Pierre Aliphat
Nikola Damjanovic
Pierre Tapie

HIGHER EDUCATION REVENUE & EMPLOYABILITY

How is access to higher education correlated to employability?

Should access to higher education be promoted to develop employment?

How is revenue correlated with access to higher education?

A global comparative Study

PAXTER

Table of contents

Preamble

As a space of learning, Higher Education carries multiple missions: the transmission of knowledge and culture; the intellectual formation of students and the development of their critical thinking; the acquisition of cognitive knowledge, disciplinary expertise, an understanding of civic life, and a sense of citizenship. Moreover, it is tasked with conferring skills that are relevant from both economic and societal standpoints, which will build graduates' employability and their capacity to contribute to economic and social development.

The following research monograph, the results of which may come as a surprise for some, was initiated ten years ago, sparked by a piece of information that runs counter to intuition most French people have. In the countries of the 2011 Arab Spring, on the southern shore of the Mediterranean, the unemployment rate among young graduates was on average 50% higher than that of young people in general; and holding a degree appeared to be a disadvantage in the search for employment. Combined with the perception among parts of the youth population that civic freedoms were insufficient, this situation was one of the underlying causes of the acts of despair that rapidly spread throughout the region, eventually evolving into a large-scale political movement. Meanwhile, the authors of this study were immersed, in France, in the seemingly self-evident notion that expanding access to higher education was the pathway to lowering unemployment and fostering economic growth.

This work therefore arose from the apparent dissonance between this observed phenomenon and a widely held French conviction. Paxter, a consultancy firm specialising in academic strategies and pedagogical engineering, has developed its own independent research production in order to strengthen its expertise. The relationship between income, employment, and higher education thus emerged as a key question to be explored.

We have chosen to communicate our findings according to the principles of open science. It would have been possible, perhaps, to extract several articles from this body of work for publication in academic journals. However, the simultaneous publication of the entirety of our findings, each shedding light on different facets of the issue under investigation, seemed to us more meaningful and more closely aligned with the role research plays in our consulting mission.

At the time of releasing this work, the brutal attacks on universities, research, and science by the new administration of a major scientific nation caused us to hesitate. As professionals in higher education and research, did we have the moral right to publish results that could potentially be misused by forces hostile to the pursuit of objective knowledge and a rational understanding of the world, some of which would even go so far as to destroy essential experimental data?

In the end, if we have decided to publish this body of work, it is precisely because we are convinced that knowledge is preferable to ignorance, that scientific research based on facts and rational analysis helps every human community to progress, regardless of its initial mental biases; that investment in research, provided it is of high quality, is what will drive both innovation and societal resilience. Our approach is solely grounded in the observation and objective analysis of reality, without preconceived ideas or premature conclusions. If our results could surprise some readers, particularly in developed countries, they can rest assured that they have been produced with the sole ambition of **understanding** and submitted to the scientific community to shed new light on certain issues. We will be most pleased by exchanges that may arise from this work.

Pierre Aliphat Nikola Damjanovic Pierre Tapie

July 2nd 2025

1 Executive Summary

This body of work examines the relationship between young people's access to higher education, average revenue in a country, and employability. Though this subject has so far been addressed in the literature mainly at a national level, this monograph seeks to explore these correlations from an international perspective.

The variables considered in this work are: GDP per capita expressed in purchasing power parity (GDP per capita (PPP)), the access rate to higher education, general unemployment rates, graduate unemployment rates, non-graduates unemployment rates, youth unemployment rates, young graduate unemployment rates, and young non-graduate unemployment rates.

The innovative aspect of our study lies essentially in the extensive dataset analyzed, covering a group of countries representing over 90% of the global youth population, and for which we were able to restate the rates of access to higher education one by one. Our results suggest that beyond the level of education, it is the relevance of qualifications to employment and the stage of a country's economic development that determines employability.

We approached this research through five sequential questions:

Question 1: How is a country's level of economic development correlated to its rate of access to higher education?

To address this, we examined potential correlations among the 140 countries for which this data was simultaneously available. We identified a strong correlation with a coefficient of R = 0.72. We also found that oil-producing countries and certain financial hubs exhibited "outlier" behaviors. Indeed, their access to higher education rates were low, relative to their levels of economic development.
This can be explained by the fact that, in oil-producing countries, oil extraction does not require a highly skilled workforce. In the case of financial hubs, financial transactions artificially inflate GDP per capita (PPP). By excluding nine countries from these two categories deemed atypical, the correlation among the remaining 131 countries between economic development and access to higher education is particularly strong (R = 0.84).

Furthermore, considering all countries with low GDP per capita (PPP), the correlation for those below the median ($15,000) is stronger than that for all 140 countries (R = 0.78 instead of 0.72). For those with GDP per capita above $15,000, economic development and access to higher education are not statistically correlated. Thus, if a relationship exists between *economic development and access to higher education for countries with low economic development, beyond a medium level of development ($15,000), economic development and the rate of access to higher education become independent.*

Question 2: Can correlations be established between economic development and various unemployment rates (total, graduates, youth, and young graduates)?

We first examined the nature of the relationship between each type of unemployment and the level of economic development of each country. Since all these variables have non-Gaussian distributions, we used Spearman's correlation to establish the existence (or not) of a monotonic relationship between them.
Our conclusions indicate that the correlations between general unemployment rates, youth unemployment rates, and young non-graduate unemployment rates with economic development are statistically very weak (-0.253, -0.22, -0.225 respectively, meaning economic development explains only 5 to 6% of variations in unemployment rates). Ultimately, the unemployment rate of non-graduates is not correlated with economic development.

Our conclusions further suggest significant negative correlations between GDP per capita (PPP) and the unemployment rate of graduates (R= -0.612) on one hand and GDP per capita (PPP) and the unemployment rate of young graduates (R= -0.641) on the other. Thus, 37.4% and 41% respectively of the variations in graduate unemployment and young graduate unemployment can be explained by fluctuations in GDP per capita (PPP). This suggests that the higher a country's level of economic development, the greater its demand for educated labor, leading in turn to lower graduate and young graduate unemployment rates. However, this correlation is only valid for countries with low to medium levels of economic development.

Question 3: What relationships can be observed between access to higher education and employability? Does this depend on the level of economic development?

3.1 Our results indicate that, in a cross-country comparison:

- The rate of access to higher education in a country has no impact on total unemployment, non-graduate unemployment, youth unemployment, or young non-graduate unemployment rates.
- However, moderate negative correlations are observed between a country's rate of access to higher education and its graduate unemployment rate (-0.38), as well as its young graduate unemployment rate (-0.46).

Thus, as the rate of access to higher education increases (meaning more graduates exist in the country) the number of unemployed graduates (or young unemployed graduates) decreases. However, only 14% of this decline in graduate unemployment and 21% of the decline in young graduate unemployment are due to variations in the rate of access to higher education; the remaining 80-85% must be explained by other factors.

3.2 Within a given country, examining the average effect of holding a degree on employability, we find that among the 88 countries studied, a degree enhances employability in 55 countries but negatively impacts it in 33, of which 25 have GDP per capita (PPP) below $15,000. The observation of trends in unemployment rates over time indicate a strong consistency in their relative positions, suggesting deep-rooted socioeconomic characteristics in these countries.

3.3. The countries where obtaining a degree has the most negative effect on employability are also those where, in adulthood, a significant gender gap in graduate employability is observed to the disadvantage of women. This gender effect is minimal or absent among young graduates. The countries with the highest unemployment rates among young graduates are the same as those with the highest unemployment among graduates in general. In these countries, older female graduates are relatively more penalized than

younger female graduates, while young male graduates face a level of disadvantage comparable to that of their female counterparts.

Question 4: Can countries with comparable wealth and employability characteristics be grouped into homogeneous clusters, and can their similar behaviors be explained by shared socioeconomic factors? [1]

The clusters were formed using three methods:

- The first method considers economic development. More precisely, we analyzed the correlations between the variables of the cluster of countries with low GDP per capita (PPP) and those with high GDP per capita (PPP) (setting the boundary at $20,000 GDP/capita (PPP)).
    - For countries with income below the median, an increase in the higher education access rate has no effect on any of the six unemployment rates.
    - For countries with income above the median, an increase in the higher education access rate affects only the general unemployment rate (by increasing it) and, more significantly, the non-graduate unemployment rate. Within the comparison among wealthier countries, an increase in the higher education access rate among these nations tends to slightly raise the general unemployment rate, has no effect on graduate unemployment, and increases non-graduate unemployment due to their relative downgrading.

- The second method partitions countries based on whether their various unemployment rates are low or high. The only significant relationships observed (around -0.60) are those linking the rate of access to higher education to both the graduate unemployment rate and the young graduate unemployment rate. For the other variables, correlations are not significant, regardless of whether the country belongs to the cluster with unemployment rates above or below the median values. By comparing countries that simultaneously exhibit high general unemployment rates, as well as high unemployment among non-graduates, youth, and young non-graduates, we observe that in wealthy countries, the positive effect of obtaining a degree on employability is significantly stronger where overall unemployment is high. Conversely, in many less developed countries where holding a degree negatively affects employability, this negative effect is more pronounced among the younger population.

- The third partitioning method used an unsupervised machine learning algorithm to explore whether the previously observed trends might not in part be averages of opposing phenomena between groups of countries. This algorithm considered all variables (rate of access to higher education, GDP per capita (PPP), and various types of unemployment) to classify countries into homogeneous groups. Five clusters were identified as a result of this calculus:

  Cluster A: These are developing countries where all unemployment rates are high. The rate of access to higher education is the lowest among all clusters, but it tends to increase as economic development

[1] **Clustering** is a statistical analysis method used to organize raw data into homogeneous groups. Within each cluster, data points are grouped based on a shared characteristic. The ordering tool is an algorithm that measures the proximity between each element based on predefined criteria.

progresses. The only significant correlation in this group is between the rate of access to higher education and GDP per capita (PPP) (0.829).

Cluster B: These are middle-income countries with relatively low or moderate unemployment rates. The rate of access to higher education is relatively low but shows a clear upward trend driven by economic development. The only significant correlation in this group is between the rate of access to higher education and GDP per capita (PPP) (0.621). Unemployment rates are not correlated with any other variable.

Cluster C: These are middle-income countries with a very high rate of access to higher education and high unemployment rates across all categories. The only significant correlation in this group is between the rate of access to higher education and the young graduate unemployment rate (-0.571).

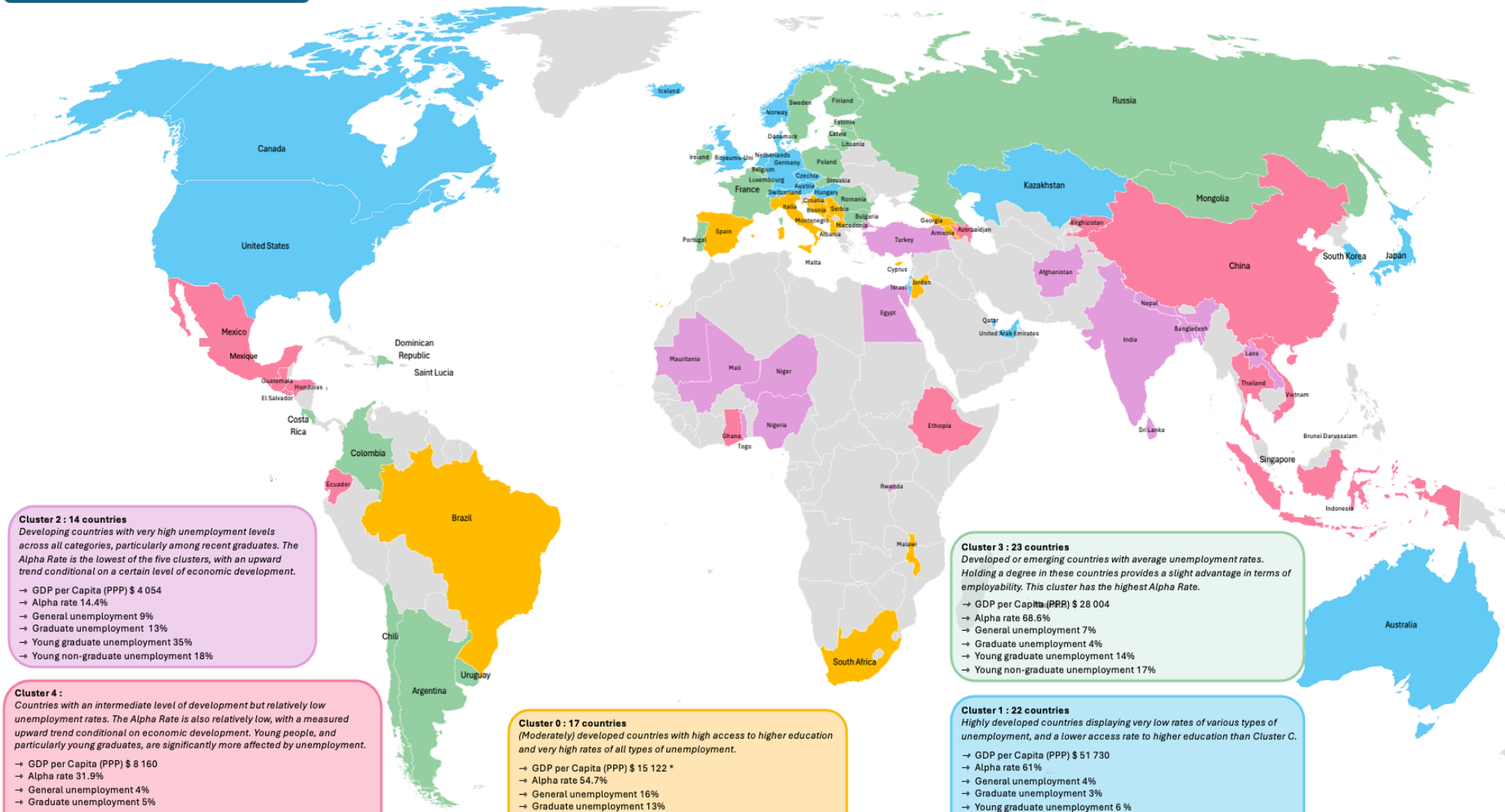
Cluster D: These are mostly highly developed countries with high non-graduate unemployment rates. In these countries, graduate unemployment is lower than non-graduate unemployment, which remains high. The rate of access to higher education is the highest in this cluster. Significant correlations include the relationship between the rate of access to higher education and the non-graduate unemployment rate (0.434) and between the rate of access to higher education and the young graduate unemployment rate (-0.528). France belongs to this cluster.

Cluster E: These are developed countries with low unemployment rates. No significant relationship exists between the rate of access to higher education and the different unemployment rates, nor between the rate of access to higher education and GDP per capita (PPP).

Employability, measured through six different unemployment rates, appears to be a stable factor within each cluster over long period or very long periods, and corresponds to the lasting results of deeply rooted socio-economic situations that are very characteristic of the different clusters. These international comparisons highlight that youth employability depends on the ability to integrate workers of all skill levels, from the least to the most qualified. Additionally, the challenge also lies in the ability of countries to adapt the level and type of skills of their workforce to the needs of their economy.

5 CLUSTERS : 89 countries

Characteristics of the countries in the 5 clusters identified



Cluster 2 : 14 countries
Developing countries with very high unemployment levels across all categories, particularly among recent graduates. The Alpha Rate is the lowest of the five clusters, with an upward trend conditional on a certain level of economic development.

- → GDP per Capita (PPP) $ 4 054
- → Alpha rate 14.4%
- → General unemployment 9%
- → Graduate unemployment 13%
- → Young graduate unemployment 35%
- → Young non-graduate unemployment 18%

Cluster 4 :
Countries with an intermediate level of development but relatively low unemployment rates. The Alpha Rate is also relatively low, with a measured upward trend conditional on economic development. Young people, and particularly young graduates, are significantly more affected by unemployment.

- → GDP per Capita (PPP) $ 8 160
- → Alpha rate 31.9%
- → General unemployment 4%
- → Graduate unemployment 5%
- → Young graduate unemployment 18%
- → Young non-graduate unemployment 9%

Cluster 0 : 17 countries
(Moderately) developed countries with high access to higher education and very high rates of all types of unemployment.

- → GDP per Capita (PPP) $ 15 122 *
- → Alpha rate 54.7%
- → General unemployment 16%
- → Graduate unemployment 13%
- → Young graduate unemployment 34%
- → Young non-graduate unemployment 32%

Cluster 3 : 23 countries
Developed or emerging countries with average unemployment rates. Holding a degree in these countries provides a slight advantage in terms of employability. This cluster has the highest Alpha Rate.

- → GDP per Capita (PPP) $ 28 004
- → Alpha rate 68.6%
- → General unemployment 7%
- → Graduate unemployment 4%
- → Young graduate unemployment 14%
- → Young non-graduate unemployment 17%

Cluster 1 : 22 countries
Highly developed countries displaying very low rates of various types of unemployment, and a lower access rate to higher education than Cluster C.

- → GDP per Capita (PPP) $ 51 730
- → Alpha rate 61%
- → General unemployment 4%
- → Graduate unemployment 3%
- → Young graduate unemployment 6 %
- → Young non-graduate unemployment 10%

* These statistics represent the median

Question 5: Where significant correlations exist, particularly between the rate of access to higher education and young graduate unemployment, how strong are these correlations, and can predictive models be built for young graduate unemployment rates?

We examined how variations in the rate of access to higher education in 2013 impacted the young graduate unemployment rate in 2017, 2017 being the average graduation year of first-cycle students who enrolled in 2013. First, we built a multiple linear regression model with a determination coefficient (R^2) of 62%, which is substantial, leading to the following equation:

Young graduate unemployment rate 2017 =
0.23 + 1.66 * Graduate unemployment rate 2013 - 0.30 * Rate of access to higher education 2013

This result indicates that a 1% increase in the rate of access to higher education in 2013 led to an average decrease of 0.3% in the young graduate unemployment rate in 2017, while young graduates experienced, on average, a 66% increase in their unemployment rate compared to the general graduate unemployment rate in 2013.

Next, we built a more powerful ensemblist model, resulting in R^2 = 87%. This model shows that increasing the rate of access to higher education in countries where it is already high (above 50%) has no impact on reducing the young graduate unemployment rate four years later. However, in countries where the rate of access to higher education is below 50%, the same increase may have a downward impact on the future young graduate unemployment rate.

At the conclusion of this study, we first demonstrated that economic development and access to higher education are positively correlated in countries with a low GDP per capita (PPP) (< $15,000). However, beyond this threshold, economic development and access to higher education become independent of one another. Therefore, implementing public policies in wealthy countries aimed at continually expanding access to higher education—especially generalist programs—under the assumption that such expansion would systematically improve access to employment, is not supported by evidence. The proportion of young people gaining employment does not increase with broader access to higher education. In contrast, in wealthy countries that are primarily characterized by their inability to provide employment opportunities to young people and non-graduates—France being one such country—increasing access to higher education improves the employability of graduates while decreasing that of non-graduates. This has no effect on overall unemployment rates and is likely due to the signaling value of diploma.

Furthermore, increasing higher education access in countries where such access is already high (above 50%) has no measurable impact on reducing the unemployment rate of young graduates after four years of education. We verified that this observation is not an artifact of clustering effects, whereby antagonistic effects between clusters of countries with homogeneous behaviors internally to clusters could distort the overall data set and bias the analysis. Such is not the case.

Finally, our study established that employability patterns are highly robust over time across different clusters, reflecting structural and enduring socio-economic characteristics.

2 Problem Statement and Research Questions

2.1 Problem statement and research objective

PAXTER's research has explored the likely demographics of students enrolled in a higher education program, this in 76 countries accounting for 90% of the world's youth over the period from 2015 to 2038. This study establishes very strong correlations between certain phenomena in past years, enabling robust forecasts of student numbers over 15-year timeframes (P. Tapie, 2014, 2015 & 2017)[2].

Reflecting on the likely number of students in the world over the next 15 to 20 years has led us to consider its consequences on employability, by exploring the correlation between access to higher education and various employment parameters.

Public officials' discourse often assume that an "ever-higher" level of education among the population would *a priori* have positive consequences on employment. Thus, at the macroscopic level in a country, the general belief is that the more qualified the population of one country compared to another, the lower the unemployment rate. Likewise, at a microscopic scale, the higher the qualification of an economic agent, the lower their unemployment rate becomes.

Our initial observations of different unemployment rates in countries with comparable levels of economic development cast doubt on this assertion. This led us to systematically explore, in as many countries as possible, how the level of diploma was or was not correlated with access to employment, within specific countries, and compared between different countries.

For this purpose, within the same country, correlations were systematically explored between a measure of access to higher education and six different employment parameters: the general unemployment rate, the general unemployment rate of graduates, the general unemployment rate of non-graduates, the youth unemployment rate (15-24 years old), the young graduate unemployment rate, and the unemployment rate of young non-graduates.

2.2 State of the art: literature review

To the best of our knowledge, the relationship between access rates to higher education and unemployment rates has not yet been thoroughly examined at the global level through a large-scale, cross-country comparison. This relationship has been explored within individual countries in numerous articles and across various research fields. While the study of a single country allows researchers to bypass certain risks associated with data heterogeneity and to focus on the analysis of a single phenomenon, the substantial differences in access to higher education from one country to another seemed to us worthy of investigation in their own right, particularly in terms of their consequences.

A key bibliographical starting point is the work of Gary S. Becker in the 1960s. In his early research, Becker developed the theory of human capital, based on the idea that individuals can acquire skills (human capital) that make them more productive. This increased productivity, in turn, leads to higher income. Becker's research

[2] Pierre Tapie, "Future Educational Directions & Challenges in Asia Pacific", Keynote Speech, Asia Pacific Deans Summitt, Séoul (Korea), August 28th, 2014 ; Pierre Tapie, "Internationalization and the student body", Keynote Speech, Canadian Federation of business School Deans, Toronto, October 17th, 2014, ; Pierre Tapie, "Higher Education Demographics and Economic New Frontiers", Keynote speech, Higher Education Summitt, 19 Octobre 2015 ; Pierre Tapie, "Singapour : Higher Education and international Mobility Trends beyond Europe", Keynote Speech, Conference TIME, October 17th, 2019

demonstrated that investments in education, vocational training, and healthcare are forms of capital investments[3]. He showed that economic returns tend to be positively correlated with skills. Moreover, he highlighted a negative correlation between education and unemployment.

In what follows, we present a set of recent and highly heterogeneous findings addressing the relationship between unemployment rates and access to higher education, as measured by the absolute number or the proportion of young people in a given age cohort entering higher education.

In 2000, Schomburg published a study on access to higher education and graduate employment in Germany[4]. He concluded that the expansion of higher education in Germany was associated with rising graduate unemployment, albeit remaining below that of non-graduates. Besides, the spread of access to higher education was coupled with growing debate about the frequent maladjustment of graduates regarding their status (income and position) and the use of their knowledge and skills.

In 2000, Mora, Montalvo and Garcia-Aracil[5] examined the relationship between access to higher education and graduate employment in Spain. They found that the surge of university students had a negative impact on unemployment among young graduates. Additionally, they reached the conclusion that unemployment was especially significant among the youngest graduates but dropped considerably for older age brackets. Between 1980 and 2000, the number of graduates in Spain was multiplied by 3.7 while the number of working graduates was multiplied by only 3.4. Put differently, despite the fact that the Spanish economy was able to create 1.5 million new jobs for graduates, 0.2 million jobs were still needed to mitigate graduate unemployment.

In 2004, Esther Duflo examined[6] the consequences of the large-scale primary school construction program in Indonesia on various attributes. Her research shows that, at a constant skill level, an increase in the proportion of children who received primary education paradoxically leads to a slight negative effect on average wages, while the proportion of workers in the formal sector increases in these areas. This counterintuitive effect, where individual wages decrease as the average education level rises, could be explained by a lower elasticity of substitution between land and labor in the informal sector compared to the elasticity of substitution between capital and labor in the formal sector. These pioneering studies thus highlight that the relationship between education level and income is not as straightforward as it may appear, even in a case where the observation of a large-scale effect was possible.

In 2006, Moreau and Leathwood published an article[7] which focused on the employability of graduate students in the UK. They observed an increase in the number of students enrolled in British universities, and at the same time a surge in graduate unemployment. Moreover, they found that UK graduates were better positioned on the labor market when compared with non-graduates in the UK or with graduates in the EU. They concluded that, contrary to assumptions that skills and personal qualities of graduates determined their

[3] Gary S.Becker, "Investment in Human Capital: A Theoretical Analysis", 1962, https://www.nber.org/system/files/chapters/c13571/c13571.pdf; Gary S.Becker, "Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education(3rd ed.)", 1993, Chicago: University of Chicago Press
[4] Harald Schomburg, "Higher Education and Graduate Employement in Germany", 2000, https://www.jstor.org/stable/1503705
[5] Jose-Gines Mora, José Garcia-Montalvo & Adela Garcia-Aracil, "Higher Education and Graduate Employment in Spain", 2000, https://www.jstor.org/stable/1503709
[6] Esther Duflo, "The Medium Run Effects of Educational Expansion: Evidence from a Large School Construction Program in Indonesia," 2004, Journal of Development Economics, Vol. 74, 163-197, https://www.sciencedirect.com/science/article/abs/pii/S0304387803001846
[7] Marie-Pierre Moreau & Carole Leathwood, "Graduates' employment and the discourse of employability: a critical analysis", 2006, European Journal of Education, https://www.tandfonline.com/doi/abs/10.1080/13639080600867083

success on the labor market, social class, gender, ethnicity, age and disability all had an impact on the opportunities available.

Also in 2006, Julia Varga[8] analyzed for Hungarian graduates, the determinants and effects of pursuing higher education on the labor market. The author also examined the motivating factors for changing fields and the determinants of the decision to specialize further.
Based on a follow-up survey of graduates, the work demonstrated that graduates who had obtained their undergraduate degree in a different field of specialization than the one they would have opted for as their first choice, were the most likely to pursue graduate studies in another field of study. (34.9% of an age bracket, while only 22.9% go deeper into their first field of education). Consequently, when compared to students who did not go beyond undergraduate studies, those who chose to further their education in a different field would, in the short term, lose part of their human capital, representing a loss of income of around 4%. whereas obtaining a higher diploma in the same field leads to an earnings gain of between 13 and 17%. For those who graduated in the same field than their undergraduate degree, earnings would be 13 to 17% higher than those that did not.

In a 2007 article, Plumper and Schneider[9] explored the correlation between university subsidies, the number of university students and unemployment rates in various German Länder. They argued that governments had kept university budgets constant, while increasing the number of students, thus, reducing the budget per student. In addition, they highlighted those states where the budget per student declined, were those where universities were facing drops in academic and administrative quality. Their analyses also revealed that German states with relatively important unemployment experienced the sharpest decline in university spending per student. Moreover, they found that states with lower unemployment were less restrictive in funding their universities. Plumper and Schneider argued that governing entities believed the higher education system to be an effective and seemingly cost-free instrument of regulation of the labor market. The authors concluded that those actions were mere unintended side-effects of what seemed to be a proper solution to youth unemployment.

In 2007, Rosa Dias and Dorrit Posel provided further valuable lessons to the issues of employment, education and skills constraints in post-apartheid South Africa from 1995 to 2003[10]. They demonstrated that, in this context, higher education provided protection against unemployment and that the returns to higher education increased over the period. However, they also showed that these aggregate trends concealed substantial variation between racial groups, within racial groups, and between men and women. Their findings provided only modest evidence of employment growth linked to high skill intensity. They observed that the increase in the formally skilled labour force was significantly greater than the rise in demand for skilled and semi-skilled labour over the same period, and that, as a result, unemployment rates–including among graduates–rose during that time.

[8] Julia Varga, "Why to get a 2nd diploma? Is it life-long learning or the outcome of state intervention in educational choices?", 2006, Budapest Working Paper on the Labour Market, BWP, Institute of Economics, Hungarian Academy of Sciences, https://vmek.oszk.hu/06300/06311/06311.pdf
[9] Thomas Plumper & Christina Schneider, "Too Much to Die, Too Little to Live Unemployment, Higher Education Policies and University Budgets in Germany", 2007, Journal of European Public Policy, https://www.researchgate.net/publication/228177374_Too_Much_to_Die_Too_Little_to_Live_Unemployment_Higher_Education_Policies_and_University_Budgets_in_Germany
[10] Rosa Dias & Dorrit Posel, "Unemployment, Education and Skills Constraints in Post-Apartheid South Africa", 2007, University of Cape Town, Development Policy Research Unit, https://open.uct.ac.za/server/api/core/bitstreams/ae8a7d7a-dd15-4ae2-932f-c35a4e27073f/content

In 2010, Nunez and Livanos[11] released a paper purely based on a quantitative approach, examining the impact of a university degree and the field of study on unemployment in 15 European countries: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Sweden, Spain, UK. Contrary to previous studies, Nunez and Livanos found that in these countries access to higher education increased the chances of getting a job. They argued that, in the long run, higher education would moderately protect from long-term unemployment.

In 2010, Lisa Kahn[12] published a study examining the labour market experiences of white male university graduates in relation to the economic conditions prevailing at the time of graduation in the United States, for the period 1979-1989. She focused on a sample of white men in order to avoid confounding factors such as maternity leave and racial discrimination. The findings of this study strongly support the hypothesis that graduating during an economic downturn has a long-lasting negative impact on individuals' wages. Specifically, individuals who graduated during years of economic growth or stagnation experienced significantly different long-term wage outcomes. An increase of 1% in the general unemployment rate at the time of graduation was found to lead to a reduction in wages of between 4% and 2.5% over an 18-year period, compared to those graduating during years of economic expansion. Moreover, Kahn found that individuals who graduated during periods of medium to high unemployment were twice as likely to be enrolled in further education one year after graduation, compared to those who graduated during years of low unemployment.

Erden and Tugcu[13] used, in 2012, a quantitative approach to explore the correlation between higher education and unemployment in Turkey. They showed that there was a statistically significant common stochastic path between access to higher education and unemployment. Those results demonstrate that, in the long term, more graduates contribute to drive up the general unemployment rate in Turkey. In addition, they found that increased access to higher education also rose unemployment levels in Turkey in the short run, albeit with a smaller impact.

Also in 2012, C. Brett Lockard and Michael Wolf provided an outlook for the projected evolution of the US economy over the 2010 decade[14]. They highlighted that although overall employment in the United States was expected to grow by 14% (from 143 to 163 million jobs), positions typically held by individuals with a high school diploma or less would continue to account for more than half of total employment. This finding serves to put into perspective the effect of educational attainment on employment. Notably, the figure of 163 million corresponds to the actual number of jobs recorded in the United States in 2017 (source: World Bank).

Marla McDaniel and Daniel Kuehn[15] analyzed in 2013 how racial factors in the US affect young people educational achievement, jeopardizing their employability. They most notably showed that white high school graduates were significantly more employed than any other racial groups at the same level on most measurements. They also found that African American high school graduates were, at best, just as employed as white high school dropouts, and at worse, less employed than the latter. African- American high school

[11] Imanol Nunez. & Ilias Livanos, "Higher education and unemployment in Europe: an analysis of the academic subject and national effects", 2009, Higher Education, https://www.jstor.org/stable/40602413
[12] Lisa Kahn, "The long-term labor market consequences of graduating from college in a bad economy", 2010, Labour Economics, https://www.sciencedirect.com/science/article/abs/pii/S0927537109001018
[13] Ekrem Erdem & Can Tansel Tugcu, "Higher Education and Unemployment: A co-integration and causality analysis of the case of Turkey", 2012, European Journal of Education, https://www.jstor.org/stable/23272491
[14] C. Brett Lockard & Michael Wolf, "Employment outlook : 2010-2020 ; Occupational employment projections to 2020", 2012, 135 Monthly Lab. Rev. 84, https://heinonline.org/HOL/LandingPage?handle=hein.journals/month135&div=10&id=&page=
[15] Marla McDaniel et Daniel Kuehn, "What Does a High School Diploma Get You? Employment, Race, and the Transition to Adulthood, The review of black political economy", 2013, Vol. 40, Issue 4, Vol. 40, Issue 4, https://doi.org/10.1007/s12114-012-9147-1.

dropouts were significantly less employed than any other racial group. Finally, the scholars concluded that the improved participation in the labor force associated with high school graduation was higher over time for young African American than their white counterparts.

In their 2014 publication which was based on a study[16] carried out in the US, Damon Clark et Paco Martorell drew a clear distinction between human capital and signaling theories by assessing the marginal return of a high school diploma in terms of income. According to D. Clark and P. Martorell, unlike most educational indicators, such as an additional year of study, a diploma cannot, in itself, affect productivity. Any return in terms of income linked to a diploma must therefore reflect the value of the diploma itself. By using regression discontinuity methods to compare the earnings of employees who narrowly passed their diploma and those who narrowly failed their high school final exams, they found little evidence of any effect of the diploma on employees' situation. Consequently, skills had more of an impact on employment than the diploma itself.

In 2015, Lavrinovicha, Lavrinenko and Teivans-Treinovskisont published an article[17] that examined the impact of education on unemployment rates and income levels among residents of Latvia during the period from 2002 to 2013. The relationship between educational attainment and labour market status, as measured by the unemployment rate, was empirically confirmed through a weak yet statistically significant correlation. The Spearman correlation coefficient stood at -0.275 in 2002 and -0.188 in 2013. Similarly, a statistically significant linear relationship was found between the level of education and household income. In 2013, each incremental level of education attained (no diploma, secondary education, higher education) was associated with an increase of €100 per household member per month.

