic culture of the countries, significantly influence the outcomes.

The results underscore the importance of targeted policy interventions to ensure that urban growth leads to equitable employment opportunities, and, in developing economies, strategies to support rural-to-urban migrants and urban women in engaging in more employment opportunities could enhance the positive effects of urbanization.

Meanwhile, in the developed economies, it is crucial to maintain policies that encourage work-life balance and gender inclusivity in order to sustain women's employment share in the labor market.

5.3 Sectoral Shifts in Women's Employments (RQ2)

This section aims to answer the RQ2, which asks the relation of urbanization to the shift of the dynamics in the major employment sectors and how the effects differentiate across all three countries. The evolution of the share of employment in the agricultural, industrial and services sectors is a key aspect of labor market transformation, and, with the expanding cities and the developing economies, the employment sectors are also expected to see structural changes. Indeed, the hypothesis related to the research question is that as urbanization increases, the share of women employed in the agricultural sector will decrease, which will correspond to an increase in the industrial and services sector. This effect is supposed to be more visible in developed countries than in developing countries. This hypothesis follows the established economic theories of structural change, where urbanization coincides with a transition from agrarian-based employment to industrialization and, ultimately, a service-driven economy. The findings of this study confirm the broader trend; however, the magnitude and pace of these changes differ significantly among the countries.

Indeed, employment in agriculture does decline across all three economies as urbanization progresses: in particular, China exhibits the steepest decline in female agricultural employment, suggesting that over the decades, China's rapid urban expansion, coupled with the increasing economic reforms that pushed towards intensive industrialization, has facilitated a massive shift from agricultural jobs to the other sectors. India follows the same trajectory as China, albeit at a slower pace, as, despite the increasing industrialization, a significant proportion of female workforce remains in agriculture, which is also shown in the different rural and urban growth rates of India compared to the other two economies. These findings suggest that in India it remains more difficult for women to shift away from agricultural labor as social norms, economic barriers and the limited absorption capacity of urban job markets continue to deter many women from moving to different types of employment.

The European Union, instead, as a highly urbanized economy, has the lowest share of women working in agriculture, and the findings do not necessarily suggest a relationship with urbanization, at least in the data in analysis. Indeed, in the European Union, the agricultural workforce is already minimized.

Industrial employment is where the findings reveal mixed results. In China and India, urbanization is positively correlated with the share of employment in the industrial sector, suggesting that to increasing urbanization levels corresponds a higher percentage of employment opportunities in industries. In China, for example, the government-driven heavy industrialization has absorbed a high number of female workers into factories and manufacturing hubs, with India showing a similar pattern.

Conversely, the European Union presents the opposite trend, with urbanization negatively correlated with industrial employment. Unlike developing economies, the EU has probably already undergone the shift from agriculture to industry, and it is now shifting more towards a more service-driven employment, with less women working in manufacturing.

Finally, the findings confirm that urbanization is strongly associated with an increase in the service-sector employment for women across all three economies, with the EU showing the highest association, and China following close behind. India, despite still maintaining a relatively high correlation, shows a slower transition, indicating existing barriers to entry for women in the sector.

The hypothesis is again *partially correct*. While it has been shown to be correct that the rise of urbanization corresponds to a decrease in women's engagement in agricultural employment and to an increase in the service-based employment, the industrial sector provides some differences across the three economies. Moreover, the shifts in employment dynamics appear more evident in the developing countries, rather than in the EU.

Therefore, it can be answered that the findings do confirm that urbanization presents strong associations with the changing employment dynamics, probably playing a role in reshaping female employment distribution, although its effects are heavily influenced by a country's stage of economic development.