Diana G. Barbu's 2015 published her doctoral thesis[18], the central question of which was whether national unemployment rates influence university enrolment rates and student success outcomes (retention and completion) in the United States during the period from 1987 to 2010. The empirical study revealed that both factors are affected by fluctuations in the national unemployment rate. A positive relationship was found between the national unemployment rate and undergraduate enrolment, full-time retention, and graduation rates. The study supports this hypothesis by showing that a 1% increase or decrease in the national unemployment rate leads to a corresponding 1.3% increase or decrease in enrolment in public universities in the US. These findings contradict those obtained by Delaney and Doyle in 2011[19], which asserted that no relationship existed between state unemployment rates and enrolment in American universities in the period between 1985-2005.

In the same year, Thierry Kamionka et Xavier Vu Ngoc published an article[20] that assessed the impact of two factors—neighborhood of origin and educational qualifications—on the career trajectory of young French people under 30 who left the education system in 1998. The study demonstrates that factors such as the neighborhood of origin (particularly in cases where housing estates or communes have been identified as

[16] Damon Clark & Paco Martorell, "The signaling Value of a High School Diploma", 2014, Journal of Political Economy Vol. 122 Number 2, https://www.journals.uchicago.edu/doi/abs/10.1086/675238
[17] Ilga Lavrinovicha, Olga Lavrinenko & Janis S. Teivans-Treinovskisont, "Influence of education on unemployment rate and incomes of residents", 2015, Procedia - Social and Behavioral Sciences, https://www.researchgate.net/profile/Olga-Lavrinenko-2/publication/277651400_Influence_of_Education_on_Unemployment_Rate_and_Incomes_of_Residents
[18] Diana Barbu, "The relation between unemployment and college enrolment and success outcomes", 2015, Electronic Theses, The Graduate School, Florida State University
[19] Jennifer Delaney & William R.Doyle, "State spending on higher education: Testing the balance wheel over time", 2011, Journal of education finance, https://www.researchgate.net/publication/236709515
[20] Thierry Kamionka & Xavier Vu Ngoc, "Trajectoire des jeunes sur le marché du travail, quartier d'origine et diplôme : une modélisation dynamique", 2015, Working Papers, Center for Research in Economics and Statistics

sensitive neighborhoods) or participation in job-assistance schemes does not significantly improve access to stable employment for individuals from these neighborhoods, even when controlling for the effects of the diploma level variable. It should be noted that the diploma level variable itself has a relatively limited impact, with less than 26% of the young people in the sample having obtained a diploma higher than the "baccalauréat".

In 2016, Mohd Sahandri Gani Bin Hamzah, Saifuddin Kumar Bin Abdulla, and Mazura Mastura Binti Muhammad published a study[21] on the employability of graduates from **Malaysia's** Polytechnic. The study indicates that the quality of academic training received by graduates of the country's flagship engineering school does not necessarily enhance their employability in the job market relative to graduates of other, less prestigious schools. These latter graduates often possess knowledge that is more readily transferable into professional skills. (The employability of graduates is evaluated at the recruitment stage based on non-technical skills, such as teamwork, communication, and leadership abilities).

In 2017, a report[22] published by Sharanjit Uppal highlighted that for a **Canadian** population under the age of thirty-four with less than a high school diploma, family factors, peer factors, school factors, individual factors, and economic factors can all affect the employability of individuals, independent of any consideration of diplomas, as this population has no diplomas at all.

In 2018, Nigusse Weldemariam Reda and Mulugeta Tsegai Gebre-Eyesus [23] released an article investigating the unemployment rate of graduates in **Ethiopia** between 1999 and 2016. This period saw a remarkable expansion of higher education in Ethiopia in terms of student numbers, academic staff numbers and scientific funding. The authors of this article have quantitatively shown that the expansion of the higher education sector has been followed by an increase in graduate unemployment and that the Ethiopian government has failed to align higher education reform with market demand.

In 2019, Deni Mazrekaj, Kristof De Witte and Sarah Vanteenkiste published a study[24] on the labour market in the Flemish region (Flanders) of **Belgium**. The authors compared labor market positions of high school dropouts with that of high school graduates who did not enroll in higher education. They find sectoral heterogeneity in the returns to upper secondary education, but **no effect of upper secondary education itself.** They do, however, find significant heterogeneity by gender and educational route. While women and those in vocational education may benefit from graduation, male graduates and students (male and female) with general education diplomas may be less well integrated into the labour market than dropouts.

In 2020, Dilrabo Jonbekova published an article[25] showing how students in **Kazakhstan and Tajikistan** are encouraged to seek ever-higher degrees in the face of declining employability, what the author calls "diploma

[21] Mohd Sahandri Gani Bin Hamzah, Saifuddin Kumar Bin Abdulla & Mazura Mastura Binti Muhammad, "The Evaluation of Employment Marketability Connectivity Skills Within Polytechnic Engineering Diploma Students in Malaysia", 2016, US-China Education Review A, Vol. 6, No. 4, p. 230-243

[22] Sharanjit Uppal, "Young men and women without a high school diploma", 2017, Statistics Canada, https://files.eric.ed.gov/fulltext/ED585313.pdf

[23] Nigusse Weldemariam Reda & Mulugeta Tsegai Gebre-Eyesus, "Graduate Unemployment in Ethiopia: the "Red Flag" and Its Implications", 2019, International Journal of African Higher Education, 5(1)

[24] Deni Mazrekaj, Kristof De Witte & Sarah Vanteenkiste, "Labour Market and consequences of a high school diploma", 2018, Applied Economics, Vol. 51, 2019, Issue 21, https://www.tandfonline.com/doi/abs/10.1080/00036846.2018.1543939

[25] Dilrabo Jonbekova, "The diploma disease in Central Asia : student's views about purpose of university education in Kazakhstan and Tajikistan", 2019, Studies in Higher Education, Vol 45, Issue 6, p. 1183-1196 https://www.tandfonline.com/doi/abs/10.1080/03075079.2019.1628199

disease". The study examines students' views on the purpose of university education and its role in their future employability and concludes that socio-economic pressures lead to the evaluation of degrees according to criteria of employment opportunities and improved income. The perceived role of university degrees in employability is declining over time at constant degree levels. Due to an oversupply of graduates and limited job opportunities, employers have raised the bar in terms of qualifications to select job applicants, thus promoting 'diploma disease' with no added employability value.

In 2021, Ming Cheng, Olalekan Adekola, Jo Clarisse Albia and Sanfa Cai published an article[26] that, based on a documentary analysis, highlights the distinction between employability and employment. The study, focused on the United Kingdom, laments the government's transfer of part of the employability issue to higher education institutions. The authors caution against the exclusive use of the employment rate as a key indicator of employability, as it encourages the practice of prioritizing employers' needs over knowledge creation and the development of academic disciplines. Such a dynamic will inevitably lead higher education to evolve in an increasingly profession-oriented way.

In 2021, Katarzyna Cieslik, Anna Barford et Bhaskar Vira published a study[27] on the situation of young people not in employment, education, or training (NEET) in Sub-Saharan Africa, in relation to the sustainable development goal No. 8, target 8.6, which represents a direct commitment to improving the dire situation of youth in the labor market by 2020. The article reviews existing literature on youth employment in the region and provides an analysis of the reasons behind the stagnation of progress. It argues that five myths about youth unemployment and underemployment have hindered understanding and progress. These myths are as follows: (1) education and training systems are flawed, (2) youth micro-entrepreneurship and self-employment are a panacea, (3) the informal sector is part of the problem rather than the solution, (4) care work for family members is equivalent to inactivity, and (5) the agricultural sector has low job creation potential.

In 2023, P. Varsha Pramod, et Remya Ramachandran published a study[28] on youth employment and inclusive growth, focusing on India. Given the rapid growth of the youth population in developing countries, which exacerbates the unemployment crisis and contributes to a cycle of harmful events, inclusive growth (IG) is presented as a cutting-edge development model. This paper specifically explores the concept of inclusive growth, the financial and socio-psychological aspects of youth unemployment, and the means to combat unemployment, such as youth entrepreneurship and skill development in the pursuit of inclusive growth.

In 2023, Noreddine Oumansour et Youb Al Edrissi published a study[29] on active labor market policies in Morocco and their effectiveness in addressing unemployment among young graduates. The article also evaluates the impact of the subsidized employment program "IDMAJ" on job creation, unemployment reduction, and working conditions, particularly in terms of wages and working hours. The findings indicate that the program was ineffective in reducing unemployment and increasing employment, while a negative effect on wages was observed.

[26] Ming Cheng, Olalekan Adekola , Jo Clarisse Albia, Sanfa Cai , “Employability in higher education: a review of key stakeholders perspectives”, https://www.emerald.com/insight/content/doi/10.1108/heed-03-2021-0025/full/html
[27] Katarzyna Cieslik, Anna Barford & Bhaskar Vira, ”Young people not in Employment, Education or Training (NEET) in Sub-Saharan Africa: Sustainable Development Target 8.6 missed and reset”, 2021, Journal of Youth Studies, 25(8), 1126‑1147, https://doi.org/10.1080/13676261.2021.1939287
[28] P. Varsha Pramod & Remya Ramachandran, “Youth employment for inclusive growth: a review and research agenda in global perspective with special reference to India”, 2023, J Glob Entrepr Res 13, https://doi.org/10.1007/s40497-023-00354-4
[29] Noreddine Oumansour & Youb Al Edrissi, “Microeconmetric evaluation of youth employment policies: empirical evidence for Morocco”, 2023, https://revues.imist.ma/index.php/JISELSC/article/view/40715

In 2023, Mohamed Niaré et Ousmane Mariko published a study[30] on the microeconomic determinants of unemployment in the countries of the West African Economic and Monetary Union (UEMOA), taking inactivity into account. The results of this empirical assessment show that certain factors increase the risk of unemployment and inactivity, particularly being a woman, single, young, living with a disability, or residing in an urban area. The study also reveals that although unemployment is higher among educated individuals, they are less likely to be inactive compared to those without education. Furthermore, university-educated women are less exposed to inactivity than their male counterparts but remain more vulnerable to unemployment. Age has little effect on male unemployment, whereas it has a significantly negative impact on female unemployment, with younger women being the most distant from employment. Additionally, the negative effect of higher education on unemployment is more pronounced in rural areas than in major cities. Finally, while disability does not seem to influence unemployment in rural areas, it exacerbates it in urban settings.

In 2023, Nesrine Djellouli and Kahina Ait Hatrit published a study[31] on the issue of unemployment and professional integration of PhD holders in Algeria. The article combines a theoretical exploration to define the scope of the subject with a practical investigation conducted among unemployed PhD holders in Algeria, aiming to identify their job search strategies (targeted sectors and job-seeking behaviour). The results reveal that the majority of respondents primarily target employment in higher education and scientific research.

In 2024, Esther Duflo, Pascaline Dupas, Elizabeth Spelke, and Mark P. Walsh published a longitudinal study[32] over a fifteen-year period on the long-term effects of a scholarship policy granted to disadvantaged children to encourage them to pursue secondary education in Ghana. The social effects are significant for girls, particularly concerning early marriage, infant mortality, and the cognitive impacts on their children. The effects of secondary education on wages are much more mixed; they only appear after more than 12 years of observation but become significant for women (+24 to +30%), while no notable impacts are observed for young men. This recent study, focusing on long-term consequences, shows that average effects may be concealed over long periods, that the relationship between wages and education level is complex, and that gender differences must be taken into account, as well as how access to education (secondary or higher) does or does not affect income…

Conclusion : The literature review thus highlights that studies conducted in different countries, with different levels of economic development, lead to very different conclusions. Sometimes, research work conducted in the same country at different points in time diverge in their final results[33]. The debate is therefore open as to whether and how access to higher education and unemployment rates are related, and in which cases they are and in which cases they are not.

[30] Mohamed Niaré & Ousmane Mariko, "Unemployment in the WAEMU Countries: A Cross-Sectional Data Approach [Le chômage dans les pays de l'UEMOA : Une approche par données transversales]" , https://ideas.repec.org/p/hal/journl/halshs-04313205.html
[31] Nesrine Djellouli & Kahina Ait Hatrit, "La problématique du chômage et de l'emploi des titulaires du diplôme de doctorat en Algérie", 2023, https://asjp.cerist.dz/en/downArticle/160/19/2/231359
[32] Esther Duflo, Pascaline Dupas, Elizabeth Spelke & Mark P. Walsh, "Intergenerational Impacts of Secondary Education: Experimental Evidence from Ghana", 2024, NBER Working Paper No. 32742 Ch 2.2, https://www.nber.org/system/files/working_papers/w32742/w32742.pdf
[33] Diana Barbu, "The relation between unemployment and college enrolment and success outcomes", 2015, Electronic Theses, The Graduate School, Florida State University; Jennifer A.Delaney & William R.Doyle, "State spending on higher education: Testing the balance wheel over time", 2011, Journal of education finance

2.3 Research questions and hypotheses

This monograph, published according to the rules of open science, aims to take a fresh look at this issue, based on a large number of countries (76 to 140, depending on the data available for each item of information in each country), countries for which all the data were available.
This leads us to ask the following five key questions

Q1 -What is the relationship between a country's level of economic development and access to higher education?
To this end, we will investigate the link between Alpha Rate[34] and GDP per capita (PPP), on all observable countries or on smaller populations, removing certain "artefact" countries.

Q2 -Can we establish correlations between economic development and different unemployment rates (general, graduate, non-graduate, youth, young graduates or non-graduates)
For this purpose, we will first examine the distributions of unemployment rates and Alpha Rates in 2017 around the world as a function of GDP per capita (PPP). Then, we will consider the possible monotonic relationships between Alpha Rates and different types of unemployment rates, on the same data.

Q3 - What relationships can be identified between access to higher education and employability, and to what extent are these relationships influenced by the level of economic development?

Q4 - Can countries with comparable levels of wealth and employability be grouped into homogeneous categories, and can their similar behaviors be explained by shared socio-economic factors?
In order to better understand the relationships between Alpha Rates and unemployment rates, different types of clustering are carried out on sets of countries that share certain socio-economic characteristics. This clustering concerns countries with low/high GDP/capita; countries with low/high unemployment rates (general and youth); free hierarchical clustering (unconstrained) according to all variables.

Q5 - In cases where there are significant correlations, particularly between access to higher education and graduate unemployment, how strong are these correlations and can we build predictive models of graduate unemployment rates?
To this end, in cases where these correlations are significant, in particular between the Alpha Rate (see definition below) and the youth unemployment rate, the strength of the relationships is examined, and predictive models of the youth unemployment rate are built to understand the strength of the impact of variations in one variable on the other.

[34] The Alpha Rate is an access rate to higher education as measured by Paxter, it is defined further mathematically

3 Methodology

3.1 Higher education access rate (Alpha Rate)

3.1.1 Measuring students in a comparable manner

A preliminary task involved determining the number of students in each country, a far more complex issue than it may seem, given the variations in the definition of "student" across different countries. To achieve this, available statistics were compiled from sources such as UNESCO (UNESCO, s.d.), national statistics (ministries responsible for higher education, national statistical agencies), and efforts were made to explain significant discrepancies in figures between different data sources. The objective was to establish reliable and comparable figures across countries, serving as a foundation for calculations and constructing representations. Some countries were excluded from the list after extensive analysis because their available statistics appeared too contradictory. Others, despite their demographic significance, had to be excluded due to political instability (conflicts, unrest, etc.), which disrupted data collection. All analyses were conducted using data from 2013, 2015, and 2017, as 2017 was the most recent year with the most comprehensive data available at the start of this study. When data from 2017 were not available, but close years were, linear extrapolations were applied to adjust figures to the 2017 reference. To specifically assess the number of students in a given country, the number of international students on its territory was excluded, while the number of national students studying abroad was added. (MOBILITY, s.d.).

For the sake of consistency across countries, student populations enrolled in short continuing education programs under an administrative student status were excluded. Policies in this regard vary significantly from one country to another, both in terms of training practices and the accounting of learners. In cases where local policies led to major distortions, these discrepancies were noted, and figures were adjusted to ensure the highest possible comparability across countries. For instance, South Korea includes as students those who temporarily suspend their studies for military service (KNSO, s.d.) ; Russia accounts for doctoral students who remain officially enrolled for extended periods while working (ROSSTAT, s.d.) ; Turkey (TUTKSTAT, s.d.) and Morocco (HCP, s.d.) classify older individuals in vocational training as students ; China (NBS, s.d.) changed its methodology in 2014 by incorporating adults in short-term continuing education programs and online short-term training cycles, leading to a sudden 23% increase in student numbers within a year (2013‑2014), equivalent to 8 million additional individuals. To ensure comparability across countries, a "PAXTER correction" was applied, which excludes these various categories of adults from the student count, as is the case in most countries. These corrections related to the definition of "student" affected eight countries (Australia, China, South Korea, Morocco, Russia, Tunisia, Turkey, and Ukraine). The number of students enrolled in initial higher education in a given country depends on its level of development, historical and national policies regarding higher education, but also, of course, on the number of young people of typical student age.

Although statistical data on the number of students in a given country are often abundant and detailed—albeit requiring nuanced interpretation—precise information on the FLOW of young individuals from a specific age cohort entering higher education remains scarce.

Based on these considerations, the PAXTER formula for measuring higher education access in a single country was defined as follows:

$$\textit{Alpha Rate}_{PAXTER} = \frac{\textit{Student count}_{UNESCO} + \textit{National students studying abroad} - \textit{Foreign students studying in the country} + \textit{Correction}_{PAXTER}}{\textit{Native youth population of student age (18 to 22 years old)}}$$

3.1.2 Alpha Rate and higher education access rate (measured in flows)

In the subset of countries for which student flow data were available, we investigated the extent to which the values of the 'Alpha Rate' aligned with observed rates of access to higher education.

The statistical verifications presented below demonstrate that this measure is correlated at 98.95% with the average higher education access rate of a student generation, in countries where this figure (measured as the access to higher education flow of one youth generation was available. The correlation analysis was conducted for 71 countries for the year 2015.

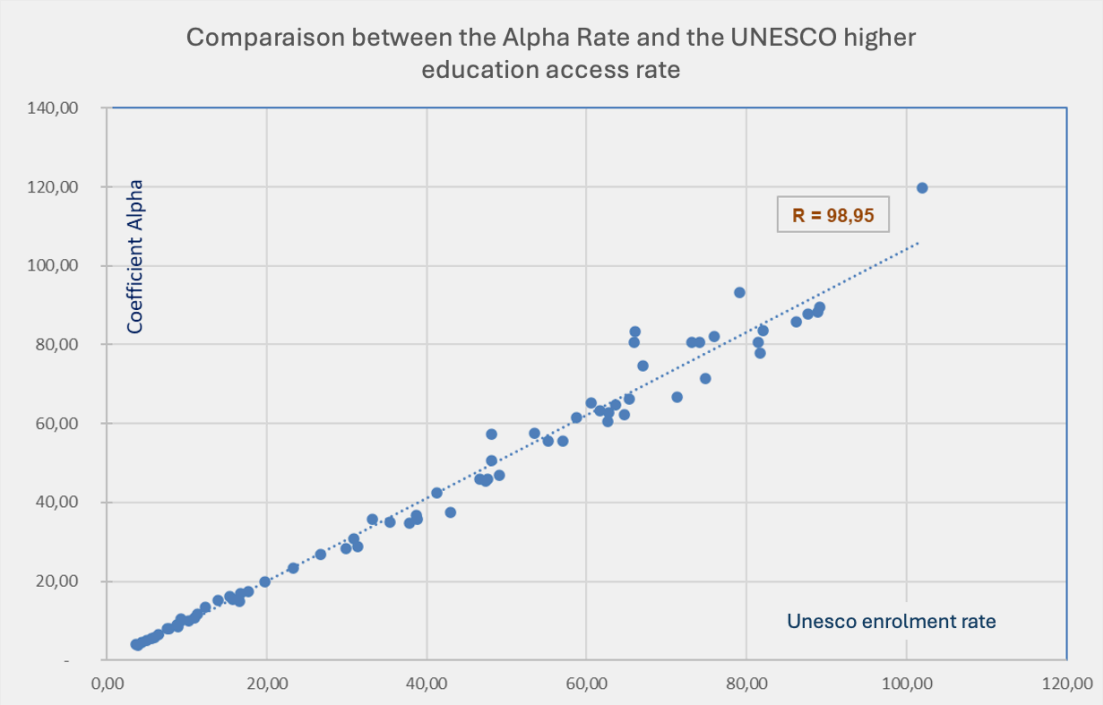


Figure 3-1 Correlation between the UNESCO higher education access rate (flow-based) and Alpha Rate

3.1.3 Conclusive remarks on measuring access to higher education

Given the strength of this correlation, we will systematically use the 'Alpha Rate' as an approximate but very statistically accurate measure, for measuring access to higher education in a given country. Indeed, data availability for estimating this rate is significantly broader, in terms of country coverage, than that for directly measuring student flows.

Moreover, the accurate estimation of this rate depends on the "PAXTER corrections" applied to the local UNESCO-reported student enrolment figures. It might seem presumptuous to adjust official data in this way. However, over nine years, the PAXTER teams have analyzed, in each of the major countries (in terms of student

population), the most evident distortions arising from different methods of counting students. It is likely that some discrepancies have escaped detection. Nevertheless, applying the "PAXTER correction" to student numbers, leading to the Alpha Rate in various countries in a comparable manner, will have had the resulted in revealing many of these distortions, which have significant impacts on reported enrolment figures.

3.2 Unemployment rate

The unemployment rate within a given population is defined as the ratio between the number of unemployed individuals in that population and the total number of active individuals (employed individuals + unemployed individuals) in the aforementioned population. The population for which the unemployment rate is observed can be selected based on various criteria, such as nationality, age, gender, or education level, etc.

The data used in this study come from the ILOSTAT database (ILOSTAT, s.d.). The initial data set contains 448,138 rows, gathering the following information: country, year, gender distribution, age bracket distribution, education level distribution, the number of unemployed individuals for each subcategory, and the corresponding percentage of the total relevant population.

Figure 3-2 presents the unique values appearing in each column of the dataset. The dataset contains unemployment-related data for 170 countries worldwide, spanning a time scale of 34 years, from 1985 to 2019. The dataset is not homogeneous in terms of data availability across the years; it does not include complete information for the entire 35-year period for each country. For instance, in France, 198 lines of data are available for the year 1999, in comparison to 469 recorded lines in 2017.

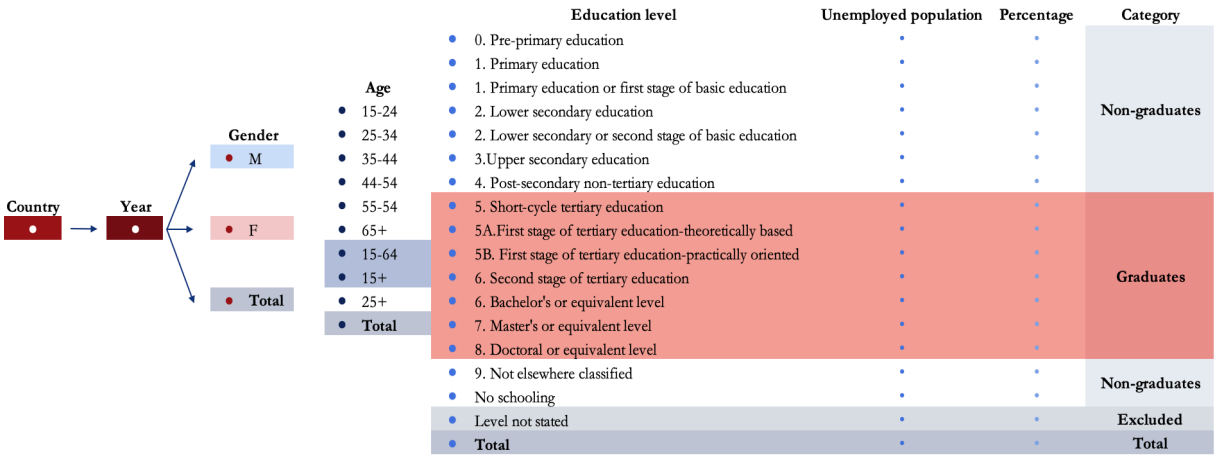


Figure 3-2 Structure of the ILOSTAT data set

The data set includes 10 different education levels across various age brackets. Exploratory analysis revealed that identical data were attributed to three different age brackets: 15-64, 15+, and Total. For the purposes of this study, two primary age brackets have been retained for analysis: ***Youth (15-24 years old)*** and ***Total*** (all ages in the population).

The classification of education levels follows UNESCO's standardized classification, (ISCED, s.d.). There are two UNESCO classifications in our data set, one from 1997 and another used since 2011. This explains the presence

of different labels for the same education level. For example, Figure 3-2 displays various naming conventions for levels 1, 2, 5, and 6. As the objective of this analysis is to assess the strength of the relationship between unemployment rates across different education levels and higher education access rates, we have grouped the previous classifications into three broad categories: Graduates (higher education degree holders), Non-Graduates, and General (entire population). Data classified under the "Level not stated" category were excluded from the analysis. After aggregating employment figures for each subcategory (e.g., USA-1999-Male-Total (age)-Master's or equivalent level is included in the category USA-1999-Male-Total (age)-Graduates) and calculating the resulting new percentages, the data set was structured as follows:

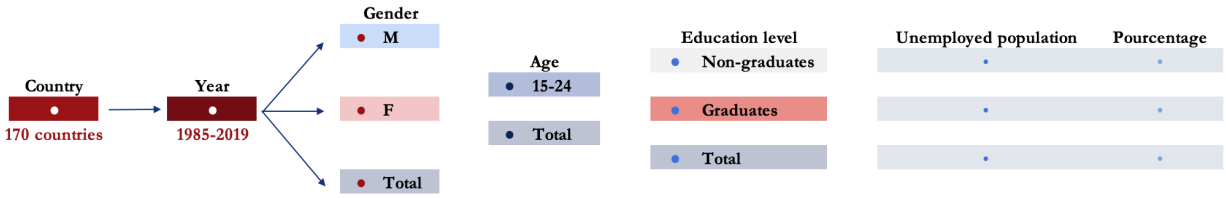


Figure 3-3 Structured data set for analysis

The analysis will consider unemployment rates for the following categories:

- General unemployment rate, defined as the ratio between the total number of unemployed individuals in a country within a given year and the total number of active individuals in that country (all ages, all genders, all education levels).

- Graduate unemployment rate, defined as the ratio between the number of unemployed individuals holding a higher education degree and the number of active individuals with a higher education degree in that country (all ages, all genders, education level: graduates).

- Non-graduate unemployment rate, defined as the ratio between the number of unemployed individuals without a higher education degree and the number of active individuals in that country without a higher education degree (all ages, all genders, education level: non-graduates).

- Youth unemployment rate, defined as the ratio between the total number of unemployed young individuals (aged 15-24) and the total number of active young individuals in that country (15-24 years old, all genders, all education levels).

- Young graduate unemployment rate, defined as the ratio between the number of unemployed young individuals holding a higher education degree and the number of active young individuals with a higher education degree in that country (15-24 years old, all genders, education level: graduates).

- Young non-graduate unemployment rate, defined as the ratio between the number of unemployed young individuals without a higher education degree and the number of active young individuals in that country without a higher education degree (15–24 years old, all genders, education level: non-graduates).

3.3 Economic data

Several economic data sources were examined and used, with the International Monetary Fund (IMF, s.d.), the World Bank (Mondiale s.d.) and the CEPII (CEPII, s.d.). All figures are expressed in 2017 US dollars. When multiple scenarios were available across these sources and studies, median assumptions were retained. In this analysis, GDP per capita (PPP) (Gross Domestic Product adjusted for Purchasing Power Parity) will be the primary indicator used to assess a country's level of economic development. The median GDP per capita (PPP) across 140 countries is $15,000; therefore, countries will be classified as high or low GDP per capita (PPP) depending on whether their level is above or below this value (except in cases where the median of the data is at $20,000).

3.4 Statistical methods

Our analysis will examine the existence of a monotonic relationship between the different variables examined (unemployment rate, access to higher education, GDP per capita (PPP)). The objective is first to determine whether there is a correlation between these variables and then to measure the strength of the correlation. To determine the correlation coefficient between two variables, it is first necessary to assess whether their distributions are Gaussian. If the distributions are Gaussian, the Pearson correlation will be used. If the distributions are not Gaussian, the Spearman correlation will be used. In both cases, it will be necessary to determine whether the correlation is statistically significant. It is possible that two variables are related by chance, and a correlation hypothesis test allows us to determine whether the observed correlation could have emerged randomly.

Thus, two hypotheses will be tested:

H_0 : " The correlation between the variables is 0 " (in other words, there is no correlation)

vs $H_{\neq 0}$: " The correlation between the variables is not 0 " (a correlation exists and needs to be explored)

Student's t-test has been used to determine whether a correlation was statistically significant. One must keep in mind that the higher the number of points in the sample, the lower the minimum value of correlation will be in order for the aforementioned correlation to be statistically significant (example: minimum 0.2 for n=100, minimum=0.28 for n=50).

Correlation Coefficient |r|	Interpretation (if the correlation is significant)
0 <|r|<0.30	Weak degree of linkage
0.30 <|r|<0.50	Moderate degree of linkage
0.50 <|r|<0.70	Noticeable degree of linkage
0.70 <|r|<0.90	Strong degree of linkage
0.90 <|r|<1	Very strong linkage

Table 3-1 Interpretation of the correlation strength, assuming the correlation has been identified as statistically significant

4 General analysis of correlations between economic development level, Alpha Rates and unemployment rates

4.1 Preamble: Background of this study

This study began in 2013, when we started investigating possible links between economic development and the Alpha Rate. This work led us to establish, for 66, then 76 countries representing approximately 90% of the global youth population, a nonlinear model of the higher education access rate as a function of GDP per capita (PPP). The successive refinements of this work resulted in very high correlation rates (R = 0.88), allowing for long-term predictions.

As early as 2014, we were able to indicate that the number of students in 2030 would increase by 30 to 35% compared to 2015 (Tapie, 2014, 2015, 2017), reaching 270 to 280 million, rather than doubling as UNESCO initially projected. In 2019, UNESCO revised its estimates, forecasting a more probable student population of approximately 300 million in 2030, compared to the 400 million initially announced in 2013.
The discovery of a strong correlation between these variables (see Figure 4-1 below), determined at lower values—despite the absence of correlation beyond a certain wealth level—led us to further explore the experimental approach by examining the relationship with employability.

For the first time, all our findings on the relationship between economic development, access to higher education, and employability are presented together in a single body of work.

4.2 Access to higher education and economic development

4.2.1 Alpha Rate distribution

The access rate to education is described by the Alpha Rate, as defined on page 20.

The statistical correlation analysis between the Alpha Rate and the higher education access rate, measured as the student flow of a given age cohort, for the 71 countries where data was available, established a linear correlation with R = 98.95%.

The data analyzed in this study corresponds to the year 2017. Figure 4-1 illustrates the global evolution of the Alpha Rate as a function of GDP per capita (PPP), for the 140 countries for which data is available.

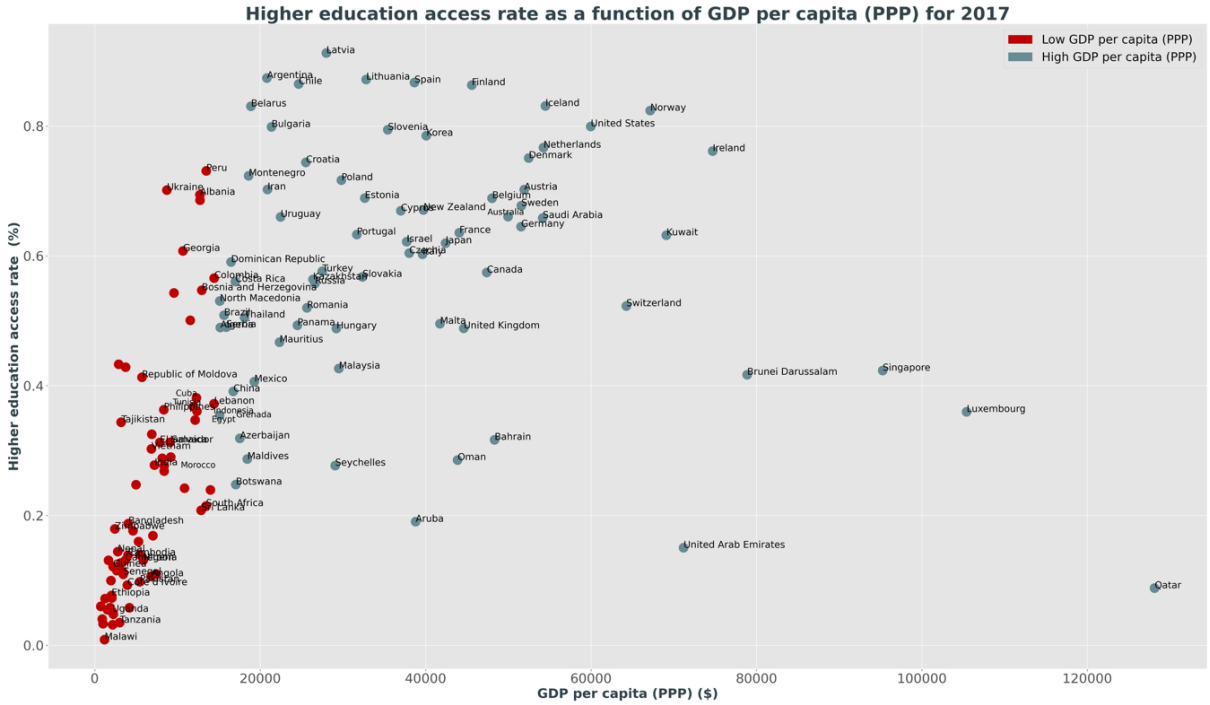


Figure 4-1 Higher education access rate as a function of GDP per capita (PPP) (2017)

- 140 countries (out of 196) are covered in the 2017 Alpha Rate data.
- The Alpha Rate ranges from 0.8% (Malawi) to 91% (Latvia). The median Alpha Rate is 41%.
- For lower Alpha Rate values, countries with a low GDP per capita (PPP) are predominant. A significant relationship between higher education access and GDP per capita (PPP) is also observed. However, countries with a high GDP per capita (PPP) do not necessarily exhibit a high Alpha Rate (Qatar, United Arab Emirates, Luxembourg, Singapore, Switzerland).

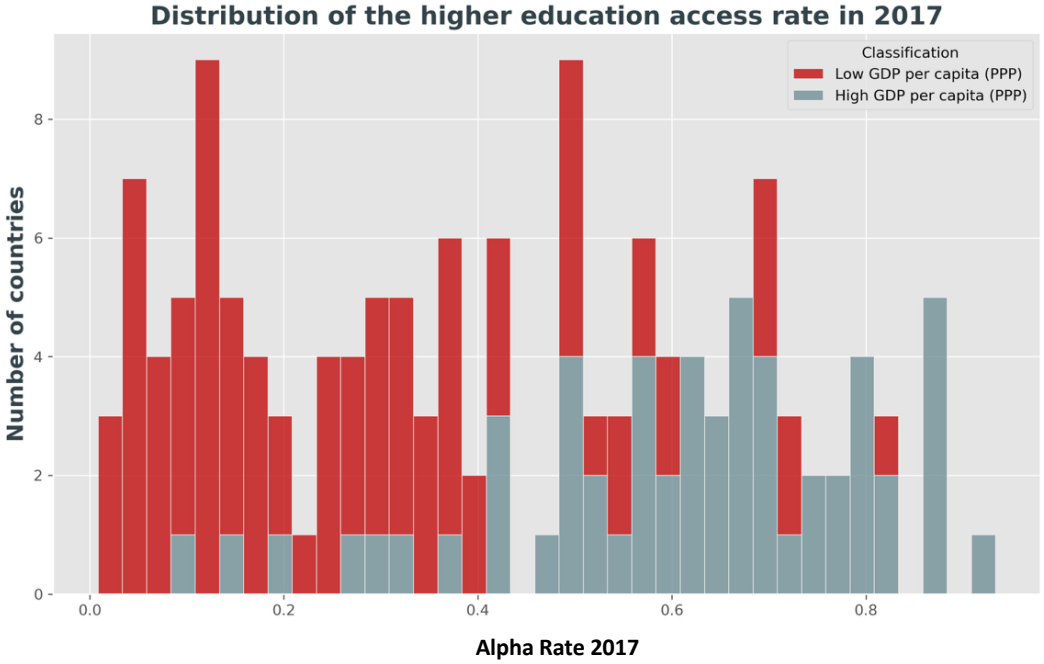


Diagram 4-2 Distribution of the higher education access rate (2017)

- A Shapiro-Wilk test indicates that the distribution of the Alpha Rate does not follow a normal distribution.

- The Alpha Rate and GDP per capita (PPP) are two variables with non-Gaussian distributions. The Spearman correlation between higher education access and GDP per capita (PPP) is significant, with a correlation coefficient of 0.72, leading to the rejection of the hypothesis (H0). This level of correlation implies a strong monotonic relationship between the variables. Under these conditions, an increase in GDP per capita (PPP) leads to a rise in the number of students. Thus, 52% of the Alpha Rate variations can be explained by GDP per capita (PPP) fluctuations (r^2=52 %).

The figure 4-3 below shows that the scatterplot of low-income countries is concentrated around a line, while the data points are more dispersed for high-income countries. Data was extracted for low-GDP per capita (PPP) countries, and their correlation with the Alpha Rate was calculated.

Variables: Alpha Rate and GDP per capita (PPP)

Correlation r	Statistical significance test result (95%)	Sample size
0.72	p=0.00, H_0 rejected	140

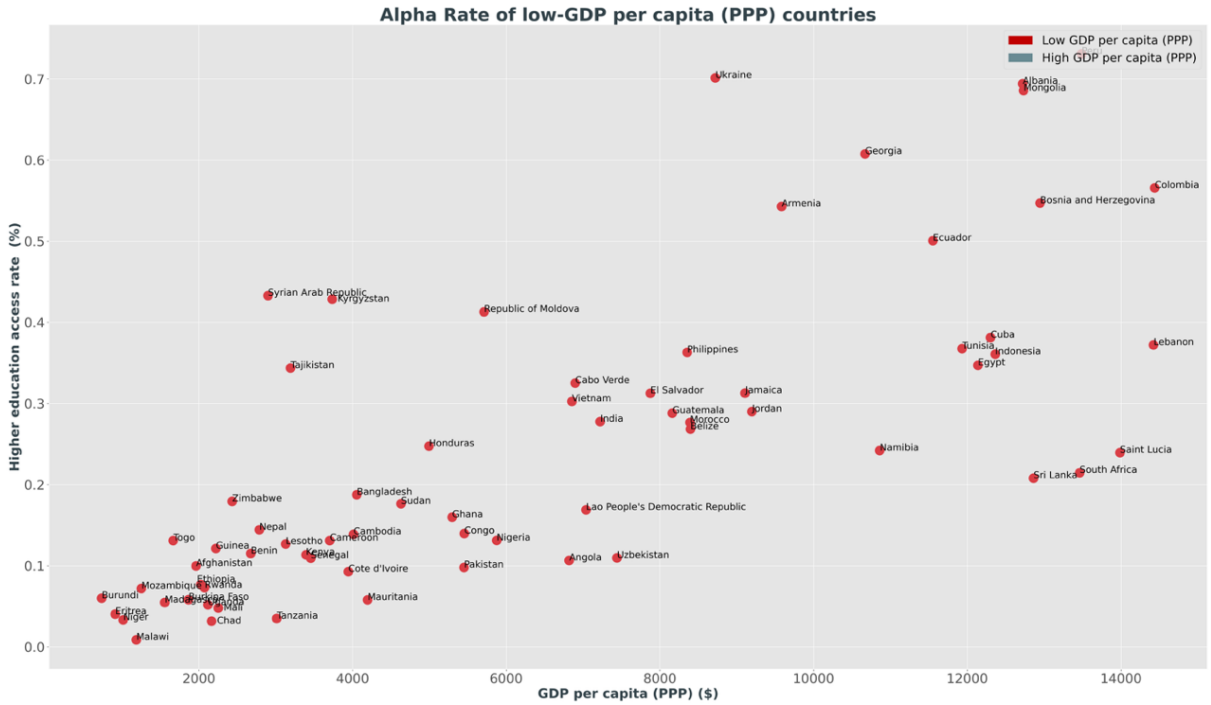


Figure 4-3 Higher education access rate of low-GDP per capita (PPP) countries (under $ 15 000)

Our analysis includes a sample of 69 countries with a GDP per capita (PPP) below $15,000 (the median of this criterion). For this sample, the Spearman correlation is strong (r = 0.78, r^2 = 0.61).

Variables : Alpha Rate et GDP per capita (PPP) (low)

Correlation r	Statistical significance test result (95%)	Sample size
0.78	p=0.00, H_0 rejected	69

Table 4-1 Correlation between higher education access rate and low GDP per capita (PPP)

This indicates that in countries with low economic development, GDP per capita (PPP) and access to higher education are strongly correlated. It can therefore be inferred that economic development facilitates access to higher education by enabling the mobilization of more resources (both public and private). Additionally, higher education likely promotes economic development, as a shortage of skilled labour would hinder further growth. In these countries, 61% of Alpha Rate variations are explained by GDP per capita (PPP) fluctuations. However, this statistical observation must be nuanced by the fact that, at equivalent GDP per capita (PPP), some countries exhibit higher education access rates that vary by a factor of 3 to 9 (examples: Tanzania vs. Syria; Uzbekistan vs. Ukraine; South Africa vs. Colombia). Divergent development models are therefore possible.

The same analysis conducted on the 72 high-GDP per capita (PPP) countries shows no statistically significant correlation in these countries.

However, figure 4-1 illustrates that oil-rich countries, which are very wealthy but have low higher education access rates, significantly distort the curve and thus the correlation. Similarly, certain financial hubs which by definition benefit from substantial economic inflows, exhibit low or very low higher education access rates relative to their wealth.

Variables: Alpha Rate and high GDP per capita (PPP) (high)

Correlation r	Statistical significance test result (95%)	Sample size
0.135	p=0.25, H_0 not rejected	72

Figure 4-4 below shows the correlation studied for countries with a GDP per capita (PPP) greater than $15,000.

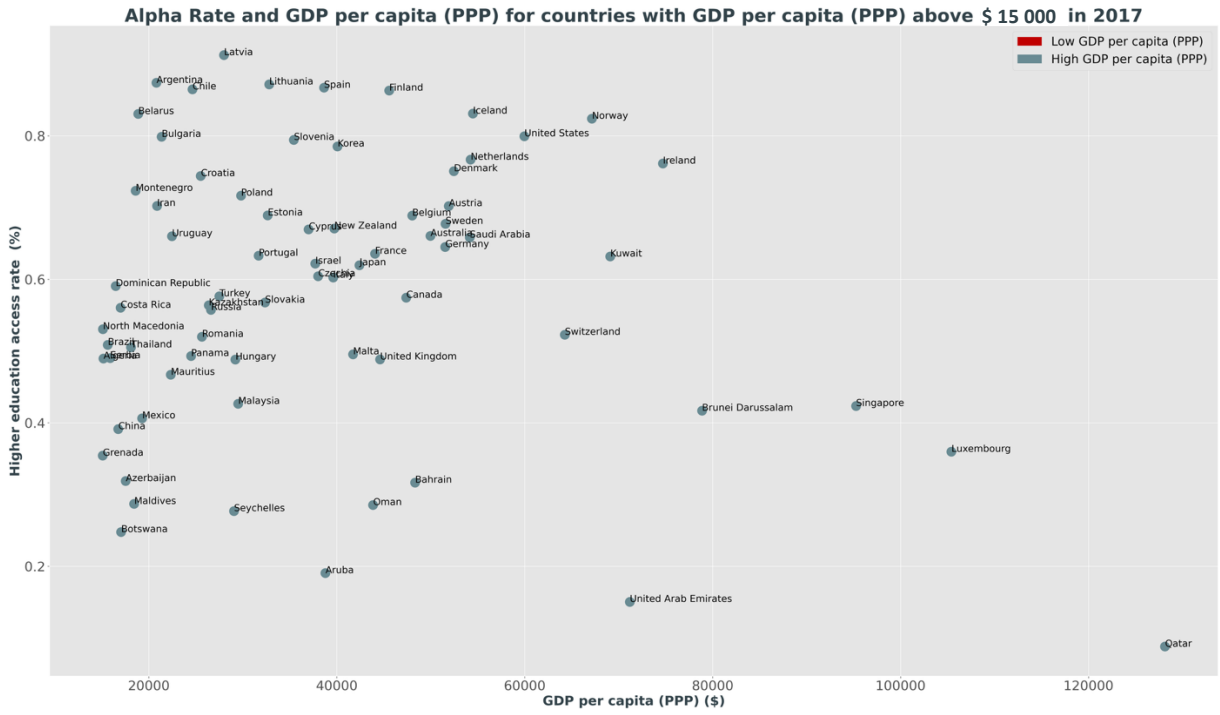


Figure 4-4 Correlation between higher education access rate and GDP per capita (PPP) for countries with GDP per capita (PPP) above $15 000

By removing oil-rich countries (Bahrain, Oman, Brunei Darussalam, UAE, Qatar, Kuwait) and financial hubs (Seychelles, Aruba, Luxembourg), totaling nine countries, the correlation coefficient increases significantly to 0.84 (compared to 0.72) for the remaining 131 countries.

Variables: Alpha Rate and GDP per capita (PPP)

Correlation r	Statistical significance test result (95%)	Sample size
0.84	p=0.00, H_0 rejected	131

Considering the new sample of 63 high-GDP per capita (PPP) countries, the correlation between the Alpha Rate and GDP per capita (PPP) yields a moderate degree correlation of 0.40 across these 63 countries. However, Figure 4-5's shape suggests that this result is primarily driven by transition economies (between $15,000 and $20,000 GDP per capita (PPP)). Figure 4-5 below shows the countries with a GDP per capita (PPP) greater than $20,000, in order to match the threshold used in our studies on unemployment. The observed phenomenon is the same, and even more pronounced (non-significant correlation decreases from 0.135 to 0.06).

When considering only the 49 countries with GDP per capita (PPP) above $20,000, a correlation coefficient of 0.06 is observed. The null hypothesis (H0) is therefore no longer rejected. For countries with GDP per capita (PPP) above $20,000, no correlation exists.

Variables: Alpha Rate and GDP per capita (PPP) > $20,000 (49 countries)

Correlation r	Statistical significance test result (95%)	Sample size
0.06	p=0.58, H_0 not rejected	49

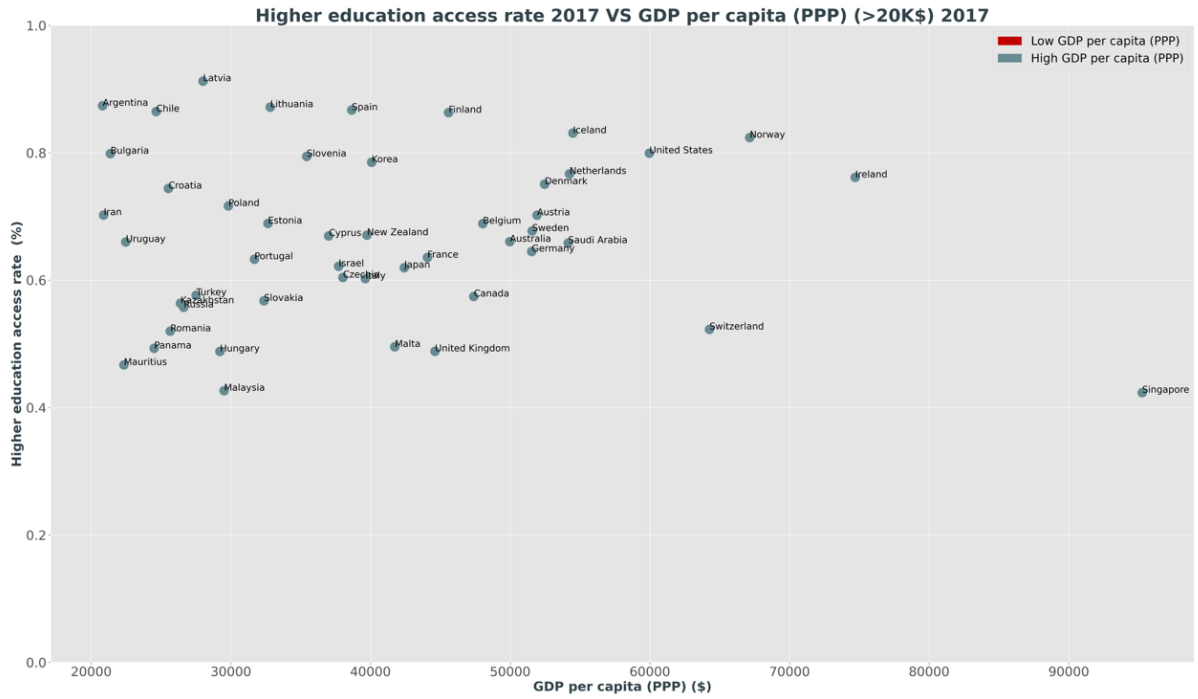


Figure 4-5 Correlation between higher education access rate and GDP per capita (PPP) for countries with GDP per capita (PPP) above $20,000

Therefore, for GDP per capita (PPP) above $20,000, statistically, increasing the number of young people in higher education has no impact on economic development. This conclusion is supported by both statistical results and the presence of significant data points, such as Switzerland, Singapore, or the United Kingdom. It is also notable that at identical education levels, countries as diverse as Bulgaria and the USA are found, despite their GDP per capita (PPP) differing by a factor of 3.
The elements provide a qualified answer to Question Q1.

4.3 Different unemployment rates and economic development levels

We have previously introduced six different types of unemployment rates, which will be addressed in the following analysis. The different levels of observed unemployment and their respective rankings relative to each other result from the economic, social, political, and cultural characteristics of each country.
We will now examine, for each type of unemployment, its relationship with GDP per capita (PPP), the primary indicator of a country's level of economic development.

General unemployment rate data is available for 91 countries (compared to 140 in Section 4.2). This data set represents 3.6 billion workers out of a total global labor force of 5.5 billion, covering 65% of the world's active population. The distribution of GDP per capita (PPP) for this sample is represented in Figure 4-6. It illustrates the heterogeneous level of development of the countries included in our analysis. The median GDP per capita (PPP) is estimated at $21,367.

For this analysis, countries with a GDP per capita (PPP) above $20,000 will be considered highly developed, while those with a GDP per capita (PPP) below $20,000 will be classified as low-development countries.

4.3.1 General unemployment rate

First, we will study the relationship between the general unemployment rate and the level of economic development.

The general unemployment rate is defined as the ratio of the number of unemployed individuals to the total labor force of a given country, across all age brackets (15 years and older) and education levels. Our analysis focuses on the data set for the year 2017. Figure 4-6 below represents the distribution of unemployment rates across countries according to their respective GDP per capita (PPP).

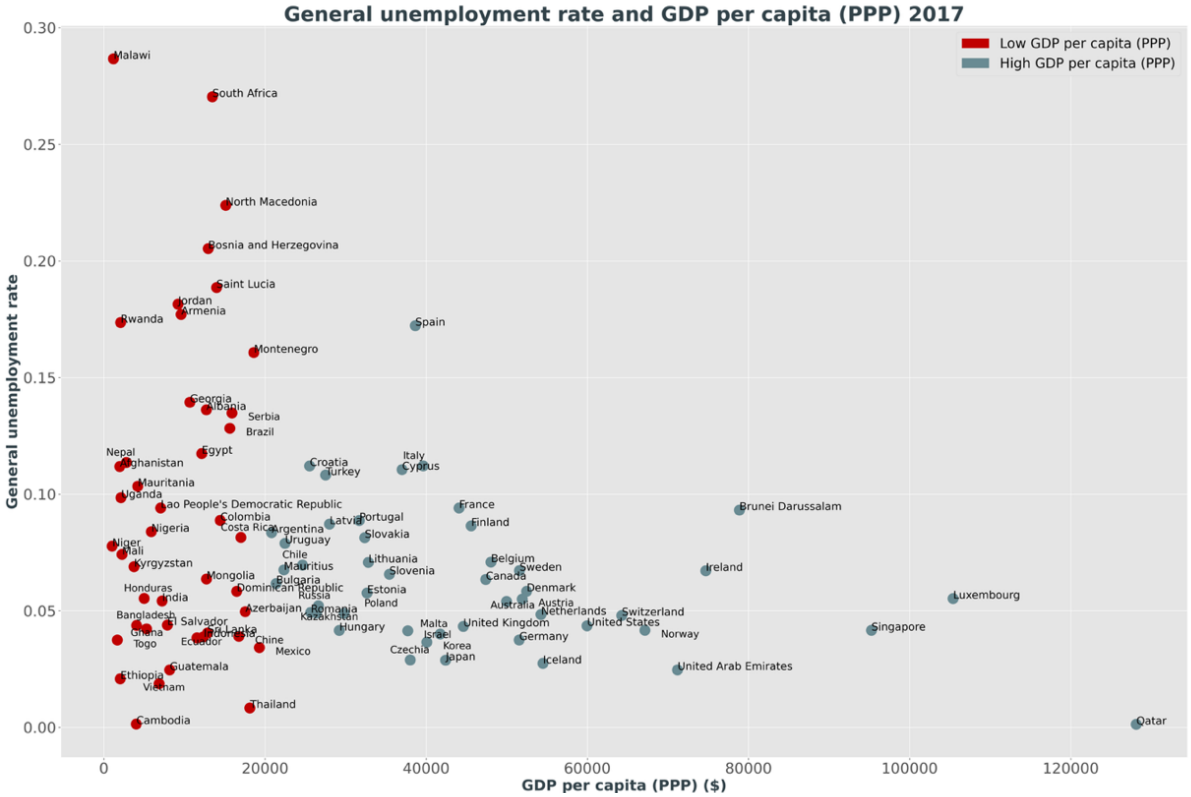
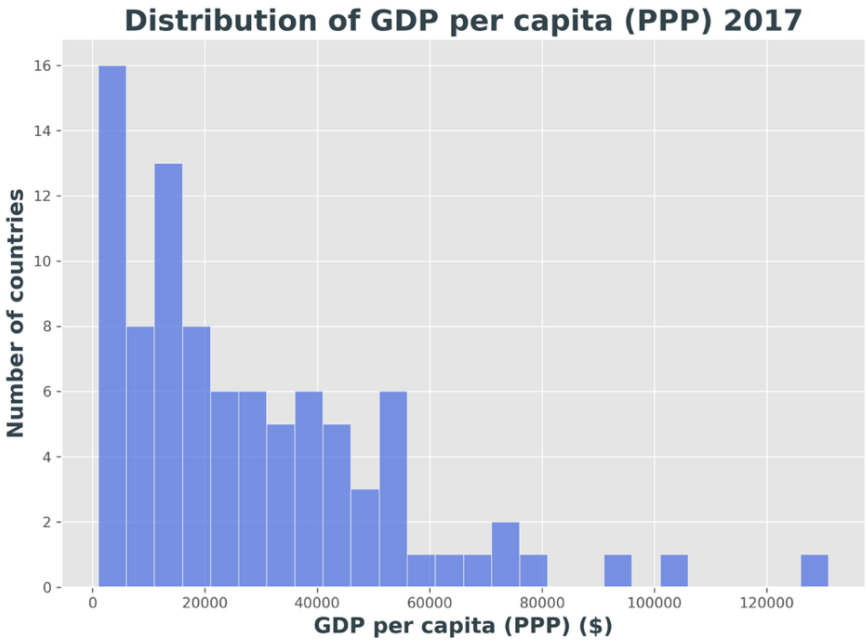


Figure 4-6 General unemployment rate and GDP per capita (PPP) (2017)



Scatterplot 4-7 Distribution of GDP per capita (PPP) (2017)

- Figure 4-7 highlights the wide range of economic development levels.
- Most countries have a GDP per capita (PPP) between $1,000 and $55,000.

- This graphical representation also underscores that the distribution of GDP per capita (PPP) does not follow a normal distribution. Therefore, we will use the Spearman correlation.

Are developed countries less affected by unemployment? If so, to what extent? To address these questions, we will examine the distribution of unemployment rates across 91 countries in 2017. We will also analyze any potential monotonic relationship between unemployment rates and economic development.

Scatterplot 4-5 does not show any concentration around a possible trend line. In other words, the linear relationship between these two variables is not strong. We will attempt to confirm or refute this aforementioned hypothesis while measuring the strength of the monotonic relationship between the two variables. Our calculations will use the scipy.stat function in Python.

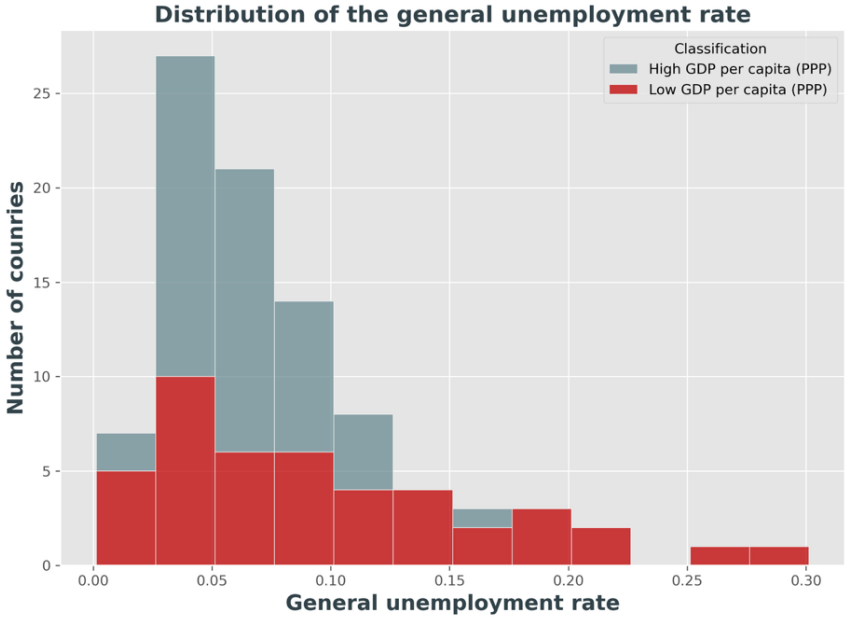


Figure 4-8 Distribution of the general unemployment rate

For the year 2017, the general unemployment rate across the 91 countries ranged between 0.13% and 28.6%. The median unemployment rate was 6%. Furthermore, the distribution of the general unemployment rate is not consistent with a Gaussian distribution. The figure shows that each level of unemployment corresponds to various degrees of economic development. For example, Figure 4-6 indicates that Qatar, the most developed country in our sample, has an unemployment rate close to that of Cambodia, which, by contrast, has one of the lowest levels of economic development.

Such a heterogeneous distribution invites us to examine the correlation between the General unemployment rate and GDP per capita (PPP). Since these two variables do not follow Gaussian distributions, we will measure their Spearman correlation.

Variables: General unemployment rate and GDP per capita (PPP)

Correlation r	Statistical significance test result (95%)	Sample size
-0.253	p=0.016, H_0 rejected	91

Table 4-2 Correlation between the general unemployment rate and GDP per capita (PPP)

The negative monotonic correlation between the general unemployment rate and GDP per capita (PPP) is thus very weak. Indeed, the statistical significance test (at 95%) demonstrates that only 6.40% of the variation

in the general unemployment rate can be explained by a fluctuation in GDP per capita (PPP). If such a correlation exists, our calculations indicate that this dependence is too weak to reveal a clear phenomenon.

4.3.2 Youth unemployment rate

In this section, we will examine the youth unemployment rate. This rate is defined as the ratio of the number of unemployed young people (aged 15-24) to the total number of young active workers in a given country, regardless of education level. Our analysis is based on the 2017 data set. The figure below (Figure 4-9) represents the distribution of the youth unemployment rate as a function of GDP per capita (PPP).

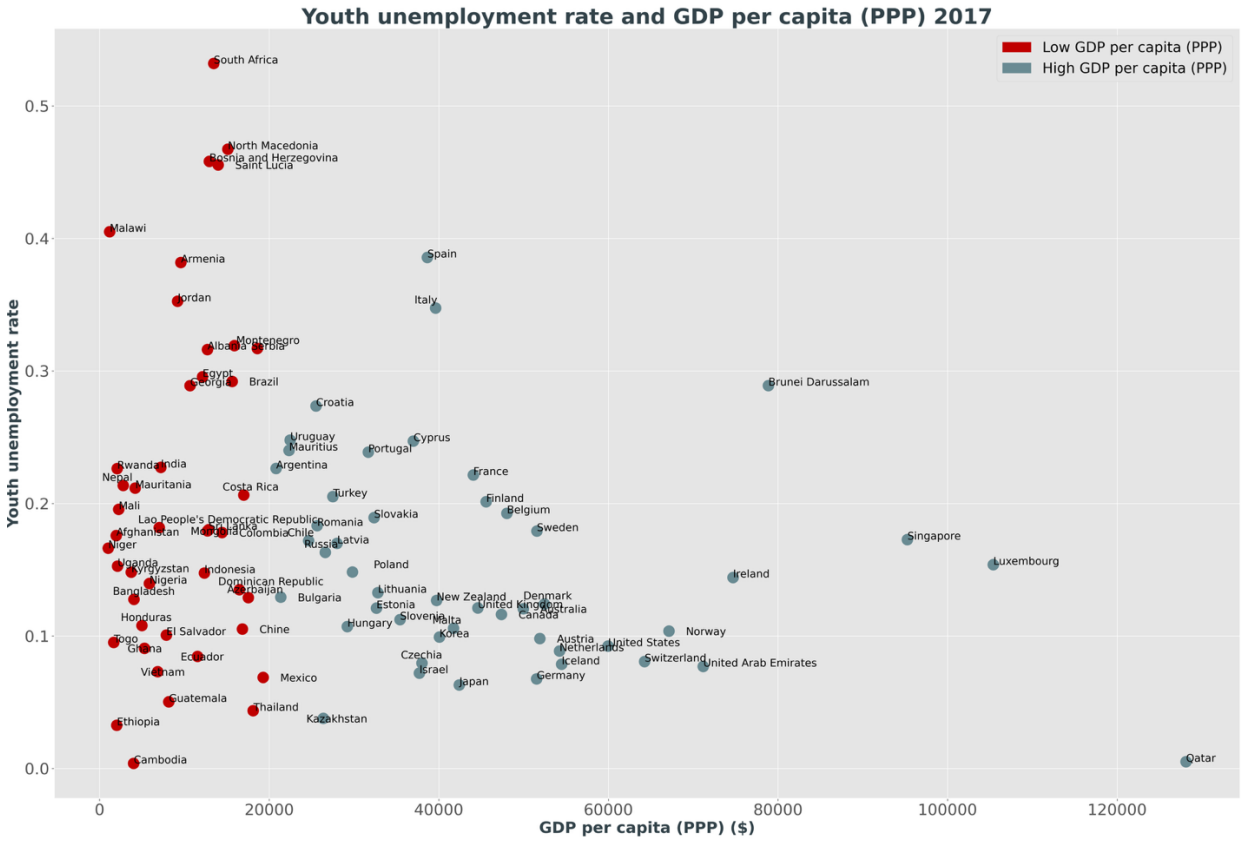


Figure 4-9 Youth unemployment rate and GDP per capita (PPP) (2017)

- Youth unemployment rates are available for 91 countries. This represents 662.2 million young active workers, approximately 55% of the total young active population, which is estimated at 1.2 billion.

- In 2017, the Youth unemployment rate ranged between 0.3% and 53.2%. The median youth unemployment rate is 15.3%, which is significantly higher than the median general unemployment rate of 6%.

In our analysis, we have separated low-development countries (GDP per capita (PPP) below $15,000) from those with a high GDP per capita (PPP). Figure 4-10 illustrates the heterogeneous distribution of youth unemployment rates between developed countries and those with lower levels of economic development.

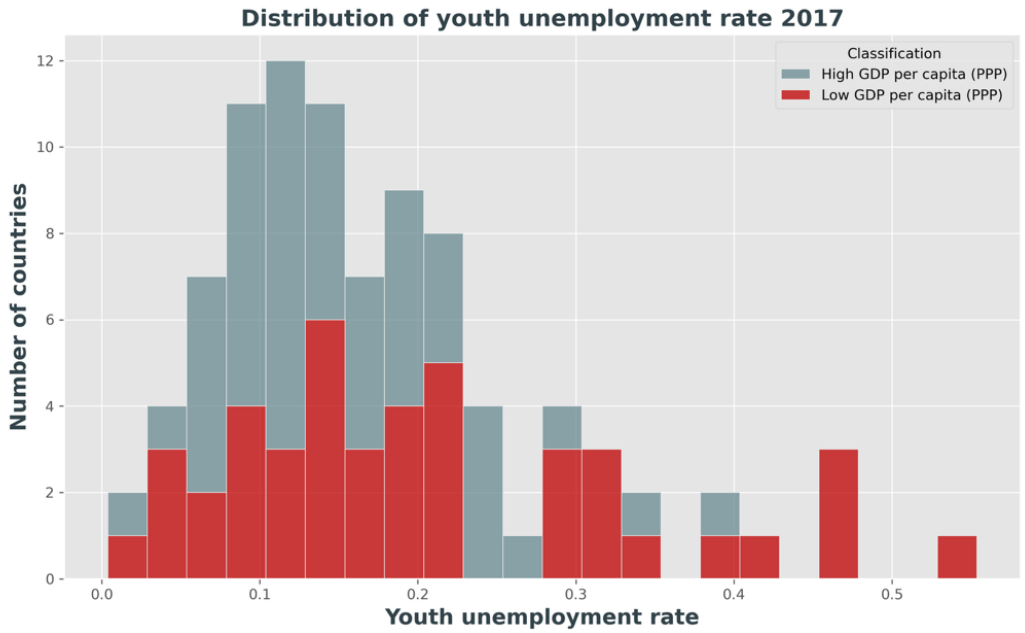


Figure 4-10 Distribution of youth unemployment rates (2017)

- The distribution of the youth unemployment rate is not Gaussian. The majority of countries in the study (with the exception of seven) have a youth unemployment rate below 35%. The median youth unemployment rate, estimated at 15%, is high. 45 countries total have a youth unemployment rate above the median.

- Except for the highest-value ranges, all other youth unemployment rate ranges show a highly heterogeneous composition in terms of the distribution between low- and high-GDP per capita (PPP) countries.

Countries where the youth unemployment rate exceeds the median (15%) exhibit diverse levels of economic development. Thus, a high level of development does not necessarily imply a low youth unemployment rate. We will therefore measure the correlation between the youth unemployment rate and GDP per capita (PPP). As stated in the previous section, GDP per capita (PPP) does not follow a Gaussian distribution. The same applies to the youth unemployment rate. We will therefore calculate the Spearman correlation between these two variables.