Moreover, the forecasting analysis provided further insights into the expected evolution of female employment in the three sectors, indicating that the transition from agriculture to other employment opportunities will continue across all countries, while keeping the employment in industry stable for India and China, indicating that these economies will rely on manufacturing as an important employer for women for the foreseeable future. However, in the EU, industrial employment is projected to decline further, keeping consistent with the already established trend. The most notable projection is the continued rise of female employment in the services sector, expected to remain or become the dominant one across the three economies, as they are all moving (or moved, in the case of EU) through the transition to an urbanized society.

This has important policy implications, particularly for developing economies, as governments need to ensure that urbanization leads to inclusive job opportunities in all sectors in order to promote a gender-equitable labor market.
Meanwhile, in developed economies, policies should focus on enhancing women's access to high-value service-sector jobs and address gender disparities in career progression.

5.4 Precarious and Formal Employment (RQ3)

The third research question focused on precarious employment, which is often characterized by low wages, lack of job security and informal work arrangements. This type of employment remains a critical challenge in labor market, particularly for women, and the question posed in this study asks how precarious employment is changing in its dynamics and how does urbanization relate to the changes in this type of employment for women. Two hypotheses derived from the questions, supposing that vulnerable employment is much higher in developing countries, particularly for India, and that there is a general decrease of precarious employment (*Contributing family workers*, *Self-employment* and *Vulnerable employment*) and rates as urbanization increases, especially in developed countries.
The findings confirm that this type of employment still remains significantly higher in developing economies, while developed economies maintain lower share of women working informal jobs. Additionally, while it generally declines over time, the rate of decline is shown to be more pronounced in developed economies, aligning with the hypothesis that urbanization and economic growth may contribute to increased formal employment opportunities, although the contribution depends on the level of development.
Indeed, H3.1 is confirmed by the findings, which demonstrate that vulnerable employment is substantially higher in India compared to the other two economies. Precarious employment, measured in this study by the indicators of the percentage of contributing family workers, self-employed individuals, and vulnerable employment, is significantly more prevalent in developing economies, where formal job opportunities remain limited and labor protections are weaker.
India has the highest level of vulnerable employment, and women in India seem to be often employed in domestic work and unpaid care work, as well as home-based productions, where labor protections are weak and non-existent. In fact, many women in India work as contributing family workers, meaning that they are employed in family businesses, possibly without receiving wages. Moreover, as theorized in the literature, the lack of access to formal employment, financial resources and skill development possibilities perpetuates this cycle.

China, on the other hand, shows lower levels of vulnerable employment compared to India, and seems to have made progress in transitioning women from informal to formal employment; the European Union, as expected, has the lowest rate of vulnerable employment reflecting a more highly regulated labor market and stronger welfare systems. It is important to note that precarious employment does still exist, but it is nowhere near the levels of developing countries.

The European Union also extensively tracked the level of part-time employment, which has been shown to be much higher in women compared to men, still reflecting deep rooted social norms of women as caretakers.

The findings also showed varying levels of salaried workers in the three economies, which highlighted the discrepancies in both men and women perceiving an income. As the EU has the highest levels of salaried workers, reflecting the highest job protection and access to resources, China and India lag behind, with India showing a significantly low percentage of both men and women employed in stable and formal employment.

All in all, these findings support H3.1, confirming that vulnerable employment is substantially more prevalent in developing economies, particularly in India.

The second hypothesis posits that precarious employment decreases as urbanization and economic development progress, with a more pronounced decline in developed countries, and it is also supported by the findings.

In general, across all the economies there is a gradual decline in this type of employment, suggesting that economic and urban growth contribute to reducing the level of vulnerability in the job market for women.

Interestingly, in all countries, the precarious employment indicators correlate negatively with urbanization. It seems that, as urban growth fosters, a decreasing trend is observed in vulnerable employment, although the pace and the extent may vary.

In the European Union, precarious employment has been consistently low and continues to decline, as labor regulations and social policies ensure that workers have access to protections and formal job opportunities.

In China, this type of employment is also declining at a steady rate, although the process is not uniform and, as it is theorized in the literature, women rural migrants and low-skilled workers continue to face job insecurity, especially as cities grow and slums tend to form due to increased overpopulation.