Variables: Youth unemployment rate and GDP per capita (PPP)

Correlation r	Statistical significance test result (95%)	Sample size
-0.22	p=0.035, H_0 rejected	91

Table 4-3 Correlation between GDP per capita (PPP) and youth unemployment rate

Thus, a correlation of approximately 22% exists between the two variables. However, this monotonic relationship is weak, as variations in GDP per capita (PPP) account for only 4.8% of fluctuations in the youth unemployment rate. This dependence is too weak to reveal a clear phenomenon.

4.3.3 Graduate unemployment rate

Besides the general unemployment rate and the youth unemployment rate, we will also examine the graduate unemployment rate. This rate is defined as the ratio of the number of unemployed graduates to the total number of graduates in a given country, regardless of age.
As with the general unemployment rate and the youth unemployment rate, we will present the graduate unemployment rate as a function of GDP per capita (PPP).

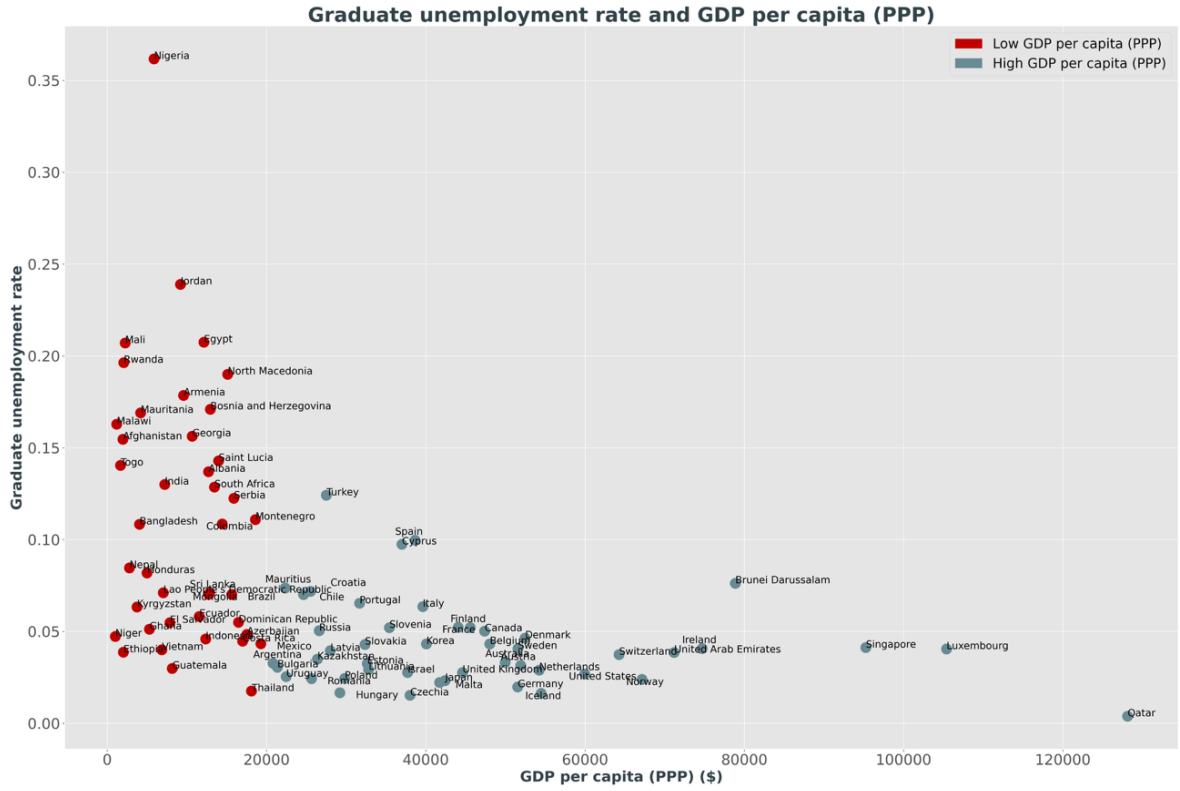


Figure 4-11 Graduate unemployment rate and GDP per capita (PPP)

- Graduate unemployment rate data for the year 2017 is available for 88 countries, covering 558.5 million graduates. The total number of graduates worldwide remains unknown. However, by comparing these 558.5 million graduates to the total number of active workers, estimated at 5.5 billion, we can assert that our sample remains statistically significant.

- The graduate unemployment rate ranges from 0.38% to 36.2%. The median graduate unemployment rate is 5%, which is lower than the median rates of general unemployment and youth unemployment.

Unlike the general unemployment rate and Youth unemployment rate, where all value ranges show a highly heterogeneous distribution between low- and high-GDP per capita (PPP) countries, the upper ranges of the graduate unemployment rate are predominantly concentrated in low-GDP per capita (PPP) countries. This trend is clearly visible in figure 4-11 and is presented in another form in the following figure (figure 4-12).

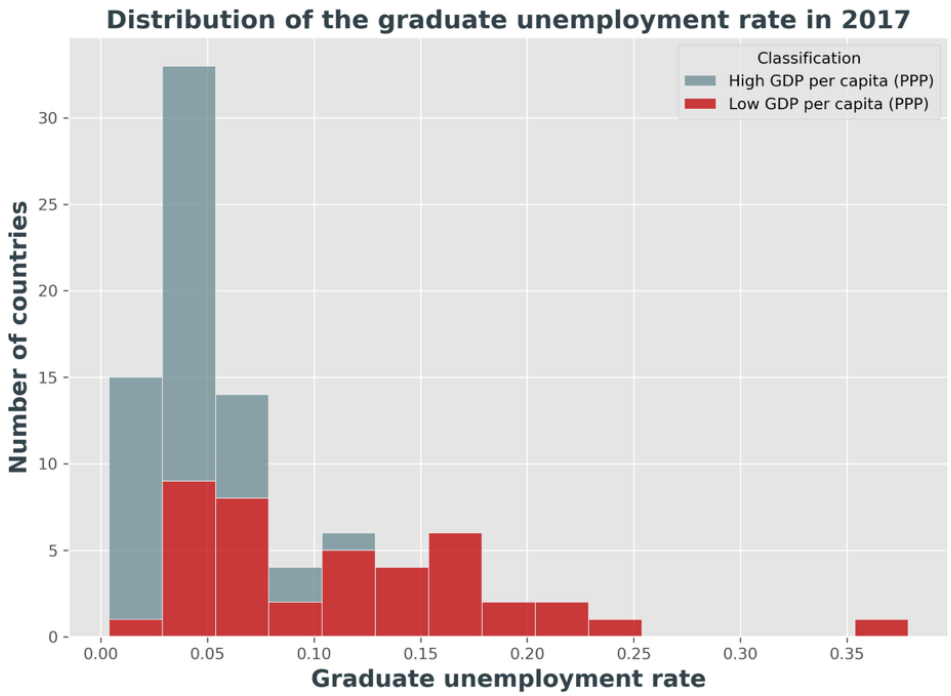


Figure 4-12 Distribution of the graduate unemployment rate (2017)

The graduate unemployment rate does not follow a Gaussian distribution. Most low-GDP per capita (PPP) countries have a graduate unemployment rate higher than the median, which is estimated at 5%. We will now calculate the Spearman correlation between GDP per capita (PPP) and the graduate unemployment rate:

Variables: Graduate unemployment rate and GDP per capita (PPP)

Correlation r	Statistical significance test result (95%)	Sample size
-0.612	p=0.0, H_0 rejected	88

Table 4-4 Correlation between graduate unemployment rate and GDP per capita (PPP)

The graduate unemployment rate is significantly more correlated with GDP per capita (PPP) than the previously considered unemployment rates. Given that this correlation is negative, an increase in GDP per capita (PPP) implies a decrease in the graduate unemployment rate. More precisely, 36% of the variations in the graduate unemployment rate can be explained by fluctuations in GDP per capita (PPP). Thus, while a country's level of development rises, its graduate unemployment rate decreases, whereas this trend has almost no effect on the general unemployment rate or the youth unemployment rate.

4.3.4 Young graduate unemployment rate

The young graduate unemployment rate is defined as the ratio of the total number of unemployed young graduates to the total number of young graduates (aged 15-24) in a given country. The figure below (figure 4-13) illustrates the relationship between the young graduate unemployment rate for the year 2017 and GDP per capita (PPP).

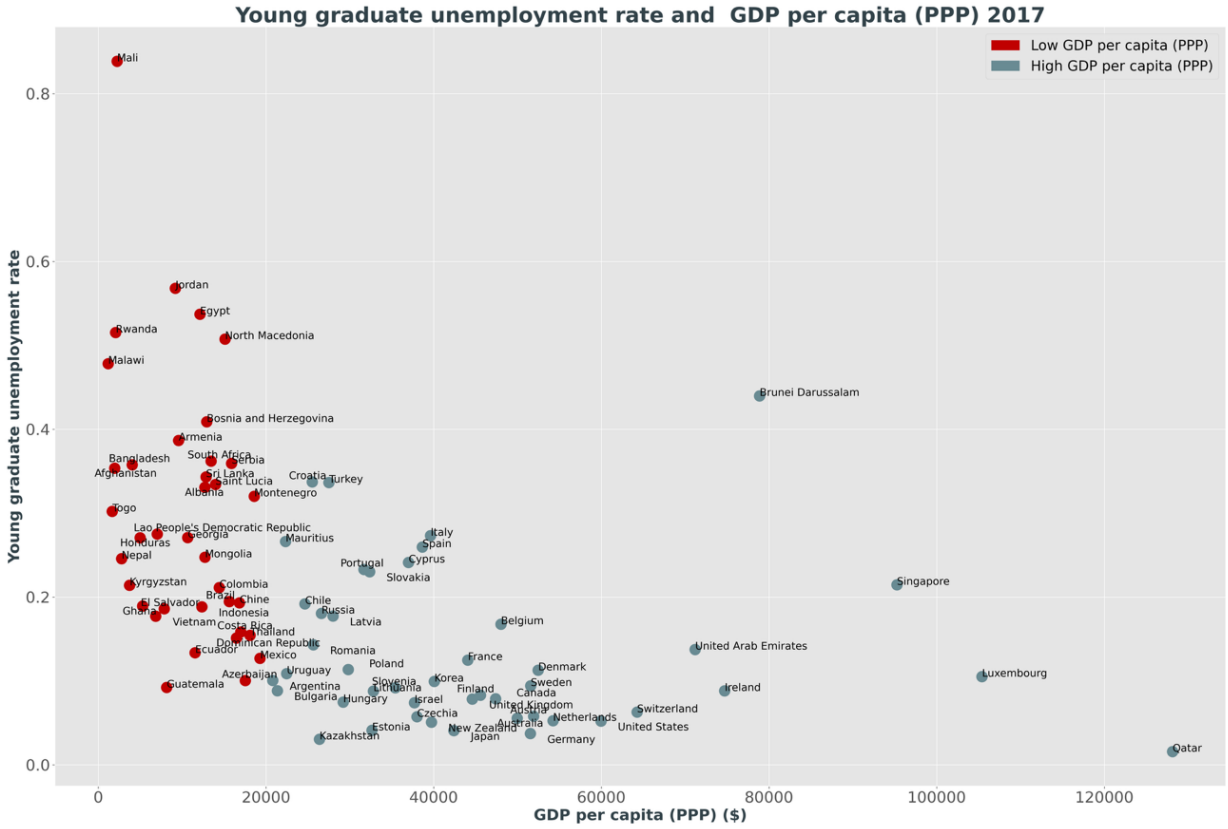


Figure 4-13 Unemployment rate of young graduates and GDP per capita (PPP) (2017)

- Young graduate unemployment rate data for the year 2017 is available for 82 countries, covering 25 million young graduates. Additionally, young graduates from China should be considered, for whom the total number is unknown, but for whom we have the unemployment rate (19.3%).

- The young graduate unemployment rate ranges from 1.57% to 83.8%. The median unemployment rate of young graduates is 17.7%. Notably, this rate is closer to the median youth unemployment rate (15.3%) than to the median graduate unemployment rate (5%).

As in the previous section, we will examine the distribution of young graduate unemployment rates among countries with low and high GDP per capita (PPP). The figure below (Figure 4-14) highlights the fact that highly developed countries tend to have relatively low young graduate unemployment rates.

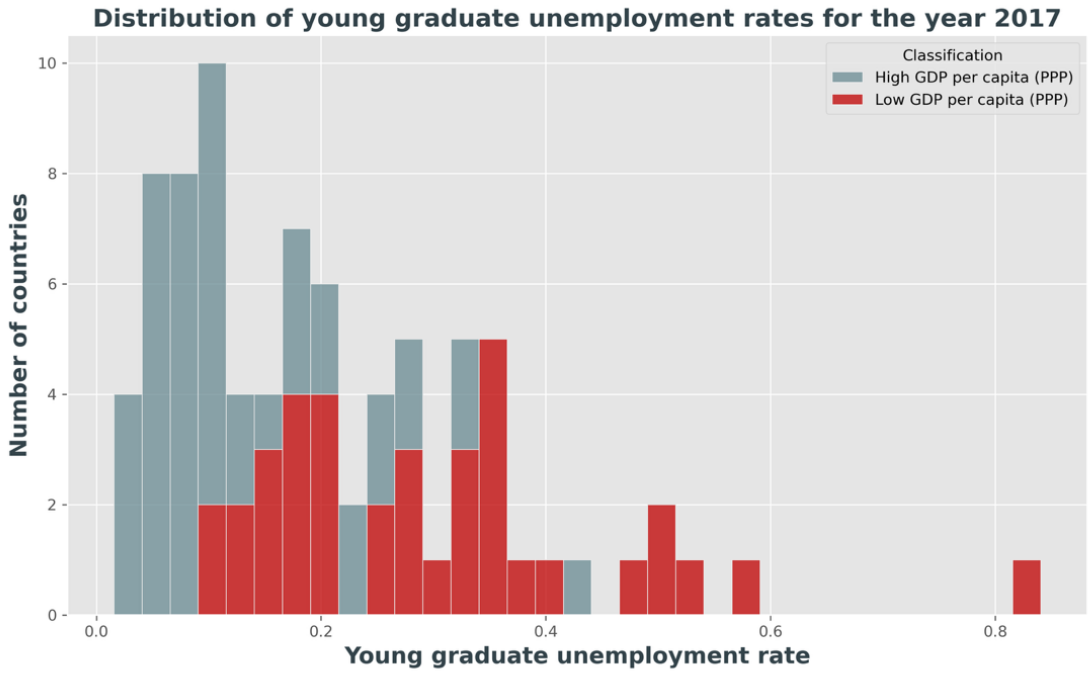


Figure 4-14 Distribution of young graduate unemployment rates (2017)

- The distribution of young graduate unemployment rates is not Gaussian.
- It is observed that only two low-GDP per capita (PPP) countries have young graduate unemployment rates below the median.

We will now measure the Spearman correlation between the young graduate unemployment rate and GDP per capita (PPP). The result is presented in the table below:

Variables: Young graduate unemployment rate and GDP per capita (PPP)

Correlation r	Statistical significance test result (95%)	Sample size
-0.641	p = 0.0, H_0 rejected	82

Table 4-5 Correlation between the young graduate unemployment rate and GDP per capita (PPP) (year 2017)

The correlation between the variables is quite significant. Indeed, 41% of the variations in the young graduate unemployment rate can be explained by fluctuations in GDP per capita (PPP). It is also observed that GDP per capita (PPP) has a greater impact on the graduate unemployment rate and the young graduate unemployment rate than on the general unemployment rate and the youth unemployment rate.

4.3.5 Non-graduate unemployment rate

The non-graduate unemployment rate is defined as the ratio of the number of unemployed non-graduates to the total number of active non-graduates. Previously, we observed that the monotonic relationship between the graduate unemployment rate and GDP per capita (PPP) was significant. We will now examine the relationship between the non-graduate unemployment rate and the level of economic development. Figure 4-15 below represents the distribution of the non-graduate unemployment rate worldwide.

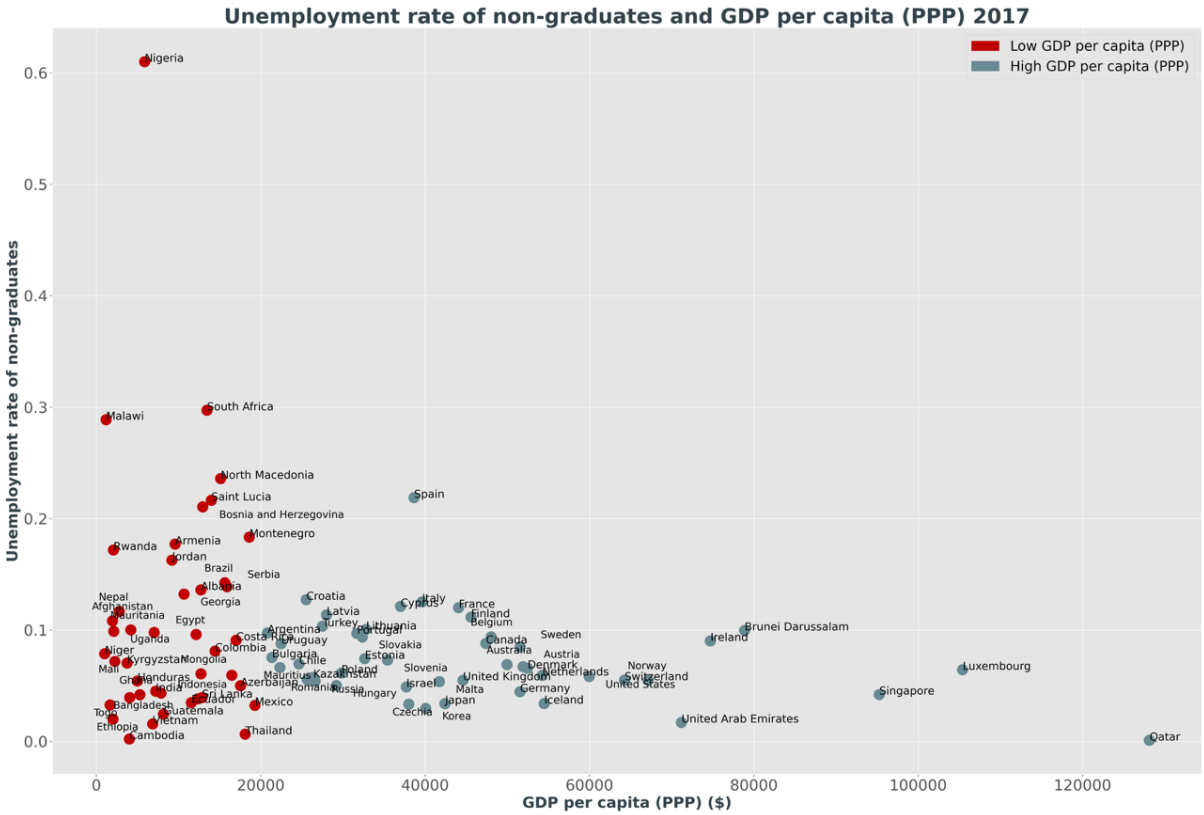


Figure 4-15 Unemployment rate of non-graduates and GDP per capita (PPP)

- The figure presents data from 90 countries. The sample covers a total of 1.89 billion non-graduates.
- The non-graduate unemployment rate ranges from 0.09% to 61%. While Nigeria records the highest non-graduate unemployment rate, the rest of the countries in our sample have rates below 30%.
- In figure 4-16 below, it is observed that unemployment rate brackets are highly heterogeneous in terms of the distribution between countries with low and high GDP per capita (PPP).

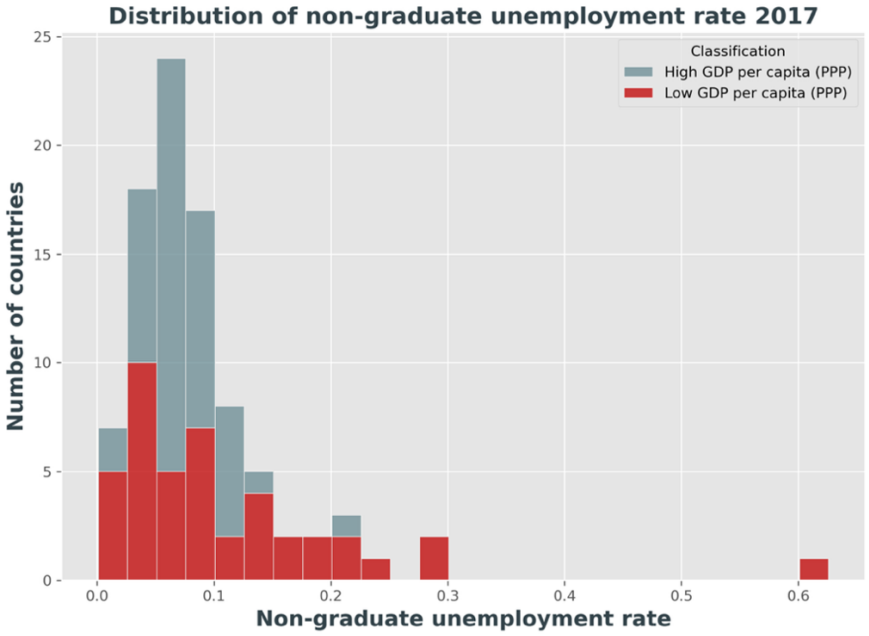


Figure 4-16 Distribution of non-graduate unemployment rate

- The distribution of non-graduate unemployment rates is not Gaussian.

- The median unemployment rate of non-graduates is 7.5%, compared to 5% for graduates and 6% across all education levels. It is also observed that the unemployment rate brackets on both sides of the median are highly heterogeneous in terms of the level of economic development.

Variables: Non-graduate unemployment rate and GDP per capita (PPP)

Correlation r	Statistical significance test result (95%)	Sample size
-0.145	p=0.173, H_0 not rejected	90

Table 4-6 Correlation between non-graduate unemployment rate and GDP per capita (PPP)

Our calculations do not establish a monotonic relationship, even a weak one, between the non-graduate unemployment rate and GDP per capita (PPP). In other words, variations in the non-graduate unemployment rate are not affected by changes in GDP per capita (PPP).

4.3.6 Young non-graduate unemployment rate

The young non-graduate unemployment rate is defined as the ratio of the number of unemployed young non-graduates (ages 15–24) to the total number of young non-graduates. Figure 4-17 below represents the distribution of young non-graduate unemployment rates across 88 countries in 2017.

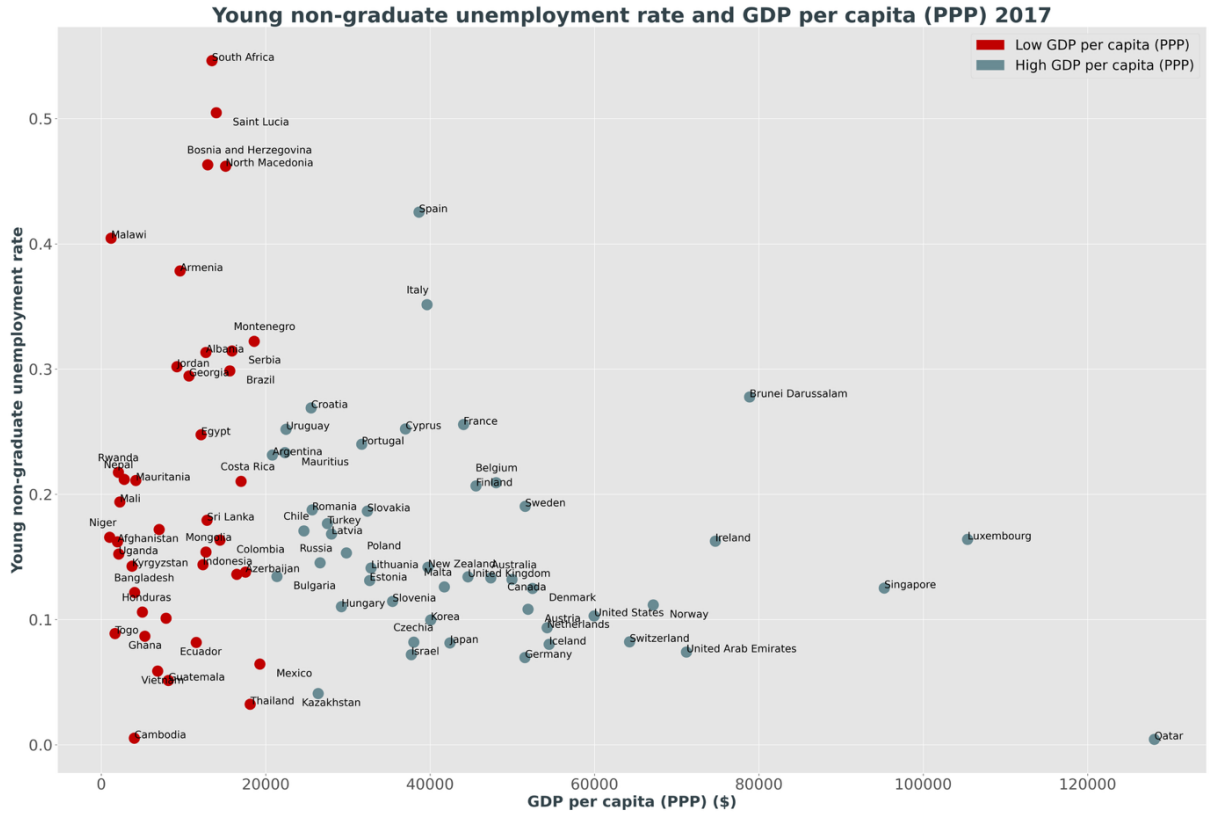


Figure 4-17 Young non-graduate unemployment rate and GDP per capita (PPP) (2017)

- Our sample thus covers 154.3 million young non-graduates.

- The young non-graduate unemployment rate ranges from 0.4% to 54%. The median value of this variable is 15.4%, which is closer to the median youth unemployment rate (15.3%) and the median young graduate unemployment rate (17.7%) than to the median non-graduate unemployment rate (7.5%).

- The different unemployment rate brackets (values on the x-axis) for young non-graduates are highly heterogeneous in terms of their distribution between countries with low and high GDP per capita (PPP), as observed in the following figure. This suggests a weak collinearity between GDP per capita (PPP) and this unemployment rate, which will be verified.

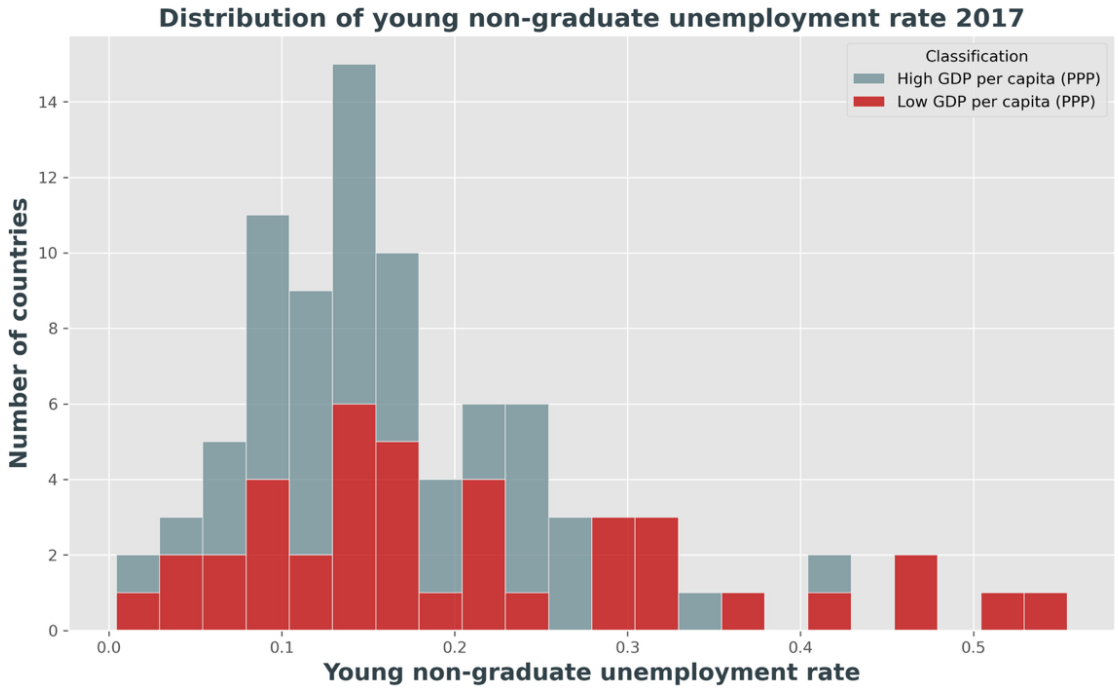


Figure 4-18 Distribution of young non-graduate unemployment rate

- The young non-graduate unemployment rate does not follow a Gaussian distribution.

Variables: Young non-graduate unemployment rate and GDP per capita (PPP)

Correlation r	Statistical significance test result (95%)	Sample size
-0.225	p = 0.035, H_0 rejected	88

Table 4-7 Correlation between young non-graduate unemployment rate and GDP per capita (PPP)

While a monotonic relationship exists between the two variables, it remains extremely weak: only 5% of the variations in the Young non-graduate unemployment rate can be explained by fluctuations in GDP per capita (PPP).

Thus, a developed country such as France may record a high unemployment rate among young non-graduates (26%), whereas countries like Ghana or Togo may exhibit rates below 10%.

4.3.7 Summary of results and conclusions

Our previous analysis aimed to answer the following research questions:

- Q1: What relationship can be observed between a country's level of economic development and access to higher education?
- Q2: Can correlations be established between access to higher education and different unemployment rates (general, graduates, youth, and young graduates) across countries at different stages of economic development?

Gross Domestic Product per capita (Purchasing Power Parity, GDP per capita (PPP)) is a conventionally used indicator for measuring a country's level of economic development.

The following table summarizes all our results, where r_s represents the Spearman correlation coefficient:

Types of unemployment rates examined	**General unemployment rate**	**Graduate unemployment rate**	**Non-graduate unemployment rate**	**Youth unemployment rate**	**Young graduate unemployment rate**	**Young non-graduate unemployment rate**
GDP per capita (PPP)	**r_s = -0.253** **r_s^2 = 0.064** H_0 rejected Sample: **91**	**r_s = -0.612** **r_s^2 = 0.375** H_0 rejected Sample: **88**	**r_s = -0.145** **r_s^2 = 0.021** H_0 not rejected Sample: **90**	**r_s = -0.22** **r_s^2 = 0.048** H_0 rejected Sample: **92**	**r_s = -0.641** **r_s^2 = 0.411** H_0 rejected Sample: **82**	**r_s = -0.225** **r_s^2 = 0.051** H_0 rejected Sample: **88**

Table 4-8 Summary of correlations between unemployment rates and GDP per capita

It is observed that all Spearman correlations (which measure the strength of a monotonic relationship between variables) are negative. With the exception of the correlation between GDP per capita (PPP) and the non-graduate unemployment rate, all correlations are also statistically significant. However, for four of them, the relationship remains very weak, with an r^2 coefficient below 6%.

Since these correlations are negative, it follows that an increase in GDP per capita (PPP) leads to a decrease in various unemployment rates. The magnitude of this decline depends on the type of unemployment considered.

Thus, 37.5% of the graduate unemployment rate and 41% of the young graduate unemployment rate can be explained by variations in GDP per capita (PPP). In contrast, for the general unemployment rate, the youth unemployment rate, and the Young non-graduate unemployment rate, only 6.4%, 4.8%, and 5%, respectively, can be attributed to fluctuations in economic development, which is extremely weak.

General unemployment, youth unemployment, non-graduate unemployment and young non-graduate unemployment are, in reality, not dependent on a country's level of economic development. While countries with very high unemployment rates are primarily poor countries (Figure 4-6), it is also observed that some low-income countries experience low unemployment rates, while certain high-income countries exhibit significant unemployment. Under these conditions, other social, political, or cultural variables may influence unemployment levels.

Policymakers often assert that a high level of qualification effectively protects against unemployment. Following the observations in Chapter 4.3, the next section, Chapter 4.4, will further explore the impact of obtaining a degree on employability within a given country, whereas the previous analysis compared different countries to one another.

4.4 The effect of graduating on access to employment

4.4.1 Detailed analysis of 2017 figures

In the previous chapter (4.3), we established that only moderate monotonic relationships linked the graduate unemployment rate and the young graduate unemployment rate with the level of economic development. Furthermore, we also demonstrated how age was a variable that also impacted the observed level of unemployment. The following table presents the medians associated with each type of unemployment rate:

	General unemployment rate	Graduate unemployment rate	Non-graduate unemployment rate	Youth unemployment rate	Young graduate unemployment rate	Young non-graduate unemployment rate
Median	0.06	0.05	0.075	0.153	0.177	0.154

Table 4-9 Summary of the medians of the different observed unemployment rates (2017)

It can be observed that the **medians of the youth unemployment rate, young graduate unemployment rate, and young non-graduate unemployment rate are closer to each other** than the medians of the young graduate unemployment rate and the graduate unemployment rate. The effect of age thus appears to be a determining factor.

The median youth unemployment rate is 2.5 times higher than the general unemployment rate. That of non-graduates is 20% higher than the median general unemployment rate. Finally, the median young graduate unemployment rate is three times higher than the median general unemployment rate. This result has a so-called "mechanical" component. Indeed, when searching for their first job, all young people are actively seeking employment, whereas only a small proportion of adults (most of whom are employed) are looking for work at any given time. Therefore, it is not abnormal that the ratios of youth unemployment rates to general unemployment rates are significantly greater than one. Qualified individuals take longer to find a job than non-qualified ones, which can partly be explained by the higher expectations of skilled workers.
In this section, we will compare the **graduate unemployment rate with that of non-graduates, systematically within the same country.** The gap between these two variables will allow us to measure the impact of education on employment access in a given country.
Considering different age brackets and education levels, our analysis will focus on two types of populations:

- (G) The first population consists of all age brackets, using education level as a differentiation criterion.
- (J) The second population consists only of young people aged 15 to 24, when taking education level as a differentiation criterion.

We used the available data for the six different types of unemployment rates for the year 2017.
We represented the impact of education on employment access using arrows. The origin of these arrows represents the non-graduate unemployment rate, while their endpoint indicates the graduate unemployment rate. Blue arrows correspond to countries where a degree facilitates employment access, meaning countries where the graduate unemployment rate is lower than that of non-graduates (blue arrows pointing downward). Red arrows symbolize countries where a degree negatively impacts employability (red arrows pointing upward).

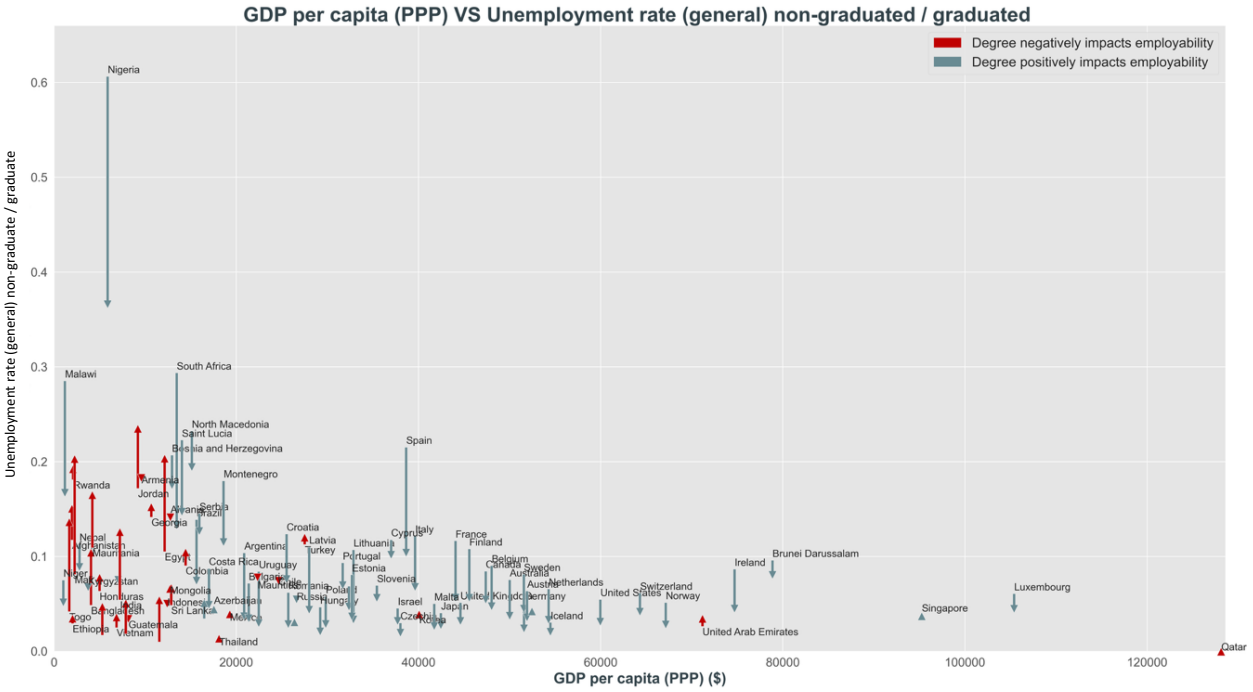


Figure 4-19 Representation of the impact of holding a higher education degree on unemployment as a function of GDP per capita (PPP) for population G (2017)

Population (G) includes data from 88 countries, representing a sample of 1.9 billion non-graduates and 558.5 million graduates, accounting for 44.5% of the global labor force.

It is observed that in 33 countries, holding a degree negatively impacts employability, while in the remaining 55 countries, it facilitates employment access. Additionally, Figure 4-19 clearly highlights that, on average, a degree is an advantage in wealthy countries and a disadvantage in poorer ones. Since these observations contradict commonly accepted assumptions, we will conduct a more detailed analysis of the 33 countries where obtaining a higher education degree negatively impacts employability. These 33 countries are represented in the following figure:

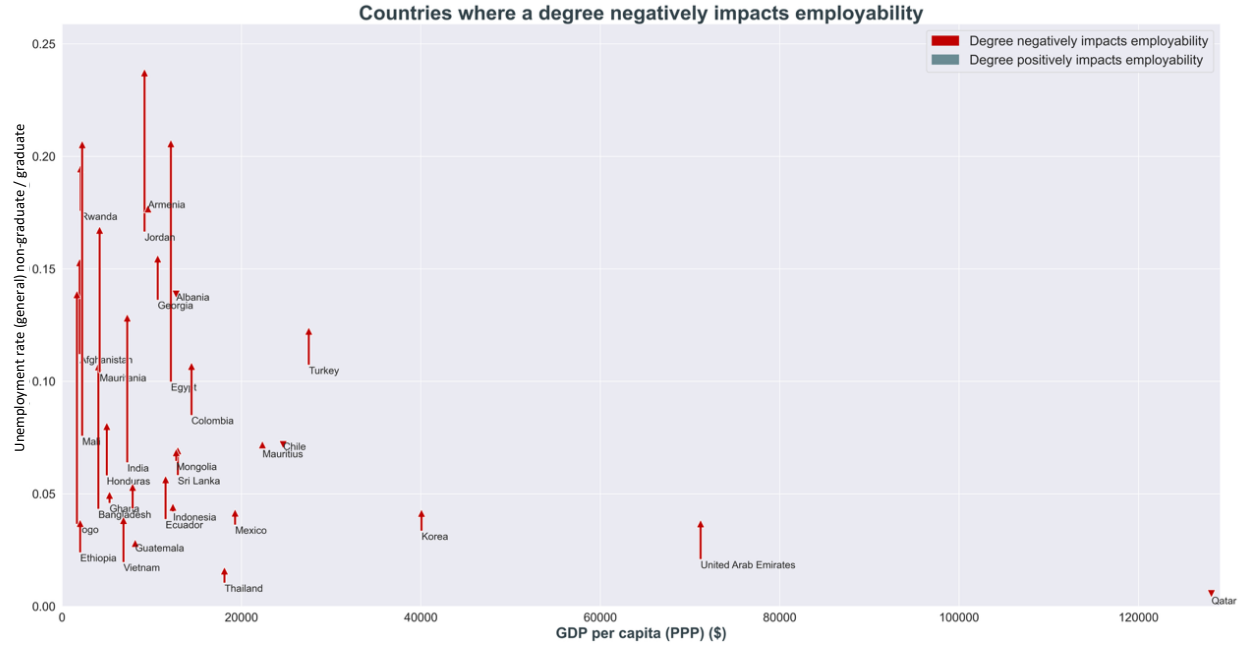


Figure 4-20 Countries where a degree negatively impacts employability (population G)

- Only eight of these countries have a GDP per capita (PPP) exceeding $15,000.

- 50% of the 33 countries studied have a graduate unemployment rate above 7.3%. For comparison, the median unemployment rate of graduates across the 88 countries in our sample is 5%.

Regarding the difference between the non-graduate unemployment rate and the graduate unemployment rate in the 33 countries presented in the figure above, the median of these differences is 2.15%, the third quartile is 3.9%, and the mean is 3.3%. The eight countries with the largest differences between the non-graduate unemployment rate and the graduate unemployment rate are Mali (13%), Egypt (11.1%), Togo (10%), India (8.5%), Jordan (7.6%), Mauritania (7%), Bangladesh (6%), and Afghanistan (4.6%). These are also countries with low GDP per capita (PPP). For these countries with very high differences, we sought to explore whether part of these disparities could be explained by gender differences and, therefore, examined employability based on gender.

Figure 4-21 shows the unemployment rates of graduates in countries where gender segmentation plays a significant role. It is worth noting that the seven countries in which gender segmentation reveals a very large disparity in graduate unemployment rates are also those with the highest general graduate unemployment rates, regardless of gender.

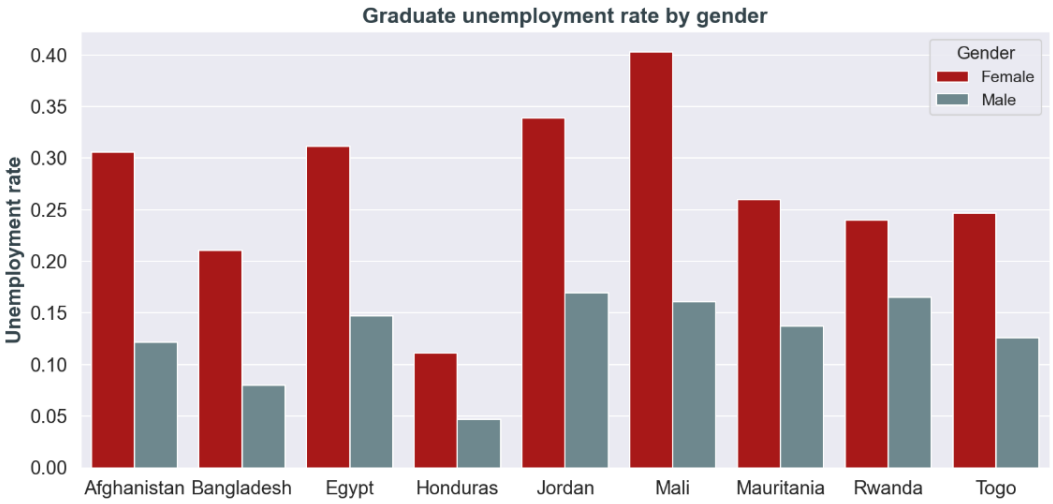


Figure 4-21 Graduate unemployment rate by gender in Afghanistan, Bangladesh, Egypt, Jordan, Mali, Mauritania, Togo, and India

In Mali, for instance, the female unemployment rate reaches 40.27%, compared to 16.10% for men. This trend is also observed in Egypt, where women face an unemployment rate of 31.14%, while the men's rate is 14.73%. In Jordan, the gap is equally striking, with 33.91% for women versus 16.96% for men. These figures highlight persistent structural inequalities in access to employment despite higher education levels.

However, some countries have lower general unemployment rates, although gender disparities persist. For example, in Honduras, women have an unemployment rate of 11.09%, while men have a rate of 4.73%.

This gender disparity underscores systemic challenges for female graduates. These obstacles may include labor market discrimination, social norms restricting access to employment, or mismatches between acquired qualifications and available job opportunities. Female graduates face significantly greater difficulties in accessing employment in countries with low GDP per capita (PPP), regardless of geographical region. Inclusive

and targeted policies are necessary to reduce these disparities and promote equal opportunities in the labor market.

So far, we have focused on population **(G),** which includes all age brackets and differentiates based on education level. We will now examine population **(J),** composed of young people aged 15 to 24, also using education level as a differentiation criterion.

	TCDF	TCDH	Delta abs(F-H)	Delta rel(F/H)
count	28.000000	28.000000	28.000000	28.000000
mean	0.138413	0.083704	0.055900	1.974613
std	0.109112	0.054927	0.068824	1.810557
min	0.011172	0.001100	0.000220	0.902042
25%	0.049711	0.042240	0.005782	1.078912
50%	0.097743	0.065134	0.021319	1.419119
75%	0.217896	0.134137	0.087011	2.028112
max	0.402761	0.177620	0.241721	10.156139

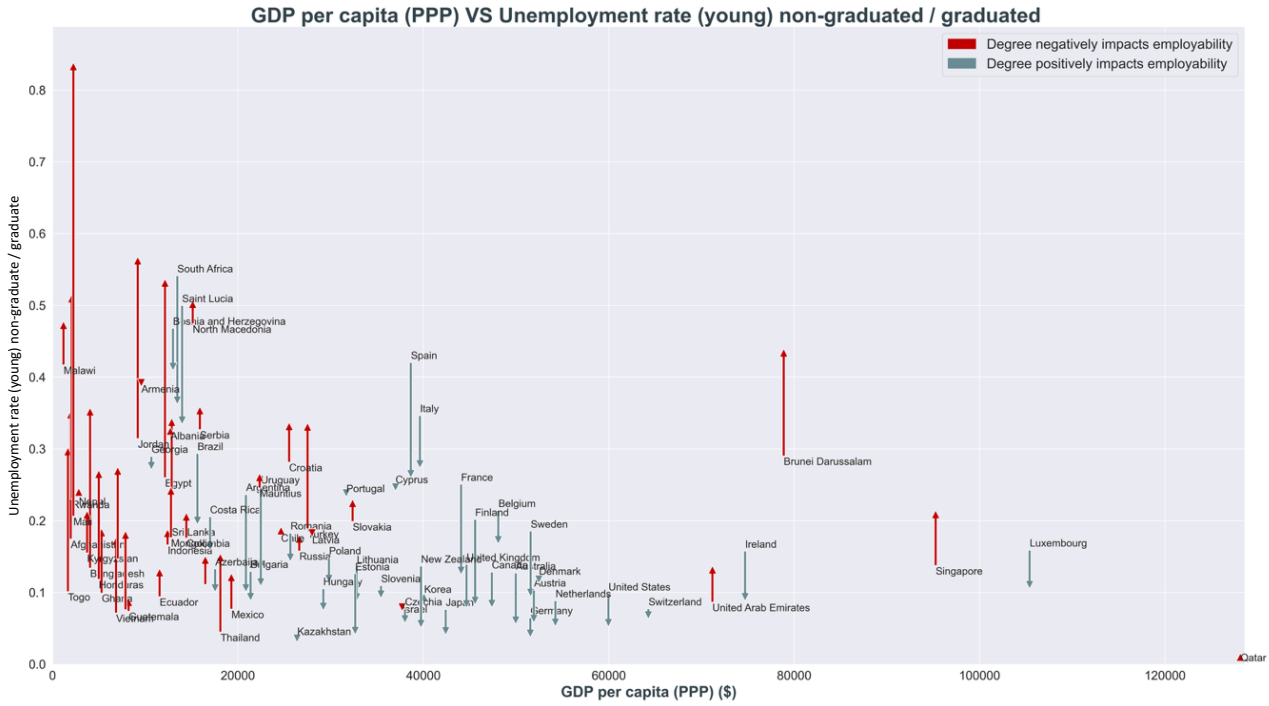


Figure 4-22 Impact of higher education degrees on unemployment as a function of GDP per capita (PPP) for the youth population (J) (2017)

Population **(J)** includes data from 80 countries, representing 149.5 million young non-graduates and 24.8 million young graduates. In half of these countries, obtaining a degree negatively impacts employability (figure 4-23). In other words, a degree negatively affects employability in seven more countries than in population **(G)**.

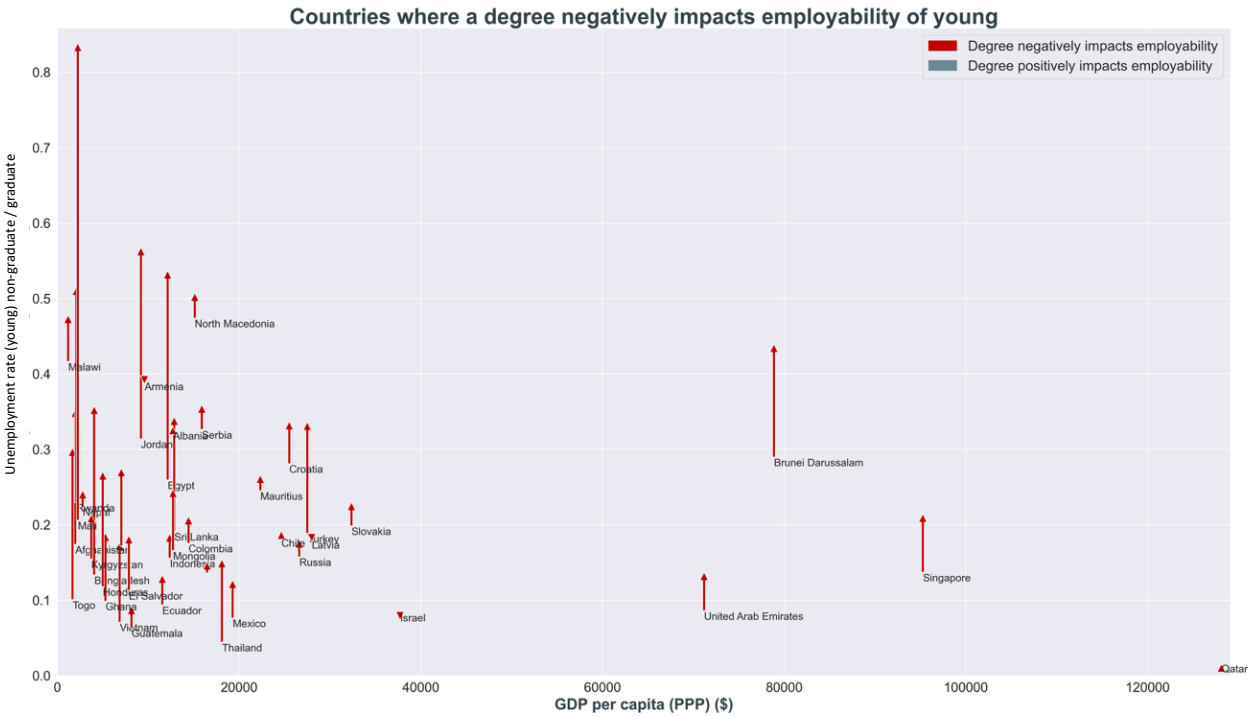


Figure 4-23 Countries where a degree does not provide additional protection against youth unemployment (population J)

Figure 4-24 shows the countries where the difference between the young graduate unemployment rate and that of young non-graduates (Unemployment Delta) is greatest.

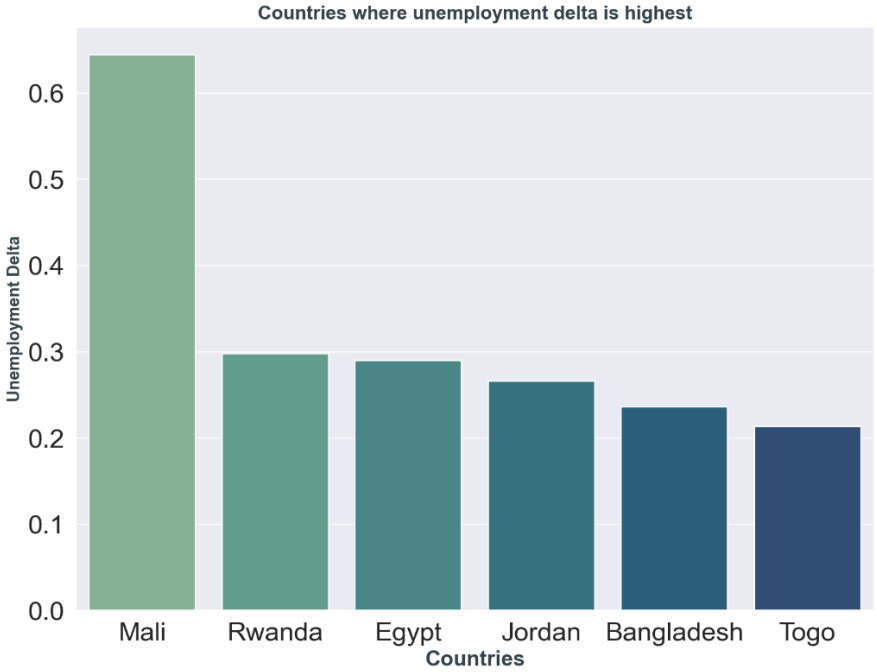


Figure 4-24 Countries where the difference between young graduate unemployment rate and young non-graduate unemployment rate (Unemployment Delta) is the largest

For the same reason as before, we will explore whether part of this observation can be explained by gender and thus compare the unemployment rate of young female graduates with that of young male graduates in the six countries that show the largest gap between the unemployment rates of young graduates (of all genders) and young non-graduates.

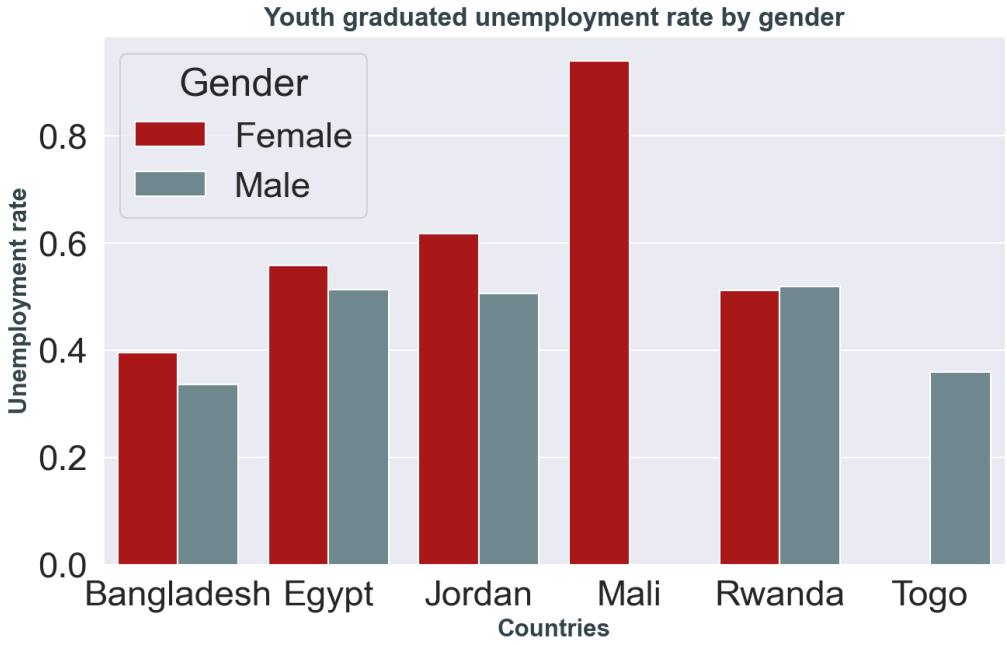


Figure 4-25 Young graduate unemployment rate by gender

- In the case of Mali and Togo, data are not available for both genders.

- With the exception of Mali, the differences between the youth unemployment rate for female graduates and that for male graduates are much smaller than the differences observed as an effect of the possession of a diploma for all ages combined (figure 4-25 versus figure 4-21) for unemployment rates in general according to age. In these countries, the Youth unemployment rate for female graduates is not the primary cause of the young graduate unemployment rate, which is observed to be much higher than the youth unemployment rate.

- The countries with the highest youth graduate unemployment rates are the same as those with the highest general graduate unemployment rates. In these countries, gender does not play an important role, while youth itself leads to an unemployment rate double that of the average. What's more, older female graduates are at a more lasting disadvantage on the job market. For these countries, gender is not a determining factor in young people's access to employment. On the other hand, youth is a variable that considerably reduces the employability of graduates. The young graduate unemployment rate is twice as high as the graduate unemployment rate (all ages combined). In addition, gender and age are interrelated. For example, older female graduates are penalized for longer periods on the job market.

We can therefore conclude that the countries where diplomas are the greatest hindrance to young people's employability are more or less the same as those where diplomas are detrimental to employability at all ages. On the other hand, the gender effect is much less marked among young graduates than among mature graduates.

4.4.2 Temporal stability of these national characteristics relative to one another : examples

All our results are based on 2017 data, the last pre-COVID figures available at the time of this study. This element may lead to a legitimate critique of our research. Do these variables depend on the period? Do these phenomena exhibit a certain consistency over the long term?

The following four figures illustrate the evolution over 10 to 30 years (depending on available data) of the six different types of unemployment rates in Canada, Colombia, Egypt, and France. These four countries were selected due to their differing levels of economic development.

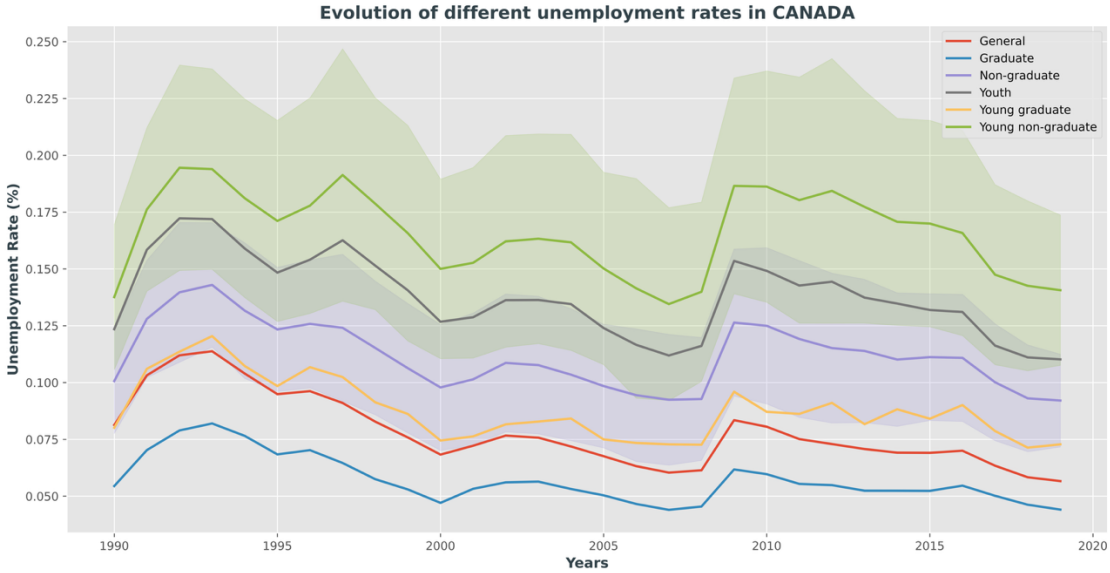


Figure 4-26 Evolution of different unemployment rates in Canada from 1990 to 2017

- In Canada, young non-graduates are the most affected by unemployment. In general, obtaining a degree provides effective protection against unemployment. The gaps between the different unemployment curves remain remarkably consistent over time. It can therefore be concluded that, in Canada, the effects of a degree on employability do not depend on the period; they are intrinsic.

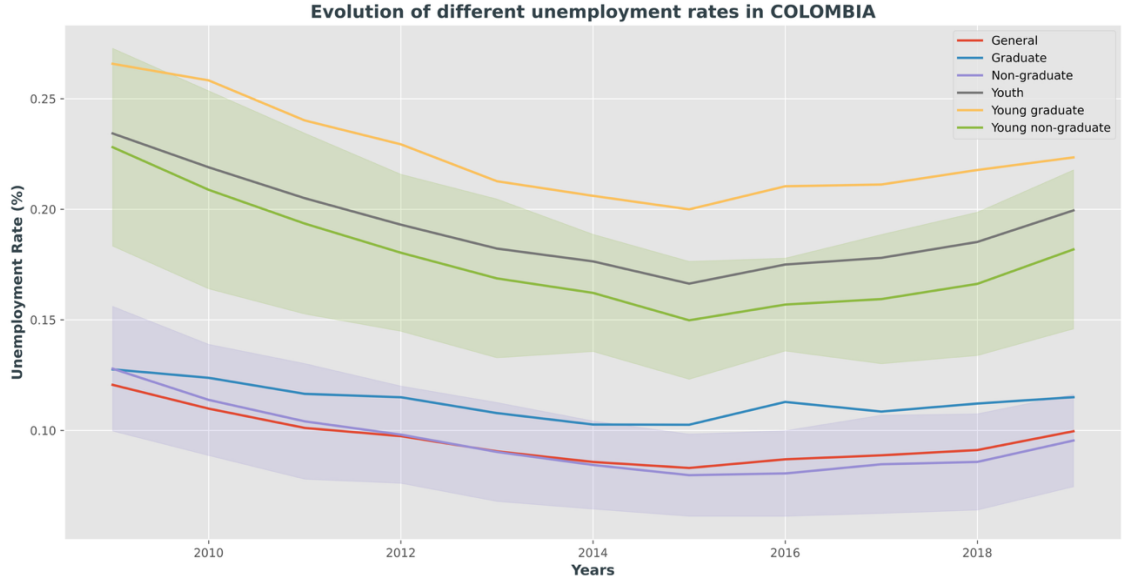


Figure 4-27 Evolution of unemployment rates in Colombia from 2009 to 2019

- The situation is quite different in Colombia, where young graduates are the most likely to be unemployed. In 2019, 22.5% of them were unemployed. By comparison, the young non-graduate unemployment rate was around 18%. It thus appears that holding a degree negatively affects employability in Colombia. Moreover, with a 20% unemployment rate, young people are more affected than older age brackets.
- However, apart from the general unemployment rate and the non-graduate unemployment rate, the relative positioning of the curves remains highly consistent over time.

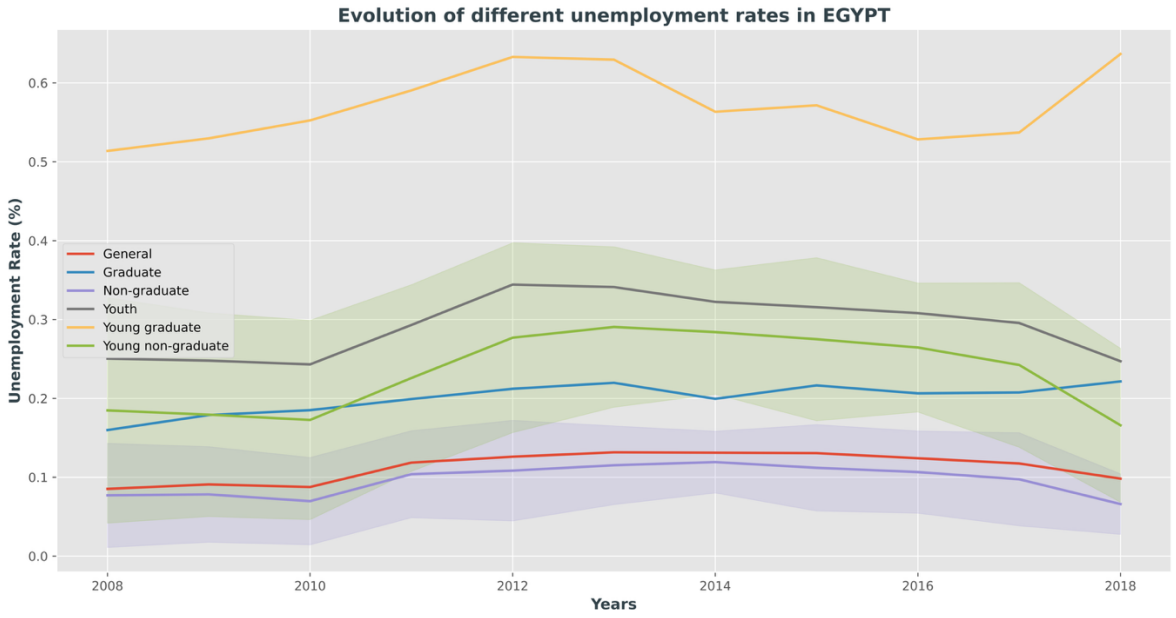


Figure 4-28 Evolution of unemployment rates in Egypt from 2008 to 2018

- Egypt presents an even more pronounced situation than Colombia. In 2019, 65% of young Egyptian graduates were unemployed, compared to 16.5% of young non-graduates.
- Nevertheless, the relative positioning of the curves remains highly stable over a long period of 20 years.

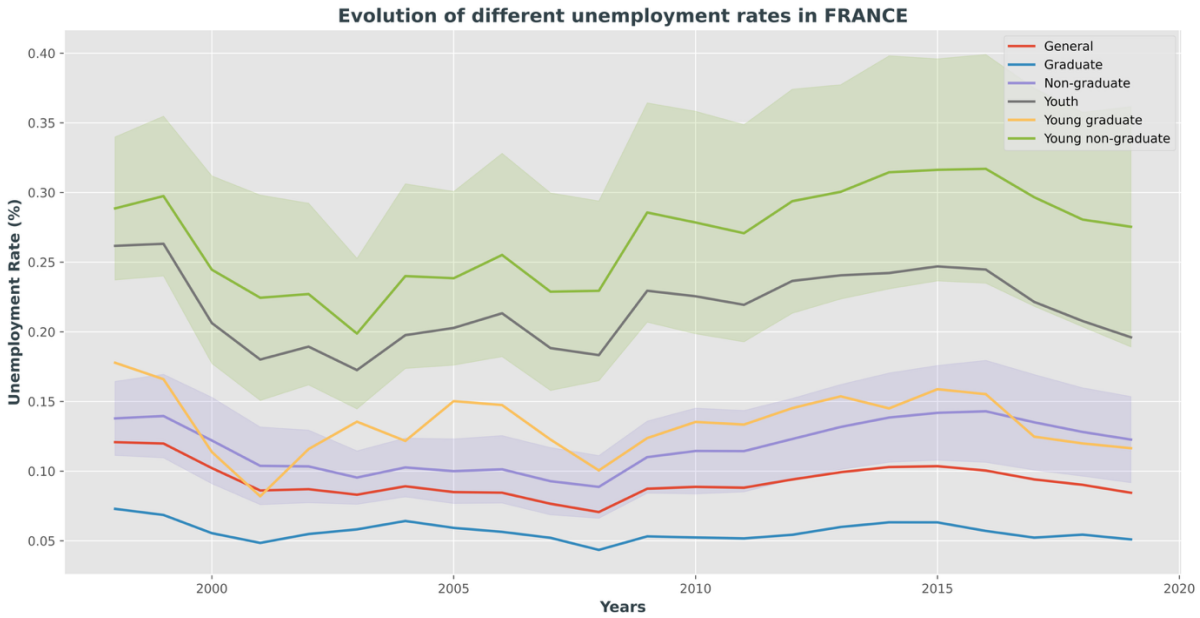


Figure 4-29 Evolution of unemployment rates in France from 1998 to 2019

- In France, in 2019, obtaining a degree significantly improved access to employment. Only 5% of graduates were unemployed, compared to 12.5% of non-graduates.
- Similarly, youth unemployment among non-graduates remained higher than that of young graduates over the same period. In 2017, the non-graduate unemployment rate in France was 12%, while the young non-graduate unemployment rate was 25.5%.
- However, the youth unemployment rate remained 2.5 times higher than the general unemployment rate. Given that the median non-graduate unemployment rate and graduate unemployment rate across all studied countries were 7% and 16%, respectively, France thus exhibits a relatively high unemployment rate among non-graduates and young non-graduates. In France, obtaining a degree provides strong protection against unemployment.
- Nevertheless, the relative positioning of these curves has remained remarkably stable over 20 years, with slightly greater variations observed in the unemployment rates of young graduates.