Conversely, India shows the slowest decline in precarious employment, indicating that structural barriers, gender norms and informal labor market dominance continue to limit women’s access to stable jobs, and suggesting that urban growth might have enabled, but is not sufficient to promote, the transition from informal to formal job market.

Therefore, the findings suggest that H3.2 is supported by the data, but while urbanization and economic development help reduce precarious employment, they are not sufficient on their own. Indeed, the presence of strong labor policies and social security systems is what is probably needed to sustain an important decline in vulnerable employment.

5.5 Limitations of the Study

While this study provides valuable insights into the relationships between urbanization and women's employment dynamics across China, India, and the European Union, several limitations must be acknowledged. These limitations primarily relate to data availability, data reliability, and external influencing factors that were not accounted for in the analysis.

One of the most significant limitations stems from inconsistencies in the collection of data for China and India, which has already been highlighted in Chapter 2. The quality and availability of labor market data vary across different sources, and while the European Union generally provides well-documented and standardized labor statistics, data from China and India have experienced inconsistencies in methodology, coverage, and frequency of updates. These inconsistencies may impact the accuracy of trends and cross-country comparisons, particularly in the case of India, where labor force surveys have undergone significant methodological changes over time.

Another critical limitation is the absence of employment data prior to 1990 for many indicators, which restricts the study's ability to analyze long-term historical trends and understand the early phases of urbanization and employment transitions in these economies. As a result, the analysis on employment largely focuses on the period after 1990, which, while still providing meaningful insights, does not capture earlier structural shifts that may have shaped current labor market trends.

A particularly notable data gap is the absence of part-time employment statistics for China and most of India. While the European Union provides relatively complete data on part-time employment, China has no recorded data on this indicator, and India only has data for the most recent years. This missing information prevents an analysis of how part-time employment opportunities influence women's labor force participation in these countries, and it also limits the study's ability to assess the role of flexible work arrangements in shaping employment dynamics. Given that part-time employment is often a critical factor in female workforce participation, particularly in balancing work and caregiving responsibilities, this omission represents a significant gap in the analysis.

Moreover, many of the employment indicators used in this study rely on modeled estimates from the International Labour Organization (ILO) rather than directly reported national statistics. While ILO estimates are widely regarded as reliable and necessary for ensuring international comparability, they are not exact reflections of actual labor market conditions. These estimates are derived using statistical modeling techniques to fill gaps in missing or inconsistent national data, meaning that some fluctuations or trends observed in the study may not perfectly correspond to real-world labor market dynamics.

Finally, while this study primarily focuses on the relationship between urbanization and employment dynamics, it is important to recognize that employment shifts are influenced by a much broader set of external factors that were not explicitly accounted for in the analysis. These factors include: macroeconomic trends, such as economic recessions, industrial policies, and trade agreements, that shape labor demand, technological advancements, particularly automation and digitalization, which impact employment structures differently across countries.

Other factors might be education and skill development, as the availability of formal education and vocational training affects women’s employment opportunities, as well as cultural and social norms, which continue to play a significant role in shaping women’s labor force participation, especially in India.

Furthermore, government policies and labor regulations influence job security, maternity leave, childcare provisions, and access to employment rights, which significantly impact the welfare of workers.

By not incorporating these factors into the quantitative analysis, the study does not provide a complete causal explanation of employment shifts, but rather analyses the relation of urbanization and labor market trends.

CHAPTER 6. Conclusions

This study has explored the complex relationship between urbanization and women's employment dynamics in three distinct economic contexts: China, India, and the European Union. By analyzing a range of demographic and labor market indicators, the research has provided empirical insights into how urbanization influences female labor force participation, employment distribution across sectors, and the transition from informal to formal employment.