These four successive figures, from countries with vastly different histories, levels of wealth, and social structures, reveal a striking parallelism in the long-term evolution of these six different curves representing different types of unemployment. Increases and decreases in various unemployment rates have followed similar patterns over the past 10 to 30 years in each country, with each maintaining its own specific characteristics. This suggests that these phenomena are deeply rooted in the sociological structures of unemployment situations, shaped by the economic and social organizations unique to each country.

4.5 Alpha Rate and unemployment rate: aggregated data

In section 4.3, we examined the relationship between different types of unemployment rates and economic development. We also established in paragraph 4.2 that the higher education access rate (Alpha Rate) and economic development were correlated, due to the strong correlation observed among countries with low levels of economic development. We will now analyze, using the same data, the correlation between the different types of unemployment rates and the higher education access rate, measured by the Alpha Rate, in various countries.

The available data for the Alpha Rate corresponds to the year 2017. These data will be aggregated with those related to unemployment.

Aggregated data sets	General unemployment rate	Graduate unemployment rate	Non-graduate unemployment rate
Alpha Rate 2017	Population observed: **91 countries, i.e. 3.57×10^9** Global active population: **196 countries, i.e., 5.5×10^9**	Observed active population: **88 countries, i.e., 0.558×10^9**	Observed active population: **87 countries, i.e., 0.96×10^9**

Tableau 4-11 Sizes of aggregated data sets, all age brackets

Aggregated data sets	General unemployment rate	Young graduate unemployment rate	Young non-graduate unemployment rate
Alpha Rate 2017	Observed active population: **92 countries, i.e., 0.662×10^9** Global active population: **196 countries, i.e., 1.2×10^9**	Observed active population: **81 countries, i.e., 0.025×10^9 + China[35]**	Observed active population: **88 countries, i.e., 0.154×10^9**

Table 4-12 Aggregate data set sizes, young people (15 - 24 years)

For each data set, the sample used is statistically very significant.

[35] Though we do not know the number of young graduates in China, we were able to find the unemployment rate for young graduates in China

4.5.1 Alpha Rate and general unemployment rate

Figure 4-30 represents the relationship between the general unemployment rate and the Alpha Rate. In previous sections, we established that the distributions of these two variables were not Gaussian.

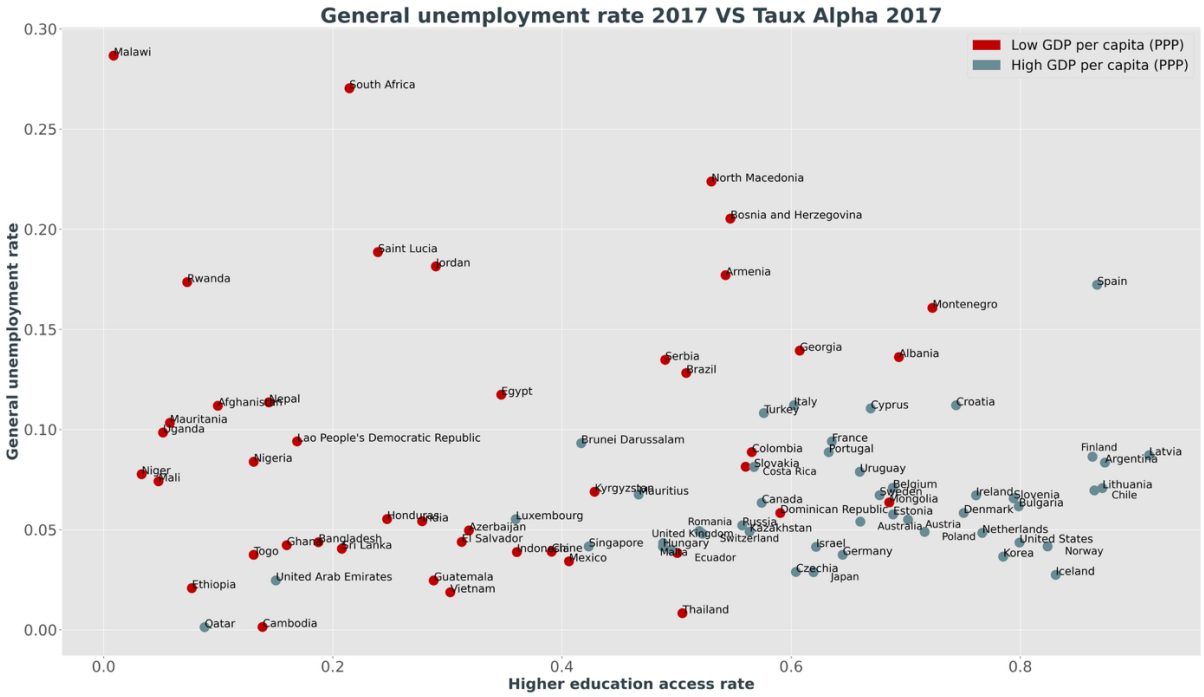


Figure 4-30 Higher education access rate as a function of the general unemployment rate (2017)

The analysis includes 91 countries, representing 3.57 billion of the 5.5 billion active individuals worldwide. The table below presents the intensity of the Spearman correlation between the Alpha Rate and the General unemployment rate.

Variables: Alpha Rate and general unemployment rate

Correlation r	Statistical significance test result (95%)	Sample size
0.046	p = 0.663, H_0 not rejected	91

Table 4-13 Correlation between the general unemployment rate and the higher education access rate (2017)

The points in Figure 4-30 are highly dispersed, suggesting a weak correlation. Similarly, Table 4-9 illustrates that no monotonic relationship exists between the general unemployment rate and the Alpha Rate.

The access rate to higher education has no impact on the general unemployment rate, and vice versa.

4.5.2 Alpha Rate and youth unemployment rate

We now consider the relationship between the Alpha Rate and the youth unemployment rate. Our sample covers 92 countries, representing 0.66×10^9 young individuals. The figure below represents the relationship between these two variables.

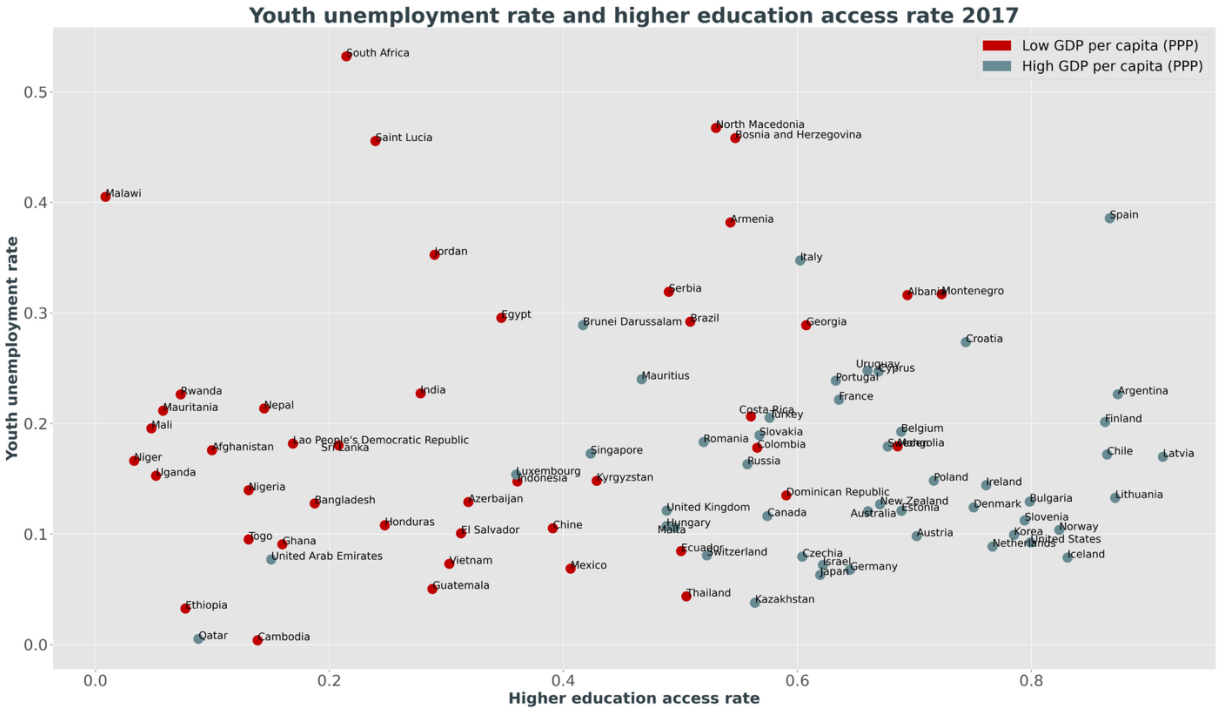


Figure 4-31 Higher education access as a function of the youth unemployment rate (2017)

The points, whether they represent rich or poor countries, are not concentrated around a straight line. This suggests that a relationship, even a nonlinear one, probably does not exist between these two variables.

Since neither variable follows a Gaussian distribution, we prefer to measure a Spearman correlation:

Variables: Alpha Rate and youth unemployment rate

Correlation r	Statistical significance test result (95%)	Sample size
0,015	p = 0.888, H_0 not rejected	92

Table 4-14 Correlation between the higher education access rate and the youth unemployment rate

The correlation, with a coefficient of 0.015, is not statistically significant.

The higher education access rate (Alpha Rate) has no impact on the youth unemployment rate and vice versa.

4.5.3 Alpha Rate and graduate unemployment rate

So far, we have established a strong correlation (-0.612) between the youth unemployment rate and GDP per capita (PPP). Furthermore, we have identified a strong link between the level of economic development and the Alpha Rate for countries with a GDP per capita (PPP) below $15,000.

We will now examine the relationship between the Alpha Rate and the graduate unemployment rate, as represented in figure 4-32 below.

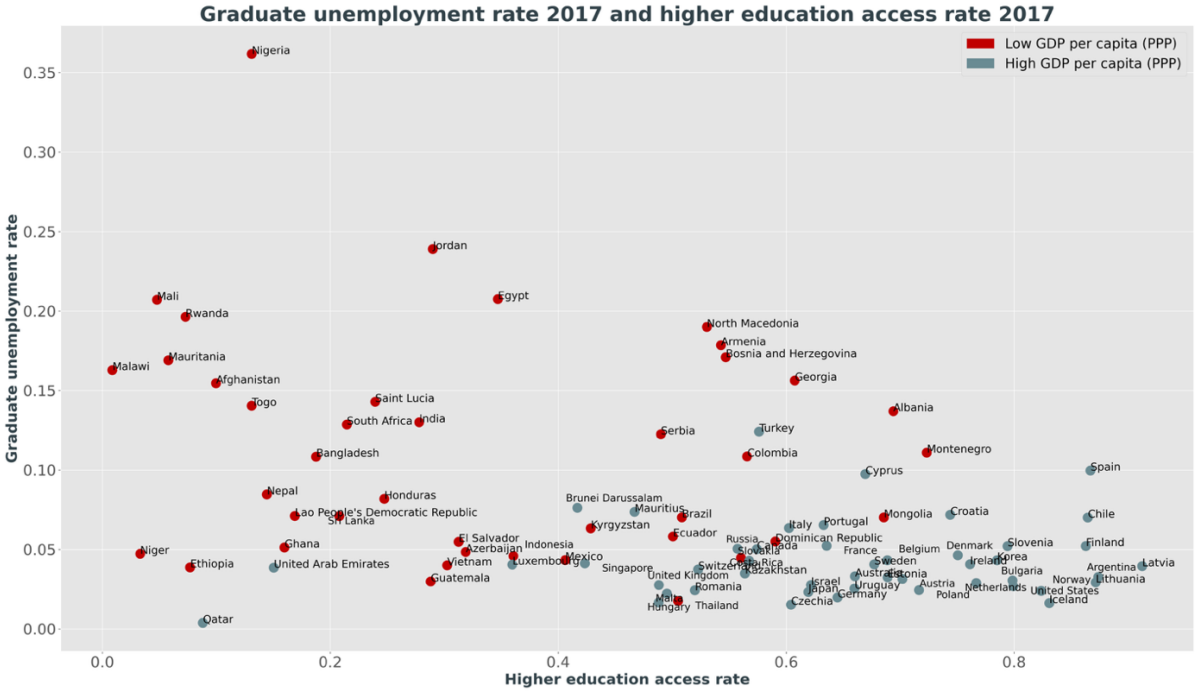


Figure 4-32 Graduate unemployment rate and higher education access rate (2017)

The sample used for our observations includes 88 countries, representing 0.558×10^9 graduates. The figure establishes a negative linear relationship between the two variables.

Variables: Alpha Rate and graduate unemployment rate

Correlation r	Statistical significance test result (95%)	Sample size
-0.38	p = 0.0, H_0 rejected	88

Tableau 4-15 Correlation between the higher education access rate and graduate unemployment rate

The correlation coefficient between the variables is -0.38, indicating a moderate degree of correlation. Thus, 14% of the variations in the graduate unemployment rate are caused by fluctuations in the Alpha Rate.

The lower the graduate unemployment rate, the higher the access to higher education.

This observation is consistent with our first intuition, which suggested that a low graduate unemployment rate encourages (or at least does not discourage) young people from pursuing higher education.

4.5.4 Alpha Rate and unemployment rate among young graduates

As with the graduate unemployment rate and GDP per capita (PPP), a negative correlation exists between economic development and the unemployment rate of young graduates (-0.641).
In this section, we will examine the relationship between the Alpha Rate and the young graduate unemployment rate, as described in figure 4-33.

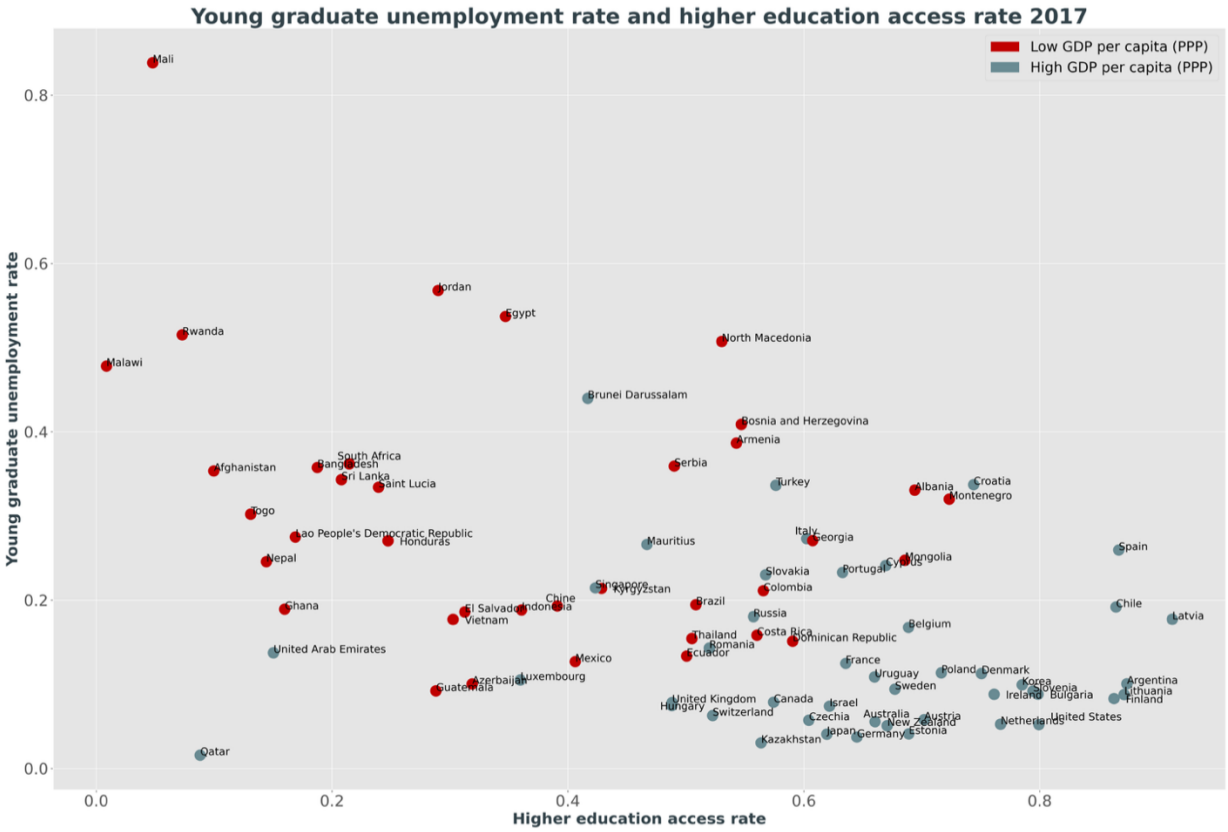


Figure 4-33 Higher education access rate and young graduate unemployment rate (2017)

Our sample includes 82 countries, representing 0.025×10^9 young graduates[36]. Some highly populated countries such as India, Ethiopia, and Nigeria are not included in the analysis due to a lack of available data.

Variables: Alpha Rate et Young graduate unemployment rate

Correlation r	Statistical significance test result (95%)	Sample size
-0.46	p =0.0, H_0 rejected	82

Tableau 4-16 Correlation between the higher education access rate and young graduate unemployment rate

There is, therefore, a moderate correlation between these two variables (-0.46). Consequently, 21% of the variations in the young graduate unemployment rate can be explained by fluctuations in the Alpha Rate.

Thus, the graduate unemployment rate and the young graduate unemployment rate exhibit a monotonic relationship with the Alpha Rate, unlike the general unemployment rate and youth unemployment rate. However, while real, the correlations between access to higher education and the graduate and young graduate unemployment rates remain weak.
The lower the young graduate unemployment rate, the higher the access to higher education.

[36] China is included though the number of young Chinese graduates remains unknown

4.5.5 Alpha Rate and unemployment rate among non-graduates

Our analysis will focus on the relationship between the Alpha Rate and the non-graduate unemployment rate. This study is based on a sample of 90 countries, representing 1.89×10^9 non-graduates.

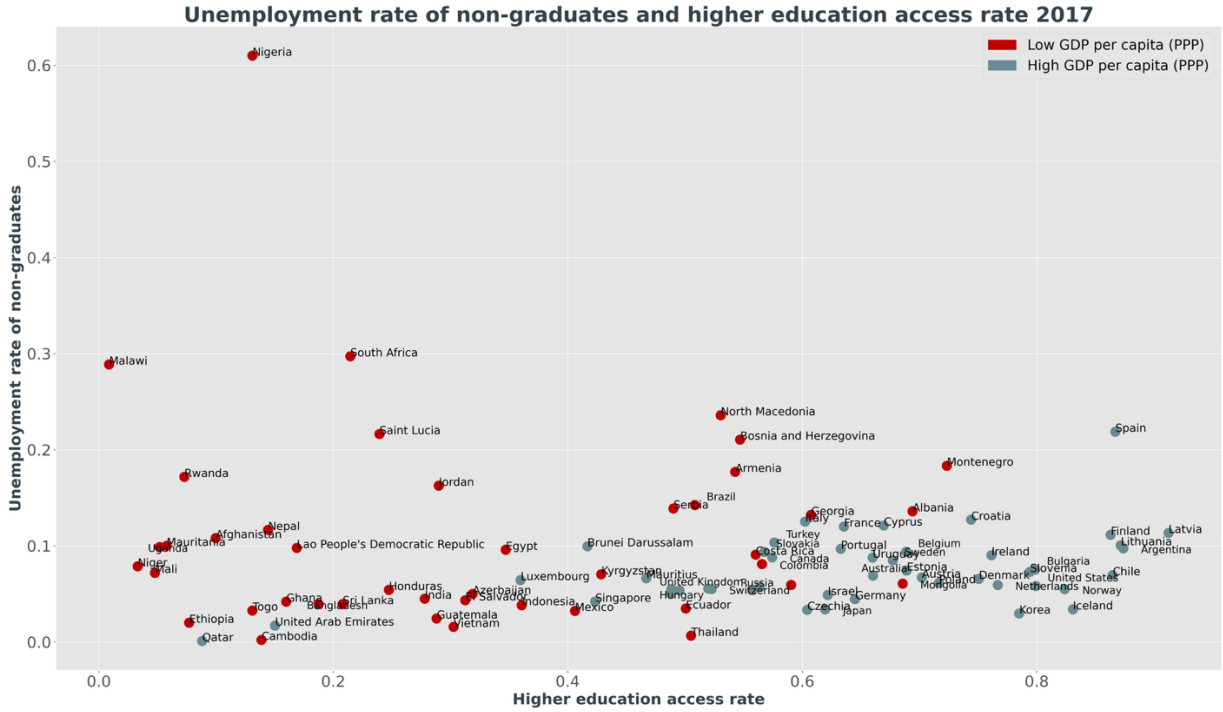


Figure 4-34 Higher education access rate and non-graduate unemployment rate (2017)

The figure highlights a weak or nonexistent linear relationship between the variables. Table 4-14 confirms this observation:

Variables: Alpha Rate and non-graduate unemployment rate

Correlation r	Statistical significance test result (95%)	Sample size
0.135	p=0.205, H_0 not rejected	90

Table 4-17 Correlation between the higher education access rate and non-graduate unemployment rate

The relationship between the Alpha Rate and the non-graduate unemployment rate is not statistically significant.
Access to higher education, therefore, has no impact on the non-graduate unemployment rate.

4.5.6 Alpha Rate and unemployment rate among young non-graduates

Finally, in this section, we will examine the relationship between the Alpha Rate and the young non-graduate unemployment rate. Our analysis covers 88 countries, representing 0.154×10^9 young non-graduates.

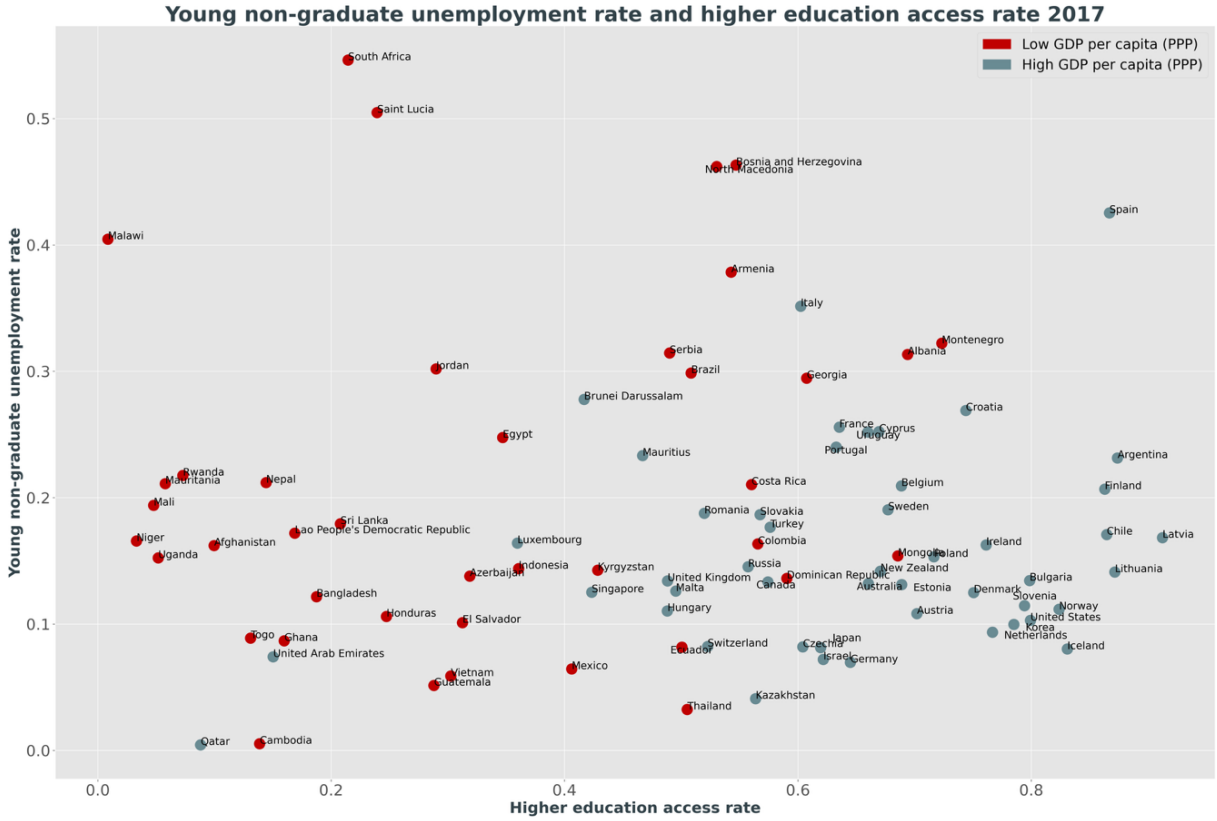


Figure 4-35 Higher education access rate and young non-graduate unemployment rate (2017)

The data points in the figure are scattered, indicating a weak or nonexistent linear dependence between the variables.

Variables : Alpha Rate and young non-graduate unemployment rate

Correlation r	Statistical significance test result (95%)	Sample size
0.027	p=0.801, H_0 not rejected	88

Table 4-18 Correlation between the higher education access rate and young non-graduate unemployment rate

The correlation is not statistically significant. The hypothesis that no correlation exists between these two variables is retained.

Access to higher education has no impact on the young non-graduate unemployment rate.

4.5.7 Closing remarks on this chapter

	General unemployment rate	Graduate unemployment rate	Non-graduate unemployment rate	Youth unemployment rate	Young graduate unemployment rate	Young non-graduate unemployment rate
Alpha Rate	**r_s=0.046** H_0 not rejected Sample: 91	**r_s=-0.38** r_s^2=0.144 H_0 rejected Sample: 88	**r_s=0.135** H_0 not rejected Sample: 90	**r_s=0.015** r_s^2=0 H_0 not rejected Sample: 92	**r_s=-0.46** r_s^2=0.212 H_0 rejected Sample: 82	**r_s=0.027** H_0 not rejected Sample: 88

Table 4-19 Summary of correlations between the higher education access rate and different unemployment rates

The objective of this chapter was to answer our second research question Q2: "Can correlations be established between access to higher education and different unemployment rates (general, graduates, youth, and young graduates) in countries at very different stages of economic development?".

Our results indicate that variations in the rate of access to higher education have no impact on general, youth, non-graduate, and young non-graduate unemployment rates.

Only moderate correlations are observed between the Alpha Rate and the graduate and young graduate unemployment rates. Specifically, variations in the Alpha Rate explain only 14% and 21% of the fluctuations in the graduate and young graduate unemployment rates, respectively. Consequently, 80% of the variations in these unemployment rates are explained by other factors.

Thus, we have established weak or nonexistent correlations between access to higher education and the different unemployment rates observed worldwide. At this stage, it appears relevant to investigate whether the observed dispersion of data points, leading to weak or null correlations, could result from antagonistic effects among different types of countries. It is possible that certain clusters of countries share similar socioeconomic behaviors, but these behaviors might offset each other when considered collectively in the data set.

Chapter 5 will explore the potential existence of such clusters to determine whether countries can be grouped based on persistent socioeconomic characteristics that lead to homogeneous behaviors.

5 Cluster analysis

We will now consider different country groupings to examine the existence of potential monotonic relationships between the Alpha Rate and various types of unemployment, based on 2017 data, within possible clusters, for the reasons mentioned at the end of the previous page. The list of countries is reordered alphabetically.

These groups were formed using the following methods:

- The first method considers economic development. Specifically, we analyzed correlations between the variables within the cluster of countries with low GDP per capita (PPP) and those with high GDP per capita (PPP), with a threshold of approximately $15,000 separating the two groups. For both clusters, no additional relationships were identified between access to higher education and the different unemployment rates beyond those observed in Chapter 4.

- The second method compares the various types of unemployment rates with their respective medians. The only significant relationships (around -0.60) are those linking higher education access rates with the graduate unemployment rate and the young graduate unemployment rate. For the remaining cases, correlations are not significant, regardless of whether the cluster consists of countries with unemployment rates above or below the medians.

- The third partitioning method used an unsupervised machine learning algorithm. This algorithm accounts for all variables (higher education access rate, GDP per capita (PPP), different types of unemployment) to group countries into homogeneous classes, without prior knowledge of their composition.

5.1 Estimation of missing data using XGBoost methods

The three partitioning methods described above will serve as the new basis for analyzing correlations between variables within the resulting subgroups (or clusters). Our data set includes 80 countries for which the following data points are fully available: GDP per capita (PPP), Alpha Rate, general unemployment rate, youth unemployment rate, graduate unemployment rate, young graduate unemployment rate, non-graduate unemployment rate, and young non-graduate unemployment rate. The countries for which these data are fully available (complete data set for 2017) are as follows:

Afghanistan, Albania, Argentina, Armenia, Australia, Austria, Azerbaijan, Bangladesh, Belgium, Bosnia and Herzegovina, Brazil, Brunei Darussalam, Bulgaria, Canada, Chile, Colombia, Costa Rica, Croatia, Cyprus, Czechia, Denmark, Dominican Republic, Ecuador, Egypt, El Salvador, Estonia, Finland, France, Georgia, Germany, Ghana, Guatemala, Honduras, Hungary, Indonesia, Ireland, Israel, Italy, Japan, Jordan, Kazakhstan, Korea, Kyrgyzstan, Laos, Latvia, Lithuania, Luxembourg, Malawi, Mali, Mauritius, Mexico, Mongolia, Montenegro, Nepal, Netherlands, North Macedonia, Poland, Portugal, Qatar, Romania, Russia, Rwanda, Saint Lucia, Serbia, Singapore, Slovakia, Slovenia, South Africa, Spain, Sri Lanka, Sweden, Switzerland, Thailand, Togo, Turkey, United Arab Emirates, United Kingdom, United States, Uruguay, Vietnam.
India, China, Ethiopia, Nigeria, Malta, Iceland, Niger, Norway, and Mauritania were not included in some parts of the previous analysis in Chapter 4 due to missing data. Specifically:

- No data exists for the young graduate unemployment rate in Norway, Niger, Mauritania, Iceland, and Malta (Group A).
- No data exists for the graduate and non-graduate unemployment rates in India, Nigeria, and Ethiopia (Group B).
- The non-graduate unemployment rate and young non-graduate unemployment rate for China are not available (Group C).

Given the demographic significance of some of these countries and their small number, it is essential to estimate the missing data values for each of them. To estimate these values, we employed Machine Learning regression models based on the data from the 80 countries for which all variables are available. The prediction technique is based on the use of the XGBoost method, integrated into one of Python's libraries. This technique is further detailed in "A Scalable Tree Boosting System"[37] by Tianqi Chen and Carlos Guestrin, published in 2016. The XGBoost method is one of the most powerful machine learning techniques in terms of both result quality and computation speed. It specifically combines Decision Trees and the Gradient Boost Algorithm, both recognized as highly efficient algorithms.

The logical structures, model implementations, and computations are explained in greater detail in the Appendices, where residual errors are also documented.
Table 5-1 presents our results:

Country	Young graduate unemployment rate
Iceland	0.0583
Malta	0.0777
Mauritania	0.5683
Niger	0.2404
Norway	0.0588

Table 5-1 Predictions by model M1 for group A (see Annex)

Country	Young graduate unemployment rate	Young non-graduate unemployment rate
Ethiopia	0.137	0.034
Inde	0.358	0.203
Nigeria	0.492	0.166

Table 5-2 Predictions by model M2 and M3 for group B (see Annex)

Country	Graduate unemployment rate	Non-graduate unemployment rate	Young non-graduate unemployment rate
China	0.057	0.034	0.096

Table 5-3 Predictions by model M3, M4 and M5 for group C (see Annex)

[37] Tianqi Chen & Carlos Guestrin, "A Scalable Tree Boosting System", 2016, In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785‑794). New York, NY, USA: ACM, https://dl.acm.org/doi/pdf/10.1145/2939672.2939785

5.2 Data partition analysis based on GDP per capita (PPP)

The first clustering method considers economic development. Countries with a GDP per capita (PPP) below $20,000 (the median GDP per capita (PPP)) are classified as low-development countries, while those with a GDP per capita (PPP) above this threshold are classified as high-development countries.
It is important to note that the sample size of countries varies moderately depending on the type of unemployment considered (detailed lists are provided in the Annex).

In the previous chapter, we measured Spearman correlations between the Alpha Rate and different types of unemployment rates. Our results are summarized in the tables below.

	General unemployment rate	Graduate unemployment rate	Non-graduate unemployment rate	Youth unemployment rate	Young graduate unemployment rate	Young non-graduate unemployment rate
Alpha Rate	**$r_s = 0.294$** **$r^2 = 0.086$** H_0 rejected Sample: 47	**$r_s = 0.06$** H_0 not rejected Sample: 47	**$r_s = 0.39$** **$r^2 = 0.152$** H_0 rejected Sample: 47	**$r_s = 0.1$** H_0 not rejected Sample: 48	**$r_s = -0.04$** H_0 not rejected Sample: 45	**$r_s = 0.127$** H_0 not rejected Sample: 48

Table 5-4 Correlation table for high-GDP per capita (PPP) countries

The only statistically significant correlations are observed between the Alpha Rate and the general unemployment rate, as well as between the Alpha Rate and the non-graduate unemployment rate. In the first case, the correlation coefficient is 0.294, and in the second case, it is 0.39. The first correlation is weak, whereas the second is of moderate strength. The latter implies that only 15% of the variations in the non-graduate unemployment rate in high-GDP countries can be explained by variations in the Alpha Rate, indicating a relatively weak factor.

When comparing the results for high-GDP-per-capita (PPP) countries with those obtained in the previous chapter (Chapter 4), we observe that the monotonic relationships between the Alpha Rate and the general unemployment rate, as well as between the Alpha Rate and the non-graduate unemployment rate, are much stronger in high-development countries. This suggests that an increase in the number of students in these high-GDP-per-capita (PPP) countries appears to be a (moderate) factor contributing to the rise in non-graduate unemployment. Thus, in wealthier countries, a higher rate of access to higher education does not increase the employability of graduates but rather reduces that of non-graduates. Furthermore, the absence of correlation between the Alpha Rate and the graduate and young graduate unemployment rates (while these parameters were correlated with the Alpha Rate across all countries, as seen in the previous chapter) could indicate differences in how high-GDP-per-capita (PPP) countries develop policies around higher education systems (e.g., more or less vocational training). This suggests that an alternative clustering method is needed to further investigate the relationship between the Alpha Rate and graduate and young graduate unemployment rates, which will be explored in Section 5.4.

Within the comparison among wealthier countries, an increase in the higher education access rate among these nations tends to slightly raise the general unemployment rate, has no effect on graduate unemployment, and increases non-graduate unemployment due to their relative downgrading.

Now, let us consider the set of low-development countries.

	General unemployment rate	Graduate unemployment rate	Non-graduate unemployment rate	Youth unemployment rate	Young graduate unemployment rate	Young non-graduate unemployment rate
Alpha Rate	**r_s=0.105** H_0 not rejected Sample: 44	**r_s=-0.149** H_0 not rejected Sample: 41	**r_s=0.067** H_0 not rejected Sample: 43	**r_s=0.186** H_0 not rejected Sample: 44	**r_s=-0.275** H_0 not rejected Sample: 37	**r_s=0.148** H_0 not rejected Sample: 40

Table 5-5 Correlation table for low-GDP-per-capita (PPP) countries

For low-GDP-per-capita (PPP) countries, no significant correlation exists between the Alpha Rate and any type of unemployment rate. The null hypothesis is not rejected for any of the six correlations.

Thus, in low-income countries, variations in the higher education access rate have no impact on the levels of different unemployment rates.

5.3 Analysis of countries with high general, youth, non-graduate and young non-graduate unemployment rates

Given the political significance of youth unemployment and the divergent discourses on this issue, it seemed relevant to further investigate countries with high levels of general unemployment, youth unemployment, unemployment among non-graduates, and unemployment among young non-graduates.

By examining the values of these four types of unemployment rates relative to their median values, we established four subgroups of countries:

- Group X includes countries where the general unemployment rate is high, meaning it exceeds the median general unemployment rate across all countries. This group consists of 44 out of the 89 countries studied.
- Group Y includes countries where the unemployment rate for non-graduates exceeds the median unemployment rate for non-graduates across all countries. Group Y comprises 44 of the 89 countries studied.
- Group Z includes countries where the youth unemployment rate is higher than the median youth unemployment rate across all countries. This set includes 45 of the 89 initial countries.
- Group W includes countries where the unemployment rate for young non-graduates is higher than the median unemployment rate for young non-graduates across all countries. This group consists of 45 countries.

5.3.1 Summary of cluster results

We examined the correlations between Alpha Rate and the unemployment rates for the general population, non-graduates, youth, and young non-graduates[38]. Our results are presented in Table 5-6 below:

	X (High general unemployment)		Y (High non-graduate unemployment)		Z (High youth unemployment)		W (High young non-graduate unemployment)	
	Graduates	**Young graduates**	**Graduates**	**Young graduates**	**Graduates**	**Young graduates**	**Graduates**	**Young graduates**
Alpha Rate	-0.607, p=0.0, H_0 rejected	-0.694, p=0.0, H_0 rejected	-0.634, p=0.0, H_0 rejected	-0.689, p=0.0, H_0 rejected	-0.476, p=0.0, H_0 rejected	-0.604, p=0.0, H_0 rejected	-0.498, p=0.0, H_0 rejected	-0.62, p=0.0, H_0 rejected

Table 5-6 Correlation table for clusters X, Y, Z, W

For each cluster of countries where at least one of the unemployment rates is high, we observe significant negative correlations between Alpha Rate and both graduate and young graduate unemployment rates. Several interpretations can be proposed for these variations:

- An increase in the graduate unemployment rate leads to a decline in student enrolment in a given country (discouragement factor).
- A decrease in the graduate unemployment rate encourages individuals to pursue higher education.
- In the same country, variations in both unemployment rates and Alpha Rate can be observed over relatively short periods (see figure 4-29). During economic crises, young graduates re-enrol in university, thereby being excluded from unemployment statistics, suggesting a countercyclical effect of university enrolment.

These are generally countries where higher education is free or nearly free.

5.3.2 Intersection of clusters X, Y, Z and W

An interesting analytical approach would be to examine the strength of the correlations between the variables across the set of countries that represent the intersection of groups X, Y, Z, and W. This intersection represents the set of countries where the general unemployment rate, the unemployment rate of non-graduates, the youth unemployment rate, and the unemployment rate of young non-graduates are all high (i.e., above their respective medians).

This set includes 32 countries: Albania, Argentina, Armenia, Belgium, Bosnia and Herzegovina, Brazil, Brunei Darussalam, Colombia, Costa Rica, Croatia, Cyprus, Egypt, Spain, Finland, France, Georgia, Italy, Jordan, Laos, Malawi, Mauritania, Montenegro, Nepal, North Macedonia, Portugal, Rwanda, Saint Lucia, Serbia, Slovakia, South Africa, Sweden, Turkey, and Uruguay.

These countries exhibit significant diversity in terms of economic development and Alpha Rate. This is illustrated in the following figures:

[38] See Annex for details

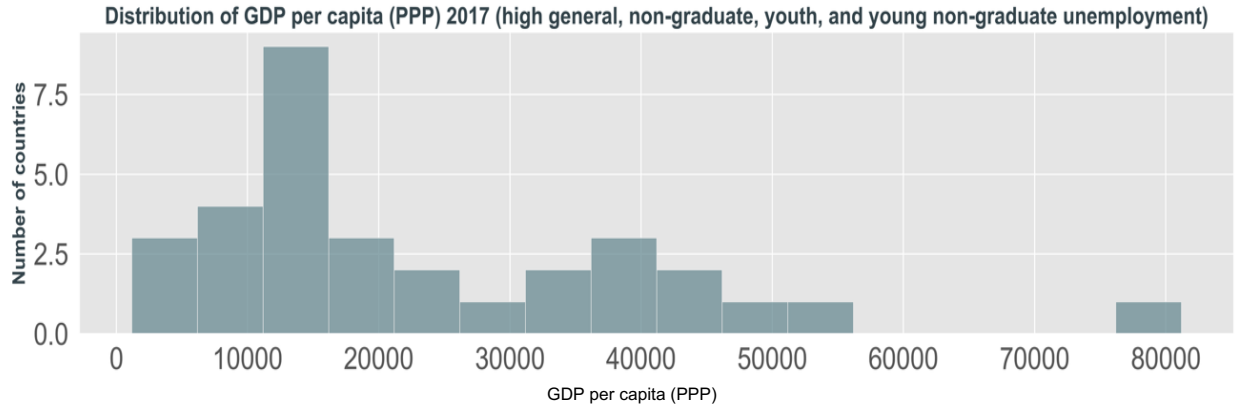


Figure 5-1 Distribution of GDP per capita (PPP) for countries with high general unemployment, high non-graduate unemployment, high youth unemployment, and high young non-graduate unemployment, 2017

We now examine the correlations between the Alpha Rate and the unemployment rates of graduates and young graduates across these 32 intersecting countries (including France). The correlation values are presented in Table 5-7 below.

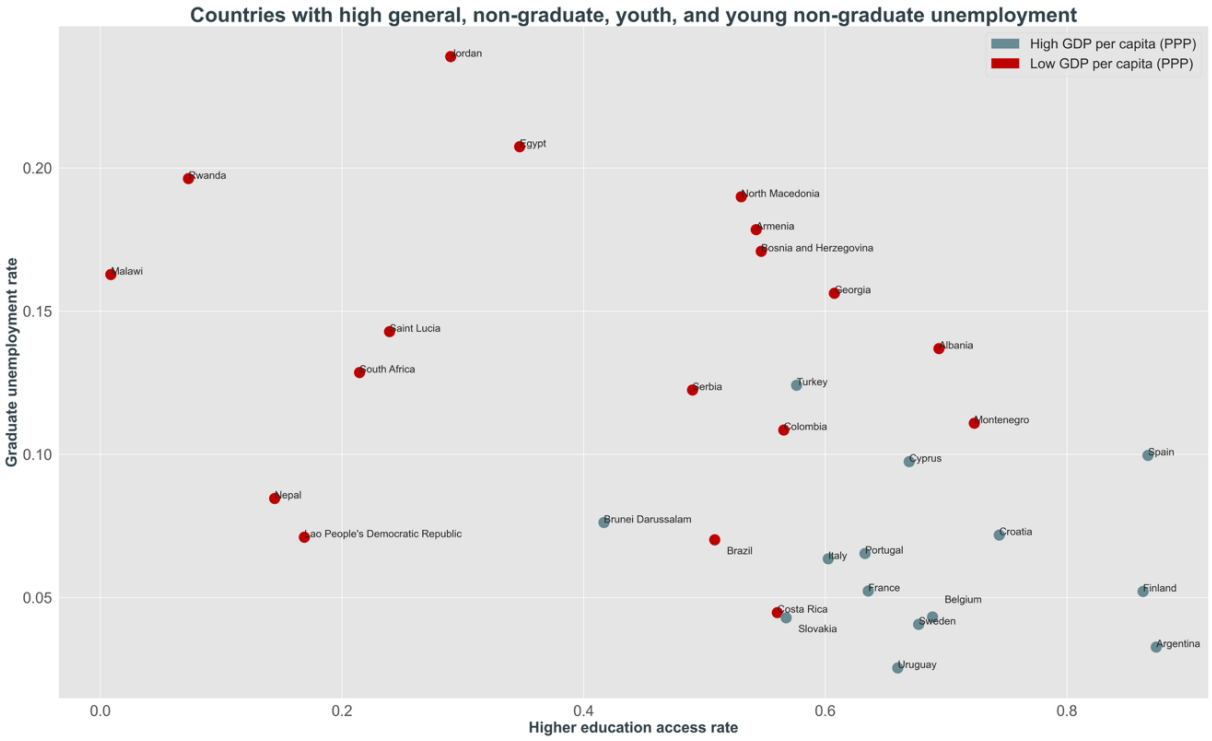


Figure 5-2 Higher education access rate and graduate unemployment rates in countries with high general, non-graduate, youth, and young non-graduate unemployment

Figure 5-3 Higher education access rate and young graduate unemployment rates in countries with high general, non-graduate, youth, and young non-graduate unemployment

	GDP per capita (PPP)	General unemployment rate	Graduate unemployment rate	Non-graduate unemployment rate	Youth unemployment rate	Young graduate unemployment rate	Young non-graduate unemployment rate
Alpha Rate	**r_s= 0.737** H_0 rejected Sample: 32	**r_s= - 0.283** H_0 not rejected Sample: 32	**r_s=-0.537** H_0 rejected Sample: 32	**r_s= - 0.067** H_0 not rejected Sample: 32	**r_s=-0.116** H_0 not rejected Sample: 32	**r_s= - 0.624** H_0 rejected Sample: 32	**r_s=-0.05** H_0 not rejected Sample: 32

Table 5-7 Correlation table for countries with high general, non-graduate, youth, and young non-graduate unemployment

For the set of 32 countries where the general unemployment rate, the unemployment rate of non-graduates, the youth unemployment rate, and the unemployment rate of young non-graduates are all high, we observe that:

- Among these countries, GDP per capita (PPP) and the Alpha Rate exhibit a strong correlation (0.737). This suggests that economic development can significantly impact student enrolment numbers and, consequently, the number of graduates in a country.
- No correlation is found between the Alpha Rate and the general unemployment rate (null hypothesis not rejected), indicating that these two variables are independent.
- The correlation between the Alpha Rate and the graduate unemployment rate is -0.537, while the correlation between the Alpha Rate and the young graduate unemployment rate is -0.624. These are relatively strong correlations within this set of 32 countries. This suggests that variations in young graduate unemployment rates significantly influence fluctuations in student enrolment numbers: a decrease in graduate or young graduate unemployment likely encourages more individuals to pursue higher education. It is important to note that these are countries where all types of unemployment are high.

Next, we will analyze the effect of "obtaining a degree" across these 32 countries.
Specifically, we will examine differences between the unemployment rates of non-graduates and graduates in countries where general, non-graduate, youth, and young non-graduate unemployment rates are all above their respective medians. Figure 5-4 illustrates the positioning of these countries concerning the effect of a degree on employment. Blue arrows represent countries where holding a degree improves access to employment, while red arrows indicate countries where holding a degree negatively impacts employability.
Among these 32 countries, 8 countries (25%) exhibit high levels of all four types of unemployment and also have a higher unemployment rate for graduates compared to non-graduates. In other words, in these countries, high unemployment coexists with the paradoxical situation where having a degree is a disadvantage for employment.

In the remaining 24 countries (75%), the unemployment rate for graduates is lower than for non-graduates. The magnitude of this difference is represented by the length of the blue arrows in Figure 5-4, averaging around 6%. For comparison, in a total of 88 countries analyzed in Chapter 4, there were 55 countries where the unemployment rate for non-graduates was higher than that of graduates. The average difference in these 55 countries was 4.5%. We observe that this difference increases by 1.5 percentage points within this subset of 24 countries compared to the broader sample of 55 countries. These 24 countries generally struggle to employ non-graduates, as their unemployment rate is above the global median for this category. Consequently, the gap between graduate and non-graduate unemployment rates is larger in these countries than in the full sample of 88 countries, making "having a degree" a stronger protective factor against unemployment.

Another interesting aspect of this data set is the correlation between GDP per capita (PPP) and the difference between non-graduate and graduate unemployment rates (i.e., the length of the arrows). This correlation is -0.143 and statistically non-significant. In contrast, across all countries, this correlation is -0.28 and statistically significant. This means that in countries where general, youth, and non-graduate unemployment rates are high, economic development is not statistically linked to the difference between graduate and non-graduate unemployment rates.

Among the 24 countries where having a degree protects against unemployment (within the subset of 32 countries), the graduate unemployment rate is, on average, 6 percentage points lower than the non-graduate unemployment rate. By comparison, among the 88 countries analyzed in Chapter 4, 55 countries had a non-graduate unemployment rate 4.5 percentage points higher than the graduate unemployment rate. The length of the 24 blue arrows in Figure 5-4 highlights the fact that holding a degree provides stronger protection against unemployment in wealthy countries with high unemployment than in the broader set of 88 countries studied in this monograph.

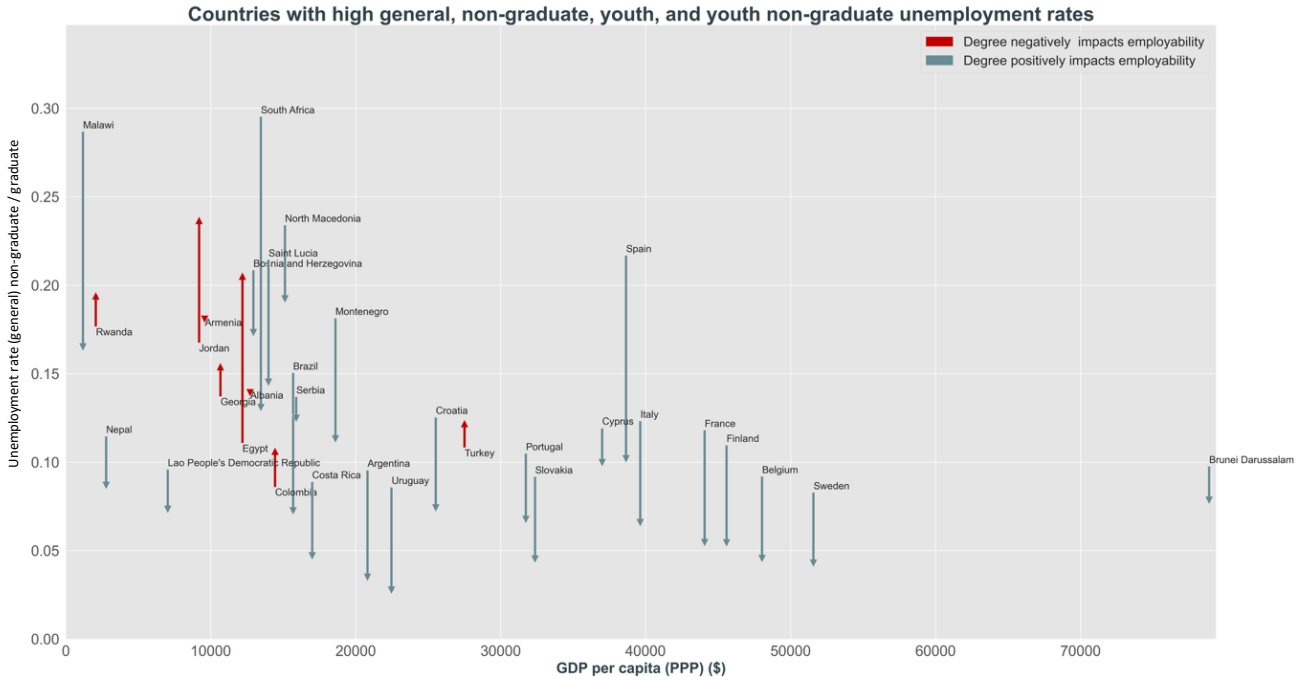


Figure 5-4 Difference between unemployment rates of non-graduates and graduates for the 32 economies with high general, non-graduate, youth, and youth non-graduate unemployment rates

In the second section, we will analyze the effect of "having a diploma" across these 32 countries, which exhibit high unemployment rates, focusing on the youth population. We will then examine the impact of a diploma on the employability of young graduates.

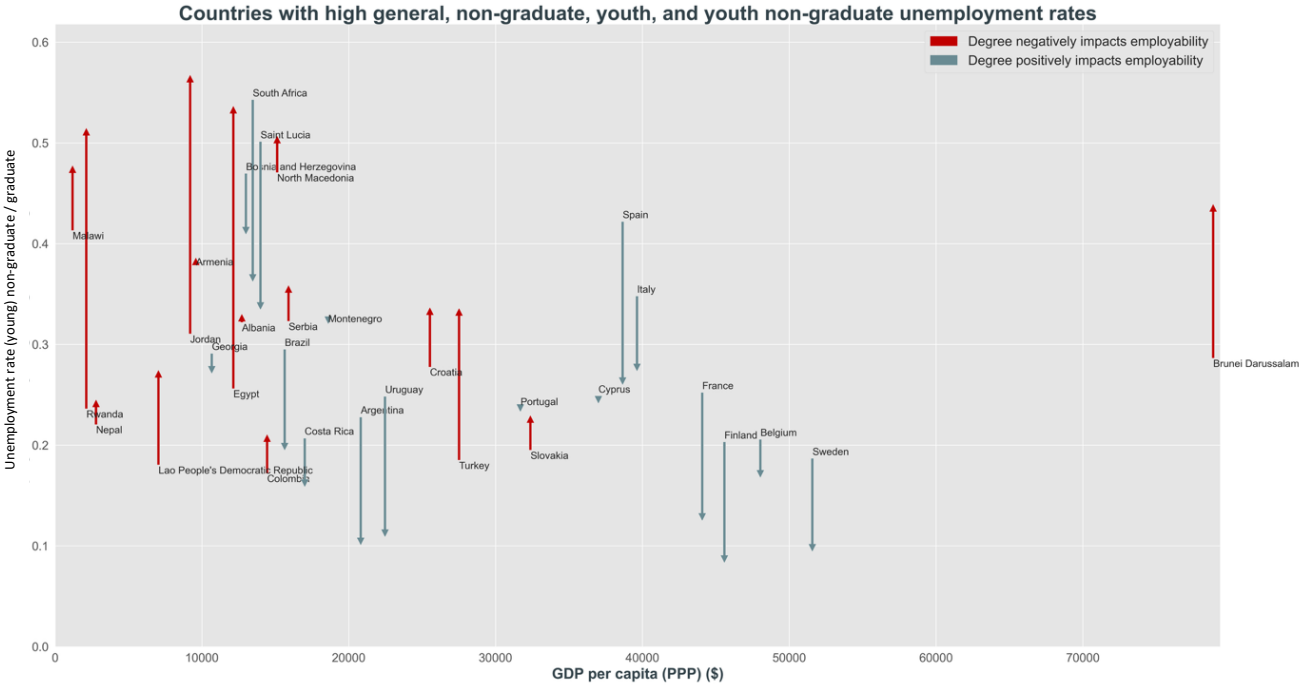


Figure 5-5 Difference between the unemployment rates of young non-graduates and young graduates for the 32 economies with high general, non-graduate, youth, and youth non-graduate unemployment rates

It is observed that among young people, there are more red arrows (50% instead of 25%) compared to the previous population (all age brackets, figure 5-4). In these countries, which generally perform poorly in employment terms, being "young and a graduate" does not provide an advantage in employability.

Furthermore, among the 32 countries where general, non-graduate, youth, and youth non-graduate unemployment rates are high, we observe that in the 16 countries where the unemployment rate of graduates is lower than that of non-graduates (blue arrows, 16 countries, 50%), the median magnitude of the difference between the unemployment rates of young graduates and young non-graduates is 9.6%.

We can observe that Figure 5-5 shows a higher number of red arrows than the previous one (Figure 5-4). Across all countries for which data on youth graduate and non-graduate unemployment rates are available (80 countries), 40 countries have a graduate unemployment rate higher than that of non-graduates (Figure 4-22, page 42). The median magnitude of the differences across these 40 countries was 5.1%, approximately half the difference observed among the 16 countries mentioned above. In other words, in countries with high unemployment levels, being young and a graduate is more detrimental to employability than being simply a non-graduate.

Thus, in high-unemployment countries, the countries where obtaining a diploma leads to lower employability show an even more pronounced effect among the youth population. In this subset of countries, the unemployment rate of young graduates is, on average, 10 percentage points higher than that of young people in general. Consequently, in countries with high general, non-graduate, youth, and youth non-graduate unemployment rates, being young significantly hinders the employability of graduates.

Country (4 Unemployment Rates above respective medians) Countries with GU, NGU, YU, and YNGU above respective medians	Alpha Rate 2017	GDP per capita (PPP) in 2017	Young graduate unemployment rate 2017	Young non-graduate unemployment rate 2017	Absolute Delta (young graduates - young non-graduates)	Relative Delta (young graduates - young non-graduates)
Albania	0.69	12719	0.33	0.31	0.02	1.06
Argentina	0.87	20815	0.10	0.23	-0.13	0.43
Armenia	0.54	9582	0.39	0.38	0.01	1.02
Belgium	0.69	48034	0.17	0.21	-0.04	0.80
Bosnia and Herzegovina	0.55	12946	0.41	0.46	-0.05	0.88
Brazil	0.51	15635	0.19	0.30	-0.10	0.65
Brunei Darussalam	0.42	78873	0.44	0.28	0.16	1.58
Colombia	0.57	14437	0.21	0.16	0.05	1.29
Costa Rica	0.56	17003	0.16	0.21	-0.05	0.75
Croatia	0.74	25526	0.34	0.27	0.07	1.25
Cyprus	0.67	37003	0.24	0.25	-0.01	0.96
Egypt	0.35	12138	0.54	0.25	0.29	2.17
Finland	0.86	45585	0.08	0.21	-0.12	0.40

France	0.64	44074	0.12	0.26	-0.13	0.49
Georgia	0.61	10669	0.27	0.29	-0.02	0.92
Italy	0.60	39630	0.27	0.35	-0.08	0.78
Jordan	0.29	9196	0.57	0.30	0.27	1.88
Laos	0.17	7038	0.27	0.17	0.10	1.60
Malawi	0.01	1186	0.48	0.40	0.07	1.18
Montenegro	0.72	18604	0.32	0.32	0.00	0.99
Nepal	0.14	2787	0.25	0.21	0.03	1.16
North Macedonia	0.53	15122	0.51	0.46	0.05	1.10
Portugal	0.63	31688	0.23	0.24	-0.01	0.97
Rwanda	0.07	2074	0.52	0.22	0.30	2.37
Saint Lucia	0.24	13986	0.33	0.50	-0.17	0.66
Serbia	0.49	15897	0.36	0.31	0.04	1.14
Slovakia	0.57	32371	0.23	0.19	0.04	1.23
South Africa	0.21	13461	0.36	0.55	-0.18	0.66
Spain	0.87	38651	0.26	0.43	-0.17	0.61
Sweden	0.68	51573	0.09	0.19	-0.10	0.49
Turkey	0.58	27510	0.34	0.18	0.16	1.90
Uruguay	0.66	22469	0.11	0.25	-0.14	0.43

Table 5-8 Comparative table of youth graduate and youth non-graduate unemployment rates for the 32 countries with high general unemployment (GU), non-graduate unemployment (NGU), youth unemployment (YU), and youth non-graduate unemployment

Without providing an exhaustive commentary on this table, it is evident that, in these countries significantly affected by unemployment across all categories, it is primarily in the poorest countries that obtaining a degree distances individuals from employment. However, among countries with similar levels of development, very different situations can be observed. In general, these countries experience high unemployment rates for both young graduates and young non-graduates. Nevertheless, the countries where obtaining a degree offers the least protection against unemployment are mainly low-income countries.

Similarly, at comparable levels of development, the impact of a degree on unemployment rates varies significantly. For example, in Uruguay and Turkey, which have similar levels of development, the unemployment rates for young non-graduates and young graduates in Turkey are 18% and 34%, respectively, whereas in Uruguay, they are 25% and 11%, demonstrating an inverse effect. A similar pattern is observed in France and Portugal, where young non-graduates have similar unemployment rates (25%). However, in France, the positive effect of a degree is very pronounced (a twofold reduction in the unemployment rate), whereas in Portugal, it is weak (23% instead of 25%).

Thus, at equivalent Alpha Rates and similar levels of development, the protective effect of a degree is more pronounced in certain economies. Portugal and France exhibit relatively close unemployment rates for young non-graduates. Yet, while obtaining a degree doubles the employability of young French graduates, in Portugal, the impact of a degree on youth employability is observed at only 12%.

5.3.3 Countries with low general and youth unemployment rates

The following two figures illustrate the "degree effect" in countries where both the general unemployment rate and youth unemployment rate are below the respective medians. In a second stage, we will measure the impact of a degree on employability in countries where general unemployment rates, as well as those for non-graduates, young people, and young non-graduates, are below the corresponding medians. These countries total 32.

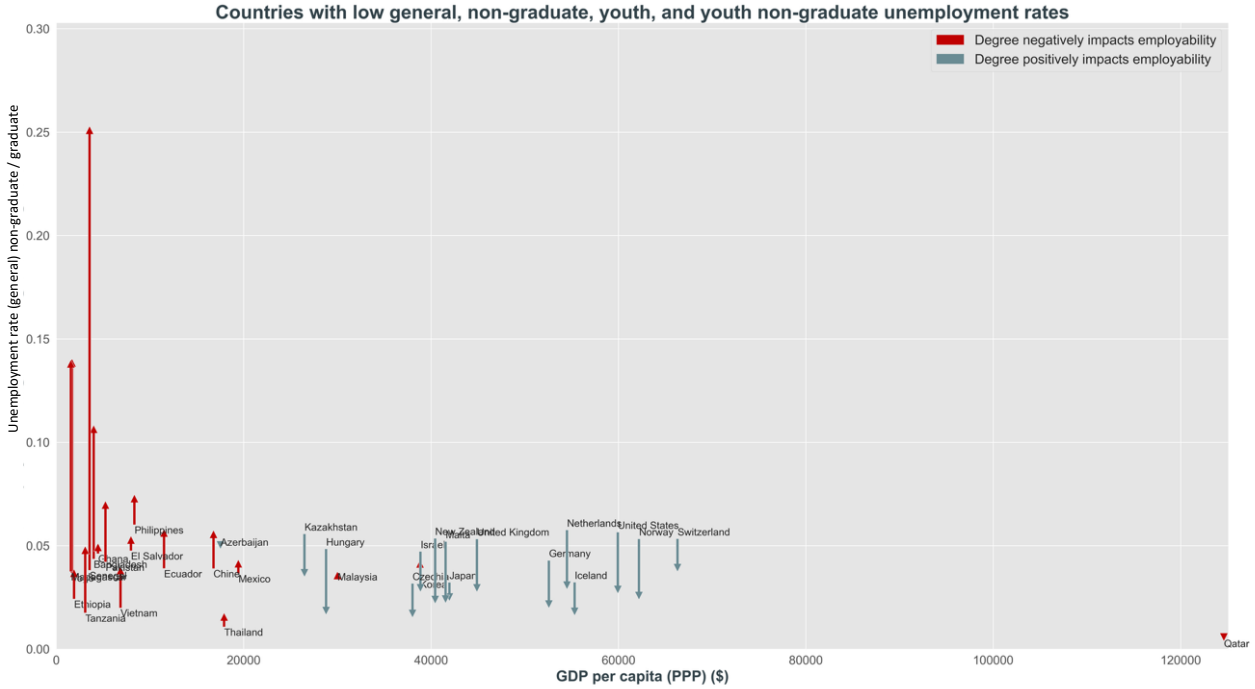


Figure 5-6 Difference between the unemployment rate of non-graduates and graduates across countries where all four unemployment rates are low (for the entire population)

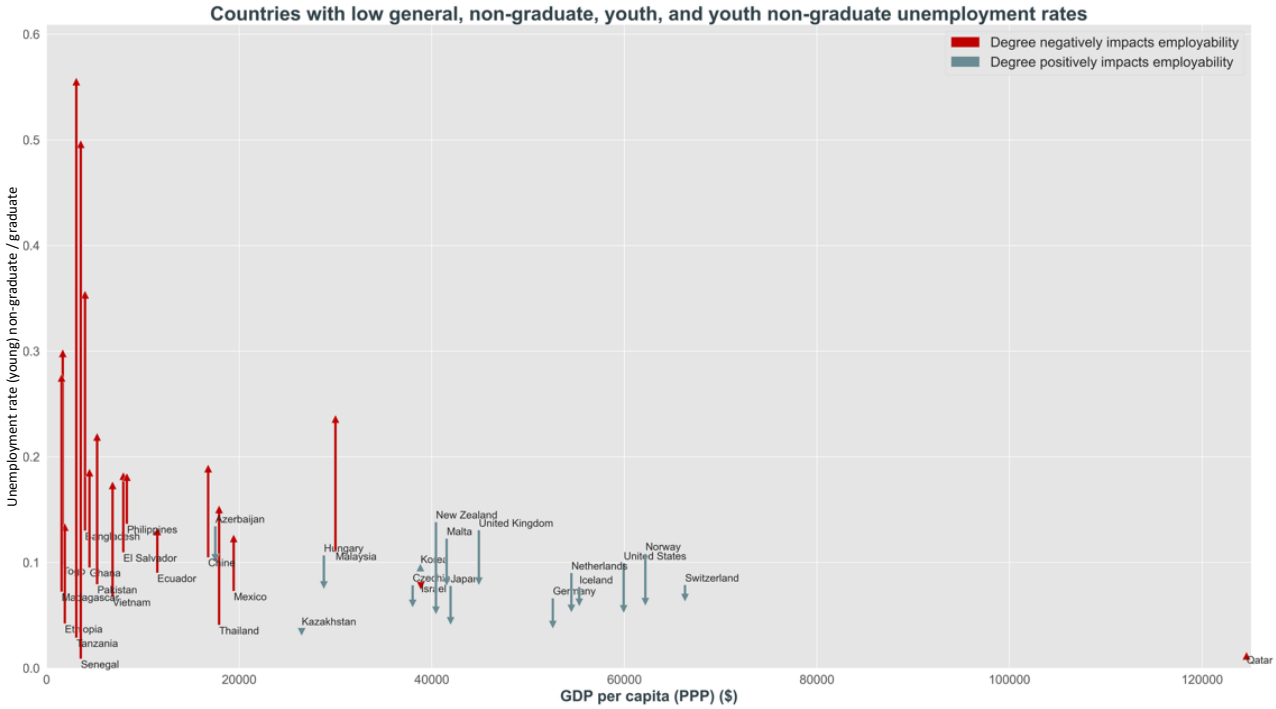


Figure 5-7 Difference between the unemployment rates of young non-graduates and young graduates in countries with low general unemployment rates, as well as low unemployment rates for non-graduates, youth, and young non-graduates

The red arrows are significantly larger than in the previous case (countries where unemployment rates are above the median), while the blue arrows are smaller than in the previous case (same reference) in both figures. This is clearly visible in the following table. In countries with low unemployment rates, as shown in the two preceding graphs, when a negative impact of holding a degree is observed—indicated by the red arrows—it tends to be more pronounced for young graduates than for graduates in general. Conversely, the positive impact of a degree is less marked in these countries than in those with high unemployment rates, whether for graduates overall or for young graduates specifically. This confirms the protective effect of higher education in advanced economies with high unemployment, which struggle to integrate less-educated labor into the workforce.