It contributes to today's global economic and social landscape, as understanding the interactions between the increasing urban growth and employment helps policymakers, economists, and development specialists in working toward a real gender equality and sustainable urban development. Moreover, the study distinguishes itself from the existing literature as it compares, through a broad range of indicators, three very different economies, which exhibit different rates and level of development.

The key findings of this study are summarized as follows: urbanization clearly relates to the labor force participation across the three economies: in the European Union, urbanization is positively correlated with female labor force participation, reflecting mature labor markets with strong employment protections and gender-inclusive policies.

China, however, has experienced a decline in female labor force participation over time, despite rapid urbanization, suggesting that labor market restructuring and rising living costs may be limiting women's workforce engagement in cities.

In India, female labor force participation remains the lowest among the three economies, with only a weak correlation between urbanization and employment, indicating that deep-rooted socio-cultural barriers and structural constraints continue to limit women's access to formal employment opportunities.

Moreover, the transition from agriculture to services seems to be a common trend, but its pace and structure vary significantly across countries.

Across all three economies, agriculture's role as an employer for women has significantly diminished, though at different rates: in China and India, industrial employment still plays a major role, particularly in urban areas, while in the European Union, deindustrialization has shifted most employment opportunities toward the service sector.

On the other hand, service-sector employment is rising in all three economies, but at different speeds. The EU leads in female service-sector employment, followed by China, where the shift is accelerating, and India, where women’s integration into services remains relatively slow due to informal labor dominance.

Finally, the prevalence of precarious employment remains high in developing economies, but urbanization may contribute to its gradual decline: India has the highest levels of vulnerable employment, with many women engaged in informal and unpaid family work, while China shows a more rapid transition toward formal employment, though certain urban labor segments, such as migrant workers, remain highly precarious.

The European Union has the lowest levels of precarious employment, though part-time and temporary work remain prevalent among women.

Furthermore, forecasting analysis suggests that the changing dynamics of the employment sectors will continue over the next decade, with agriculture's role diminishing, industry stabilizing, and services continuing to grow as the dominant employer of women. However, the extent of this transition will depend on policy interventions, economic conditions, and labor market reforms.

The findings of this study provide several important policy implications that can help governments, labor market institutions, and policymakers develop targeted strategies for fostering inclusive and equitable employment opportunities for women.

In India, policies should focus on reducing socio-cultural restrictions that limit women’s workforce entry, improving transportation safety, childcare support, and legal protections.

On a similar note, in China urban labor policies must be restructured to ensure that women have equal access to stable employment, particularly as state-owned enterprises decline, and private-sector job markets expand.

Conversely, in the European Union policies should continue to focus on gender wage equality, career advancement opportunities, and reducing occupational segregation in traditionally male-dominated sectors.

Policymakers should also encourage the sectoral transition and skill development, investing in vocational training and education for women, and, especially in India and China, targeted urban employment programs should focus on integrating women into high-value service industries, digital employment, and entrepreneurial ventures.

Governments should also support female entrepreneurship by increasing access to credit, networking opportunities, and regulatory frameworks that reduce barriers to business ownership.

It is crucial to reduce precarious employment in India and China and strengthen labor protections, especially through minimum wage laws and social security coverage, as well as legal recognition of informal workers.

As mentioned, while this study provides valuable insights, several limitations should be acknowledged, such as limitations and inconsistencies in data collection, mostly for India and China data, but also missing data on certain indicators, which leads to rely on estimates from the ILO. This results in cross-country comparisons to be less precise and does not give the possibility to address the reality of the situation.

Furthermore, the study does not account for economic crises, automation, globalization, or political shifts, all of which significantly influence employment patterns.

Future research should address these limitations by incorporating a broader range of economic indicators, including wages, education levels, and occupational mobility, but also analyzing qualitative factors, such as social norms, employer biases, and work-life balance policies, to gain deeper insights into women's employment challenges.

All in all, this study has demonstrated that urbanization significantly relates to women's employment patterns, but the extent and direction of this relationship depend on economic structures, labor market policies, and societal norms.

China is undergoing rapid employment shifts, with rising service-sector jobs, declining agricultural work, and a mixed industrial employment trend, while India continues to struggle with low female labor force participation, high informal employment, and slow transitions into formal job markets.

The European Union, instead, provides a relatively stable and structured employment environment for women, though gender disparities still persist.

Ultimately, achieving inclusive and gender-equitable labor markets will require coordinated policy efforts, improved labor protections, and targeted economic reforms. As urbanization continues to reshape labor markets, ensuring that women have access to quality, stable, and well-paid jobs will be essential to promoting sustainable economic development and social progress.

CHAPTER 7. References

Acker, J. (1990). Hierarchies, Jobs, Bodies: A Theory of Gendered Organizations. *Gender and Society*, *4*(2), 139–158. https://www.jstor.org/stable/189609

Anne de Bruin & Na Liu. (2019). The Urbanization-Household Gender Inequality Nexus: Evidence from Time Allocation in China. *China Economic Review*, *60*. https://doi.org/10.1016/j.chieco.2019.05.001

Buchholz, M. (2023). Does Urbanization Increase Inequality? Race, Gender and the Urban Wage Premium. *Journal of Urban Affairs*. https://doi.org/10.1080/07352166.2023.2252538

Chang, H., & Brada, J. C. (2006). The Paradox of China's Growing Under-Urbanization. *Economic Systems*, *30*, 24–40.

Christiansen, L. E., Lin, H., Pereira, J. M., Topalova, P., & Turk, R. (2017). *Unlocking the Potential of Greater Female Employment in Europe*. *2017*(1), 5–16. https://www.intereconomics.eu/contents/year/2017/number/1/article/unlocking-the-potential-of-greater-female-employment-in-europe.html

Cook, I. G., Gu, C., & Hu, L. (2017). China's Urbanization in 1949-2015: Processes and Driving Forces. *Chinese Geographical Science*, *27*, 847–859. https://doi.org/10.1007/s11769-017-0911-9

Davis, K. (1965). The Urbanization of the Human Population. *Scientific American*, *213*(3), 40–53. http://www.jstor.org/stable/24931110

Dhamija, Guarav, Binay, & Roychowdhury. (2023). Urbanization and Women Empowerment: Evidence from India. *Global Labor Organization*, *1323*. https://hdl.handle.net/10419/276225

Dobbs, R., & Sankhe, S. (2010). *Comparing Urbanization in China and India*. McKinsey Quarterly. https://www.mckinsey.com/~/media/McKinsey/Featured%20Insights/Urbanization/Comparing%20urbanization%20in%20China%20and%20India/Comparing%20urbanization%20in%20China%20and%20India.pdf

Eurocities. (2024). *In a world that fails women, cities step up*. https://eurocities.eu/latest/in-a-world-that-fails-women-cities-step-up/

European Commission. (n.d.). *Gender equality strategy*. Retrieved 22 January 2025, from https://commission.europa.eu/strategy-and-policy/policies/justice-and-fundamental-rights/gender-equality/gender-equality-strategy_en

ILO Modelled Estimates (ILOEST database). (n.d.). ILOSTAT. Retrieved 27 December 2024, from https://ilostat.ilo.org/methods/concepts-and-definitions/ilo-modelled-estimates/

Jaysawal, N., & Saha, S. (2014). Urbanization in India: An Impact Assessment. *International Journal of Applied Sociology*, *4*(2), 60–65. https://doi.org/10.5923/j.ijas.20140402.04

Masika, R. (1997). *Urbanisation and Urban Poverty: A Gender Analysis*.

Müller, B., Rößler, S., Jiang, M., & Schiappacasse, P. (2019). *Report on the Current Framework and Situation of Urban Renewal, Urban Expansion as Well as Land Management and Banking (incl. Land Administration) in China and Europe* [Technical Report].

Onda, K., Sinha, P., Gaughan, A., Stevens, F., & Kaza, N. (2019). Missing Millions: Undercounting Urbanization in India. *Population and Environment*, *41*, 126–150. https://doi.org/10.1007/s11111-019-00329-2

Ren, X. (2022). Urbanization, Migration and Families in China and India. *Division for Inclusive Social Development*. https://social.desa.un.org/sites/default/files/inline-files/2022-Ren-UN-memo-V2-Sep-12.pdf

Satterthwaite, D. (2007). *The transition to a predominantly urban world and its underpinnings*.

Tacoli, C. (2012). *Urbanization, Gender and Urban Poverty: Paid Work and Unpaid Carework in the City*.

Third Plenary Session of the 18th CPC Central Committee. (n.d.). Retrieved 21 January 2025, from http://dk.china-embassy.gov.cn/eng/ztbd/18ththirdplenary/#

Thorbek, S. (1988). Women and Urbanization. *Acta Sociologica*, *31*(4), 283–301. https://www.jstor.org/stable/4200719

Unpaid Care Work – A Gateway for Gender Equality. (n.d.). UNDP. Retrieved 19 January 2025, from https://www.undp.org/papp/blog/unpaid-care-work-gateway-gender-equality

World Bank. (2024). *Who We Are*. https://www.worldbank.org/ext/en/who-we-are

World Development Indicators (World Bank). (n.d.). European University Institute. Retrieved 27 December 2024, from https://www.eui.eu/Research/Library/ResearchGuides/Economics/Statistics/DataPortal/WDI

Zhang, S. (2024). Analysis of Female Employment in China's New-Type Urbanization. *Business, Economics and Management*, *43*.

CHAPTER 8. Appendix

Indicator Name	Description
Population growth (annual %)	Represents the exponential rate of growth of midyear population from year t-1 to t, expressed as a percentage. The population counts all residents regardless of legal status or citizenship.
Population density (people per sq. km of land area)	Calculated dividing midyear population by land area expressed in square kilometers. Land area is a country's total area, excluding area under inland water bodies or exclusive economic zones.
Population in urban agglomerations of more than 1 million (% of total population)	Represents the percentage of a country's population living in metropolitan areas that in 2018 had a population of more than one million people.
Population in the largest city (% of urban population)	Represents the percentage of a country's population living in that country's largest metropolitan area.
Population, total	Based on the de facto definition of population, which counts all residents regardless of legal status or citizenship. The values shown are midyear estimates.
Population, female	Based on the de facto definition of population, which counts all female residents regardless of legal status or citizenship. The values shown are midyear estimates.
Population, male	Based on the de facto definition of population, which counts all male residents regardless of legal status or citizenship. The values shown are midyear estimates.
Rural population growth (annual %)	Refers to the people living in rural areas as defined by national statistical offices. It is calculated as the difference between total population and urban population.
Urban population growth (annual %)	Refers to the people living in urban areas as defined by national statistical offices. It is calculated using World Bank population estimates and urban ratios from the United Nations World Urbanization Prospects.
Urban population (% of total population)	Refers to the people living in urban areas as defined by national statistical offices.

Table 8.1 – Population Statistics Definitions

Indicator Name	Description
Labor force, female (% of total labor force)	Represents female labor force as a percentage of the total labor force and shows the extent to which women are active in the labor force.
Wage and salaried workers, female (% of female employment) (modeled ILO estimate)	Represents the percentage of women who hold the type of jobs defined as "paid employment jobs" where the incumbents hold explicit employment contracts that give basic remuneration.
Wage and salaried workers, male (% of male employment) (modeled ILO estimate)	Represents the percentage of men who hold the type of jobs defined as "paid employment jobs" where the incumbents hold explicit employment contracts that give basic remuneration.

Contributing family workers, female (% of female employment) (modeled ILO estimate)	Share of female workers who hold "self-employment jobs" as own-account workers in a market-oriented establishment operated by a related person living in the same household. They are also known as unpaid family workers.
Contributing family workers, male (% of male employment) (modeled ILO estimate)	Share of male workers who hold "self-employment jobs" as own-account workers in a market-oriented establishment operated by a related person living in the same household. They are also known as unpaid family workers.
Employers, female (% of female employment) (modeled ILO estimate)	Female employers are workers who hold the type of jobs defined as "self-employment jobs" (i.e. jobs where the remuneration is directly dependent upon the profits derived from the goods and services produced) and have engaged one or more persons to work for them as employees.
Employers, male (% of male employment) (modeled ILO estimate)	Male employers are workers who hold the type of jobs defined as "self-employment jobs" (i.e. jobs where the remuneration is directly dependent upon the profits derived from the goods and services produced) and have engaged one or more persons to work for them as employees.
Employment in agriculture (% of female employment) (modeled ILO estimate)	Defined as women of working age who were engaged in any activity to produce goods for pay or profit in the agricultural sector, which consists of activities in agriculture, hunting, forestry, and fishing.
Employment in agriculture (% of male employment) (modeled ILO estimate)	Defined as men of working age who were engaged in any activity to produce goods for pay or profit in the agricultural sector, which consists of activities in agriculture, hunting, forestry, and fishing.
Employment in industry (% of female employment) (modeled ILO estimate)	Defined as women of working age who were engaged in any activity to produce goods or services for pay or profit in the industrial sector, which consists of activities of mining, manufacturing, construction and public utilities.
Employment in industry (% of male employment) (modeled ILO estimate)	Defined as men of working age who were engaged in any activity to produce goods or services for pay or profit in the industrial sector, which consists of activities of mining, manufacturing, construction and public utilities.
Employment in services (% of female employment) (modeled ILO estimate)	Defined as women of working age who were engaged in any activity to produce goods or services for pay or profit in the service sector, which consists of wholesale and retail trade, restaurants, hotels, transport, storage, communications, financing, business services and insurance.
Employment in services (% of male employment) (modeled ILO estimate)	Defined as men of working age who were engaged in any activity to produce goods or services for pay or profit in the service sector, which consists of wholesale and retail trade, restaurants, hotels, transport, storage, communications, financing, business services and insurance.
Vulnerable employment, female (% of female employment) (modeled ILO estimate)	Defined as female contributing family workers and own-account workers as a percentage of total employment.
Vulnerable employment, male (% of male employment) (modeled ILO estimate)	Defined as male contributing family workers and own-account workers as a percentage of total employment.

Part time employment, female (% of total female employment)	Refers to the percentage of women employed in a regular employment in which working time is substantially less than normal.
Part time employment, male (% of total male employment)	Refers to the percentage of men employed in a regular employment in which working time is substantially less than normal.
Self-employed, female (% of female employment) (modeled ILO estimate)	Refers to female workers who work on their own account or with one/few partners or in a cooperative and hold the jobs where the remuneration is directly dependent upon the profits derived from the goods and services produced.
Self-employed, male (% of male employment) (modeled ILO estimate)	Refers to male workers who work on their own account or with one/few partners or in a cooperative and hold the jobs where the remuneration is directly dependent upon the profits derived from the goods and services produced.
Unemployment, female (% of female labor force) (modeled ILO estimate)	Refers to the share of the female labor force that is without work but available for and seeking employment.
Unemployment, male (% of male labor force) (modeled ILO estimate)	Refers to the share of the male labor force that is without work but available for and seeking employment.

Table 8.2 – Employment Statistics Definition

Equation 1 - First Order Differencing

$$y'_t = y_t - y_{t-1}$$

Equation 2 - Second Order Differencing

$$y''_t = y'_t - y'_{t-1} = (y_t - y_{t-1)}) - \big(y_{t-1} - y_{t-2)}\big) = y_t - 2y_{t-1} + y_{t-2}$$