Country (4 unemployment rates below respective medians)	Alpha 2017	GDP per capita (PPP) 2017	Young graduate unemployment rate 2017	Young non-graduate unemployment rate 2017	Absolute Delta (graduates – non-graduat)	Relative Delta (graduates / non-graduates)
Azerbaijan	0.32	17 525	0.100	0.138	-0.038	0.73
Bangladesh	0.19	3 998	0.358	0.122	0.236	2.94
China	0.39	16 782	0.193	0.096	0.097	2.00
Czechia	0.60	38 020	0.057	0.082	-0.025	0.70
Ecuador	0.50	11 501	0.133	0.082	0.052	1.63
El Salvador	0.31	7 973	0.186	0.101	0.085	1.84
Ethiopia	0.08	1 897	0.138	0.034	0.104	4.09
Germany	0.65	52 574	0.037	0.070	-0.032	0.53
Ghana	0.16	4 457	0.189	0.087	0.103	2.18
Hungary	0.49	28 799	0.075	0.110	-0.036	0.68
Ireland	0.83	55 322	0.058	0.080	-0.022	0.73
Israel	0.62	38 868	0.074	0.072	0.002	1.03
Japan	0.62	41 959	0.041	0.081	-0.041	0.50
Kazakhstan	0.56	26 491	0.030	0.041	-0.010	0.74
Korea	0.79	38 824	0.099	0.100	0.000	1.00
Madagascar	0.05	1 560	0.279	0.064	0.215	4.37
Malaysia	0.43	30 004	0.240	0.102	0.137	2.34
Malta	0.50	41 549	0.078	0.126	-0.048	0.62
Mexico	0.41	19 432	0.127	0.064	0.062	1.97
Netherlands	0.77	54 503	0.052	0.093	-0.041	0.56
New Zealand	0.67	40 439	0.051	0.142	-0.091	0.36
Norway	0.82	62 183	0.059	0.112	-0.053	0.53
Pakistan	0.10	5 249	0.223	0.071	0.152	3.15
Philippines	0.36	8 340	0.185	0.128	0.057	1.45
Qatar	0.09	124 609	0.016	0.004	0.011	3.62
Switzerland	0.52	66 300	0.063	0.082	-0.019	0.76
Tanzania	0.03	3 090	0.559	0.020	0.539	27.66
Thailand	0.51	17 917	0.154	0.032	0.122	4.77
Togo	0.13	1 683	0.302	0.089	0.213	3.40
United Kingdom	0.49	44 896	0.078	0.134	-0.056	0.58
United States	0.80	59 928	0.052	0.103	-0.051	0.51
Vietnam	0.30	6 858	0.177	0.059	0.118	3.01

Table 5-9 Comparative table of youth graduate and youth non-graduate unemployment rates for the 32 countries with low general unemployment rates (TU), non-graduate unemployment rates (NGU), youth unemployment rates (YU), and youth non-graduate unemployment rates (Y)

This table of the countries least affected by unemployment (32 countries) highlights, through its size, the extent to which countries are polarised between those that manage the issue of employability effectively and those that manage it "less effectively". Representing a total of 32 + 32 = 64 out of the 89 countries studied, we observe that countries positioned above or below the medians for one of the rates are generally positioned similarly for all of the rates.

Among these countries, in half of them (16/32), holding a degree leads to better employability, and conversely, in the other half, it does not.
Similar to the previous table, in countries that manage employment better, holding a degree is particularly penalizing in poorer countries. In some cases, this manifests in extreme ratios (graduate unemployment to non-graduate unemployment ratios exceeding 3 in Madagascar, Ethiopia, and Pakistan; a factor of 28 in Tanzania). However, this mismatch between graduates and the local labor market is also evident in more advanced economies such as Mexico, Thailand, and Malaysia (ratios of 2, 4.8, and 2.3, respectively), in countries with GDP per capita (PPP) ranging from $18K to $30K per capita. Nonetheless, the misalignment between young graduates and the labor market is also observed in more developed economies such as Mexico, Thailand, and Malaysia.

5.4 Unsupervised machine learning analysis: hierarchical clustering using Ward's Method

5.4.1 Ward's Method

Sections 5.1 to 5.3 explored country partitions based on predefined criteria. It seemed important to investigate whether these observations might actually conceal opposing phenomena in different categories of countries, as averaging effects could obscure certain realities.

Thus, the third partitioning method (dividing data into multiple disjoint sets) is based on the use of a hierarchical agglomerative clustering (HAC) algorithm. The purpose of such an algorithm is to classify elements into the most homogeneous clusters possible, presenting comparable properties without prior selection bias. This method relies on the classical principle of "good clustering," defined by the following conditions:
A good clustering ensures that:

- Elements within the same cluster share strong similarities.
- Elements from different clusters exhibit weak similarities.

Statistically, this translates to:

- Low variability within a cluster (or low intra-cluster variability), meaning there is little variation within a single cluster.
- High variability between different clusters (or high inter-cluster variability), indicating significant variation between clusters.

Reminder of a few definitions:

- The total inertia of the data set is defined as the sum of the variances[39] of all variables in the data set.
- The intra-cluster inertia is defined as the sum of the variances of all variables among elements within a cluster.
- The inter-cluster inertia is defined as the sum of the variances of the cluster centroids (the centroid of a cluster corresponds to the point whose coordinates represent the mean of the variables of the points in the cluster).

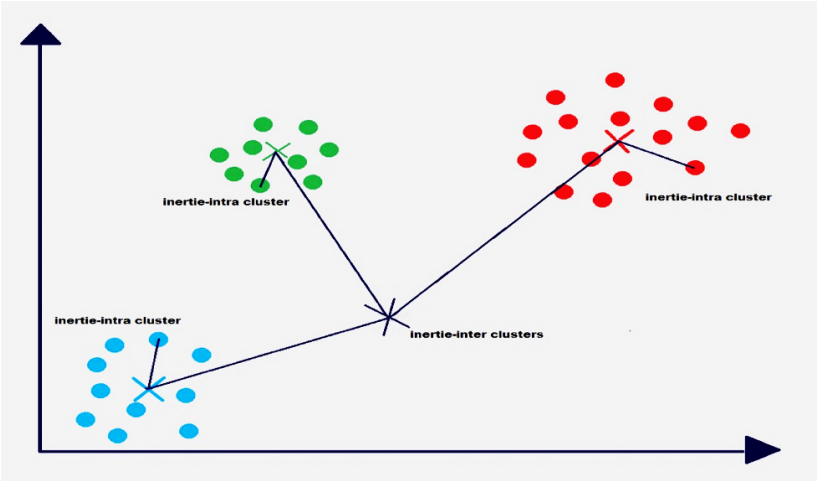


Figure 5-8 Clustering using Ward's Method

The principle of Ward's algorithm is to:

- Distribute elements into clusters in such a way that inter-cluster inertia is maximized while intra-cluster inertia is minimized. Thus, the goal is to maximize inter-cluster inertia and minimize intra-cluster inertia.

Huygens' theorem implies that the total inertia, i.e., the sum of variances of all elements, is decomposed into the sum of intra-cluster and inter-cluster inertias.

Thus :

$$\sum_{q=1}^{Q}\sum_{i=1}^{n_q} d^2(x_{iq}, G) = \sum_{q=1}^{Q}\sum_{i=1}^{n_q} d^2(x_{iq}, g_q) + \sum_{q=1}^{Q} n_q d^2(g_q, G)$$

Total inertia = Inertia-intra + Inertia-inter

Where:

- x_{iq}, is the i-th element of cluster q
- Q, is the total number of clusters
- n_q, is the number of elements in cluster q
- g_q, is the centroid of cluster q, i.e., the mean of the variables in cluster q
- G, i s the centroid of the entire data set

[39] Variance is caracterised as a measure of dispersion around the mean (cf Annex 8.4.5).

Therefore, maximizing inter-cluster inertia is equivalent to minimizing intra-cluster inertia since total inertia remains constant. Consequently, the algorithm can focus on a single parameter, either maximizing inter-cluster inertia or minimizing intra-cluster inertia.

The specificity of Ward's Method lies in applying the previously described algorithm while initially assuming that each element (each vector in the data set) is its own cluster. Thus, at the beginning, there are as many clusters as there are rows in the data set. At each step, the hierarchical agglomerative clustering algorithm is applied by merging two clusters in a way that minimizes the increase in inter-cluster inertia due to their aggregation. Hence, clusters a and b can be merged if their aggregation results in the smallest decrease in inertia between a and b, thereby allowing for maximum inter-variability.

Mathematically, the inertia of a new cluster formed after merging clusters a and b is given by:

$$Inertia(a \cup b) = Inertia(a) + Inertia(b) + \frac{n_a * n_b}{n_a + n_b} * d(g_a, g_b),$$

Where:

- n_a and n_b, are the sizes of clusters a and b
- $\boldsymbol{d(g_a, g_b)}$, is the Euclidean distance between the centroids of clusters a and b

To merge clusters a and b, the objective is to minimize the intra-cluster inertia of the new cluster $\boldsymbol{a \cup b}$. This is achieved by minimizing the term $\frac{n_a * n_b}{n_a + n_b}$ (first term) and the term$\boldsymbol{d(g_a, g_b)}$ ((second term). The minimization of these two terms suggests different approaches to merging these clusters.

- Minimizing the first term implies grouping objects of comparable sizes (it is more likely that a cluster with three elements will be merged with a similarly sized cluster rather than with a much larger one).[40]
- Minimizing the second term implies grouping clusters whose centroids are close. Thus, clusters with the most similar means are more likely to be merged.

The hierarchical agglomerative clustering algorithm using Ward's Method therefore merges, at each step, the two clusters that result in the partition with the smallest intra-cluster inertia.

We applied this algorithm to our data set of 89 countries described in section 5.1, which includes only those for which complete data on the six unemployment rates is available. To effectively present the hierarchical clustering of elements, we chose to use a dendrogram representation.

Reminder, Eight variables were considered for this clustering: GDP per capita (PPP) in dollars, Alpha Rate, general unemployment rate, youth unemployment rate, graduate unemployment rate, young graduate unemployment rate, non-graduate unemployment rate, and the unemployment rate of young non-graduates.

[40] If we have two clusters of identical size equal to m, the **first term** is m2/2m=m/2. If we have two clusters of sizes m and n=m+d (where the second cluster is larger than the first), the **first term** is given by (m2+md)/(2m+d), which is greater than (m2+md)/(2m+2d)=m(m+d)/2(m+d)=m/2 (since increasing the denominator decreases the ratio). This implies that, in order to minimize the first term, it is preferable to merge clusters of similar sizes rather than clusters with significantly different sizes.

Before applying the algorithm, data preparation was necessary. This preparation involved transforming the data using the standardized normalization technique, which ensures that all variables are distributed within the same order of magnitude. This prevents any variable from dominating others during the clustering process. For instance, if one variable ranges from 1,000 to 10,000 while another ranges from 10 to 100, the first variable would overshadow the second. Applying this algorithm to our data set of 89 countries resulted in the clusters described in figure 5-9.

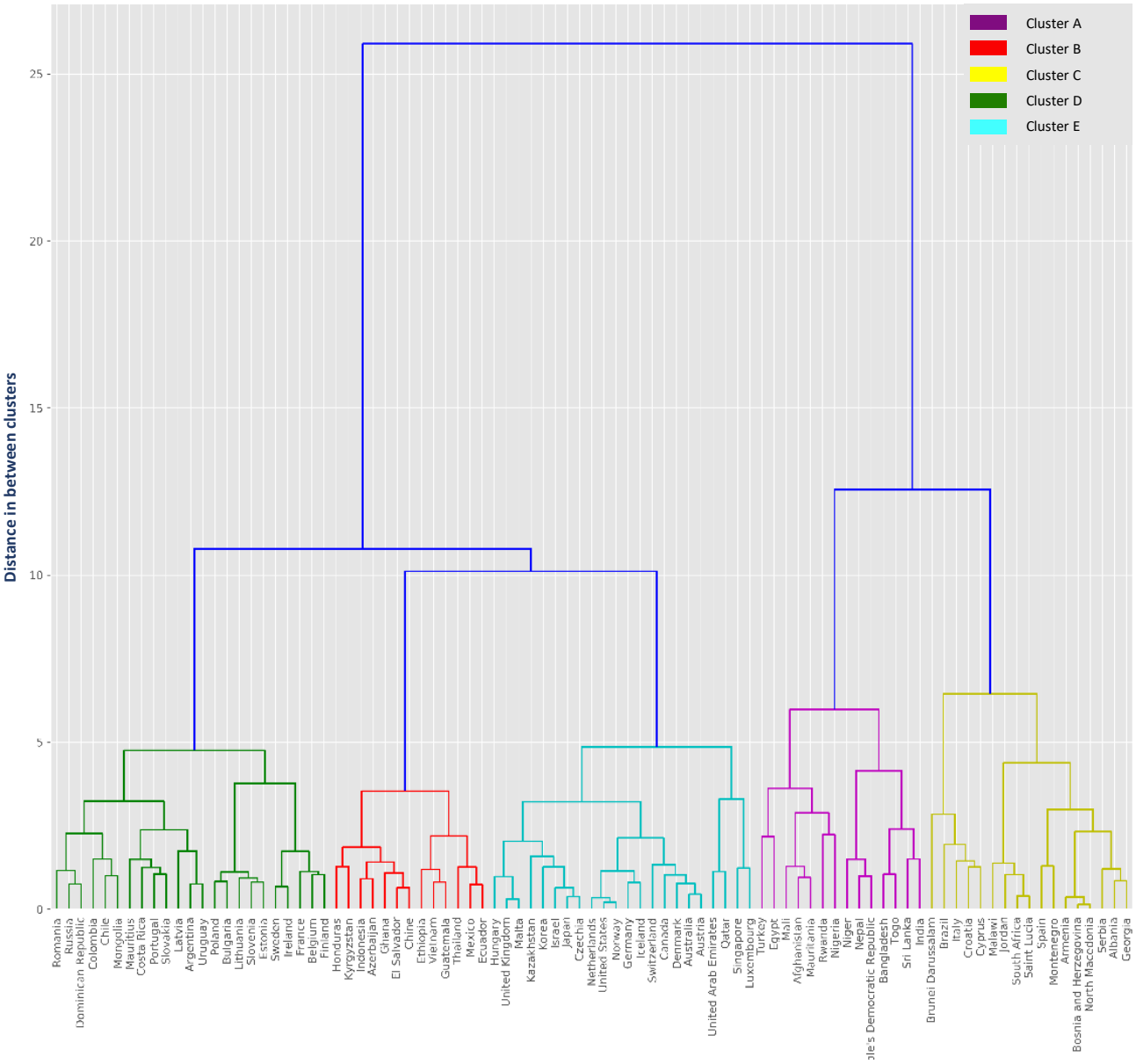


Figure 5-9 Distribution of 89 countries into clusters using Ward's method

Following the principles of the "best clustering" approach explained earlier, we observe that after applying the algorithm to our data set, the countries are statistically grouped into five very distinct clusters, each sharing specific characteristics, on the basis of the values of the eight indicated variables. The following patterns are observed within the resulting clusters:

- Based on the selected variables, France is closest to Belgium and Finland. These three countries belong to the same cluster as Ireland and Sweden (cluster D).

- The United States is closest to the Netherlands and Norway. Additionally, these countries, along with Canada, Denmark, Iceland, Switzerland, Germany, Australia, and Austria, form a large cluster that is closely related to the cluster containing South Korea, Israel, Japan, and the United Kingdom (cluster E).

- Russia forms a sub-cluster with Romania, Chile, Colombia, Mongolia, Costa Rica, the Dominican Republic, Mauritius, Slovakia, Uruguay, Poland, Bulgaria, Lithuania, and Slovenia. This sub-cluster is more similar to the cluster containing France than to the one containing the United States.

- Brazil, Italy, Croatia, and Cyprus form a distinct sub-cluster, separate from the previously mentioned groups, and closer to the cluster containing Spain, Montenegro, Albania, Georgia, Armenia, Bosnia and Herzegovina, and Serbia. Together, they form cluster C.

- Cluster C is closer to cluster A, and together they form a broader grouping. The other grouping consists of cluster E, B, and D, where cluster E and B are more similar to each other than to cluster D.

Ward's algorithm thus leads to the following distribution of the 89 countries:
Cluster A: Afghanistan, Bangladesh, Egypt, India, Laos, Mali, Mauritania, Nepal, Niger, Nigeria, Rwanda, Sri Lanka, Togo, and Turkey.
Cluster B: Azerbaijan, China, Ecuador, El Salvador, Ethiopia, Ghana, Guatemala, Honduras, Indonesia, Kyrgyzstan, Mexico, Thailand, and Vietnam.
Cluster C: Albania, Armenia, Bosnia and Herzegovina, Brazil, Brunei Darussalam, Croatia, Cyprus, Spain, Georgia, Italy, Jordan, Malawi, Montenegro, North Macedonia, Saint Lucia, Serbia, and South Africa.
Cluster D: Argentina, Belgium, Bulgaria, Chile, Colombia, Costa Rica, Dominican Republic, Estonia, Finland, France, Ireland, Latvia, Lithuania, Mauritius, Mongolia, Poland, Portugal, Romania, Russia, Slovakia, Sweden, and Uruguay.
Cluster E: Australia, Austria, Germany, Canada, Czech Republic, South Korea, Denmark, Hungary, Iceland, Israel, Japan, Kazakhstan, Luxembourg, Malta, Norway, Netherlands, Qatar, Singapore, Switzerland, United Arab Emirates, United Kingdom, and United States.

The cartographic illustration of this distribution is detailed below for each cluster.

5.4.2 Analysis of resulting clusters

In the previous section, Ward's Method allowed us to group 89 countries into five distinct clusters. The following figure represents the medians of each variable for each cluster.

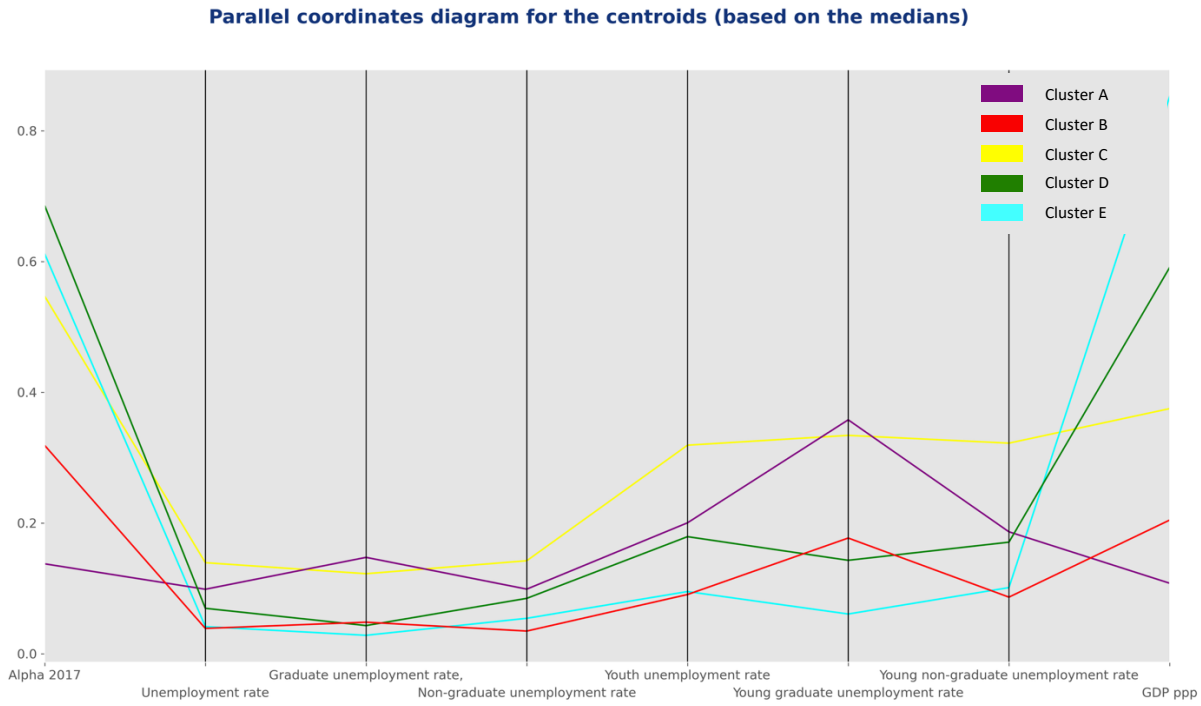


Figure 5-10 Medians of higher education access rate, general unemployment rate, youth unemployment rate, graduate unemployment rate, young graduate unemployment rate, non-graduate unemployment rate, and young non-graduate unemployment rate by clusters

For ease of interpretation, we have named each of the clusters from A to E according to increasing average GDP per capita (PPP). The hierarchical structure resulting from Ward's method (figure 5–9) reveals proximities between the various clusters that do not follow this linear order A–B–C–D–E. However, for the purposes of the subsequent interpretation, this mode of presentation proved much easier to follow.

It should be noted that the average GDP per capita (PPP) increases by a factor of approximately two from one cluster to the next, while very different patterns are observed in terms of unemployment rates.

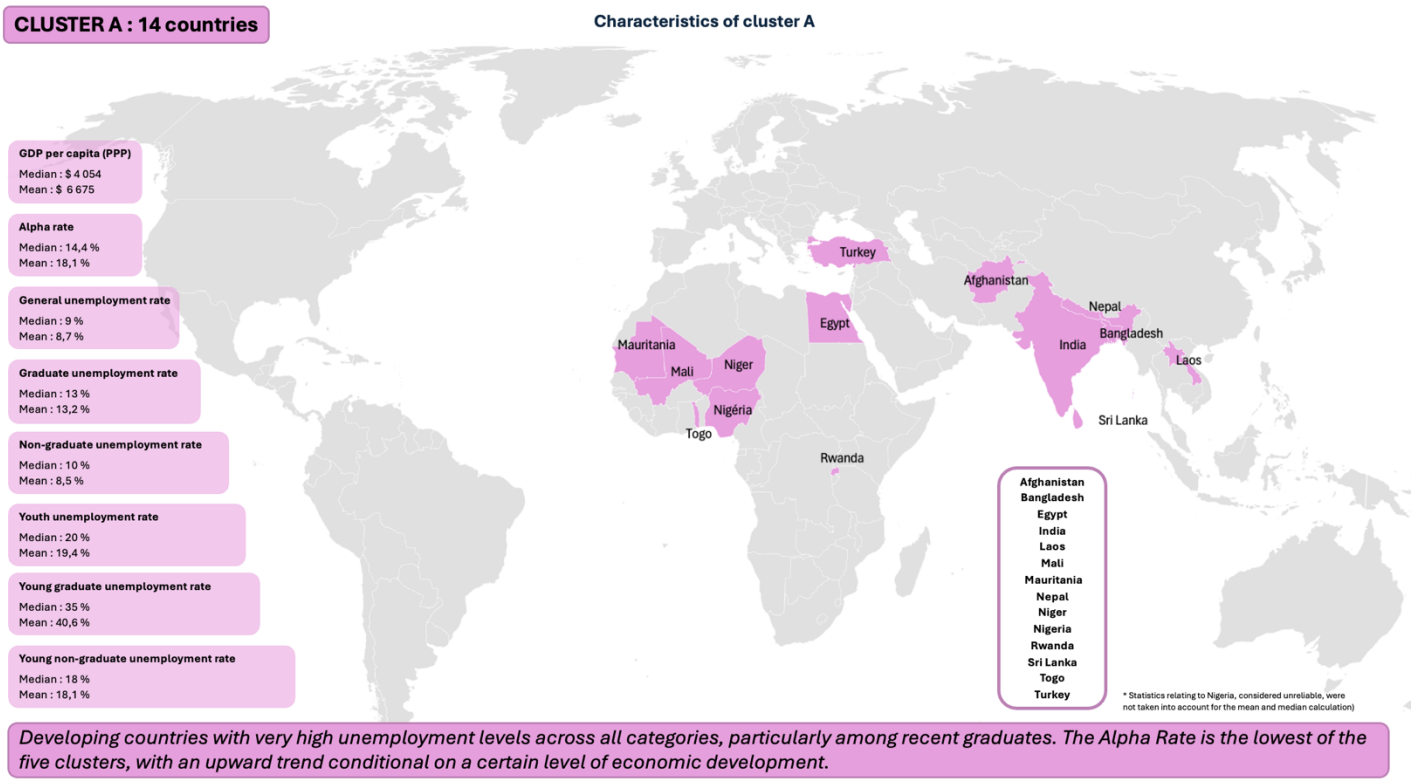


Figure 5-11 World Map, Cluster A

Cluster A: Afghanistan, Bangladesh, Egypt, India, Laos, Mali, Mauritania, Nepal, Niger, Nigeria, Rwanda, Sri Lanka, Togo, Turkey.

- In this cluster, GDP per capita (PPP) is the lowest, with a median of $4,054 and a mean of $6,675 (GDP per capita (PPP) ranges from $1K (Niger) to $27.5K (Turkey)).
- The Alpha Rate is also the lowest, with a median of 14.4% and a mean of 18.1% within the cluster (ranging from 3% (Niger) to 57% (Turkey)).
- Graduate unemployment rates (13%, 13.2%) and recent graduate unemployment rates (35%, 40.6%) are the highest.
- General unemployment rate (9%, 9%) and non-graduate unemployment rate (10%, 9%) are average.
- Youth unemployment rate (20%, 19.4%) and youth non-graduate unemployment rate (18%, 18.1%) are relatively high.
- There is a dual employment disadvantage: one affecting young people and another affecting graduates.

	GDP per capita (PPP)	**General unemployment rate**	**Graduate unemployment rate**	**Non-graduate unemployment rate**	**Youth unemployment rate**	**Young graduate unemployment rate**	**Young non-graduate unemployment rate**
Alpha Rate	**r_s=0.829** H_0 rejected p=0.0	**r_s=-0.007** H_0 not rejected p=0.982	**r_s=-0.13** H_0 not rejected p=0.659	**r_s=-0.147** H_0 not rejected p=0.615	**r_s=0.279** H_0 not rejected p=0.334	**r_s=-0.103** H_0 not rejected p=0.725	**r_s=0.147** H_0 not rejected p=0.615

Table 5-10 Correlation table, Cluster A

The following observations can be made regarding the correlations between variables in Cluster A:

- GDP per capita (PPP) shows a strong correlation with the Alpha Rate. Thus, economic development in these countries significantly impacts young people's access to higher education, and vice versa.
- There is no statistically significant correlation between the Alpha Rate and any other unemployment-related variable.

In Cluster A, composed of low- to middle-income countries, youth unemployment rates and graduate unemployment rates are significantly higher than the general unemployment rate. Economic development and access to higher education are strongly correlated, whereas access to higher education and unemployment rates (regardless of type) are independent.

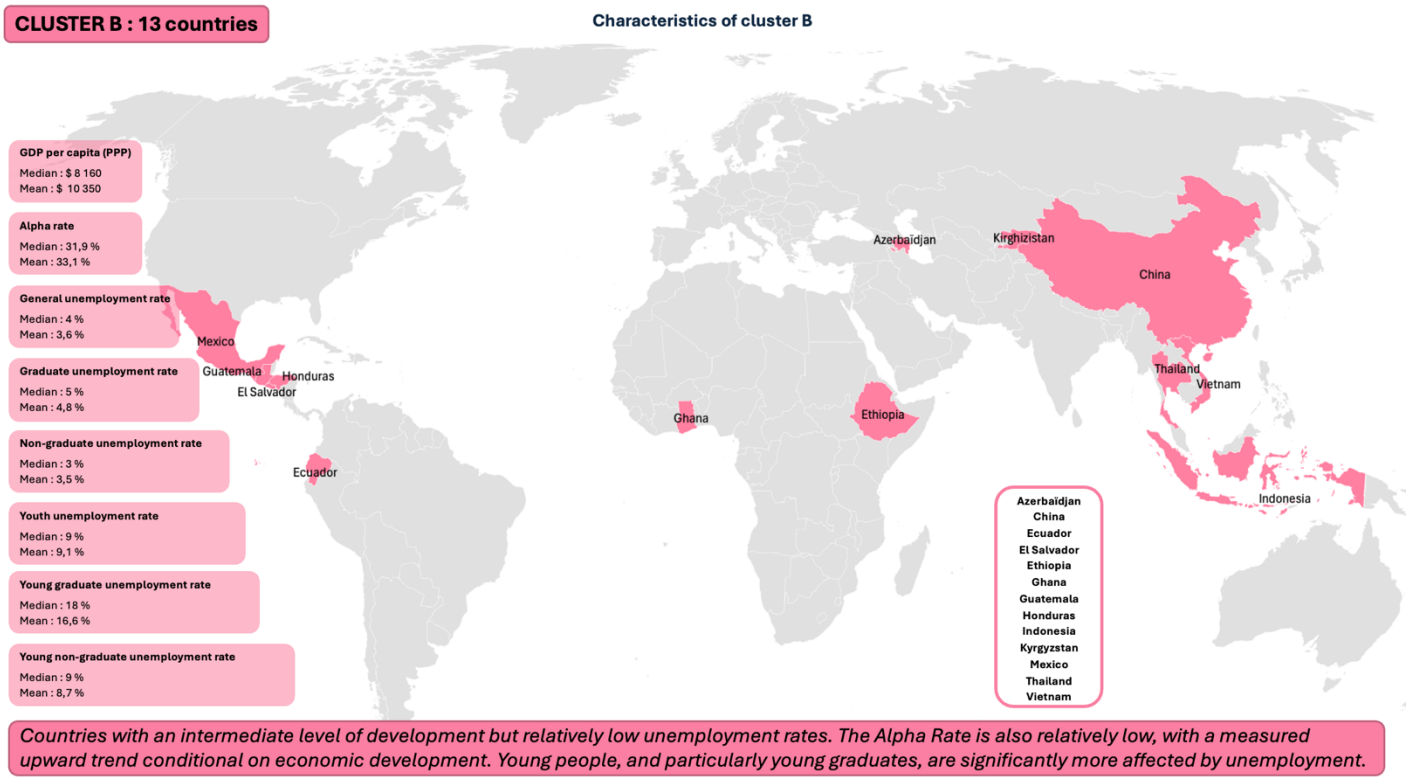


Figure 5-12 World Map, Cluster B

Cluster B: Azerbaijan, China, Ecuador, El Salvador, Ethiopia, Ghana, Guatemala, Honduras, Indonesia, Kyrgyzstan, Mexico, Thailand, Vietnam.

- The median GDP per capita (PPP) of the countries in this cluster is relatively low ($8.2K). The mean is $10.3K (GDP per capita (PPP) ranges from $2K (Ethiopia) to $19K (Mexico)).
- The median Alpha Rate (31.9%) in Cluster B is relatively low. The mean is 33.1% (ranging from 7% (Ethiopia) to 50% (Thailand)).
- The general unemployment rate (4%, 3.6%), the non-graduate unemployment rate (3%, 3.5%), the youth unemployment rate (9%, 9.1%), and the youth non-graduate unemployment rate (9%, 8.7%) are the lowest.
- The graduate unemployment rate (5%, 4.8%) is low and close to the general unemployment rate.
- However, the recent graduate unemployment rate (18%, 16.6%) is twice the youth unemployment rate and four times the general unemployment rate. These are countries where recent graduates are clearly at a disadvantage.

	GDP per capita (PPP)	General unemployment rate	Graduate unemployment rate	Non-graduate unemployment rate	Youth unemployment rate	Young graduate unemployment rate	Young non-graduate unemployment rate
Alpha Rate	r_s=0.621 H_0 rejected p=0.024	r_s=-0.06 H_0 not rejected p=0.845	r_s=0.093 H_0 not rejected p=0.762	r_s=-0.049 H_0 not rejected p=0.873	r_s=0.154 H_0 non-rejected p=0.616	r_s=-0.071 H_0 non-rejected p=0.817	r_s=0.082 H_0 not rejected p=0.789

Table 5-11 Correlation table, Cluster B

The following observations can be made regarding the correlations among the variables for the set of countries in cluster B:

- GDP per capita (PPP) shows a strong correlation with the Alpha Rate, as is the case in all emerging countries: economic development positively impacts access to higher education among young people, and vice versa.
- There is no statistically significant correlation between the Alpha Rate and any of the unemployment rates.

In these emerging low- or middle-income countries (cluster B), all unemployment rates are low or very low. Young graduates appear to be the group that faces the most difficulty accessing employment. While access to higher education and economic development are strongly linked, this access does not correlate with observed unemployment rates.

Comparison between Cluster A and Cluster B

- A twofold difference is observed between the average GDP per capita (PPP) of the two clusters, in favour of Cluster B.
- Cluster B shows an Alpha Rate approximately twice as high as that of Cluster A.
- Cluster B displays unemployment rates around half those of Cluster A, with an even greater gap for graduate unemployment.
- In both clusters, holding a degree appears to be a disadvantage for employability: the unemployment rate of recent graduates is twice that of non-graduate youth and four times the general unemployment rate.

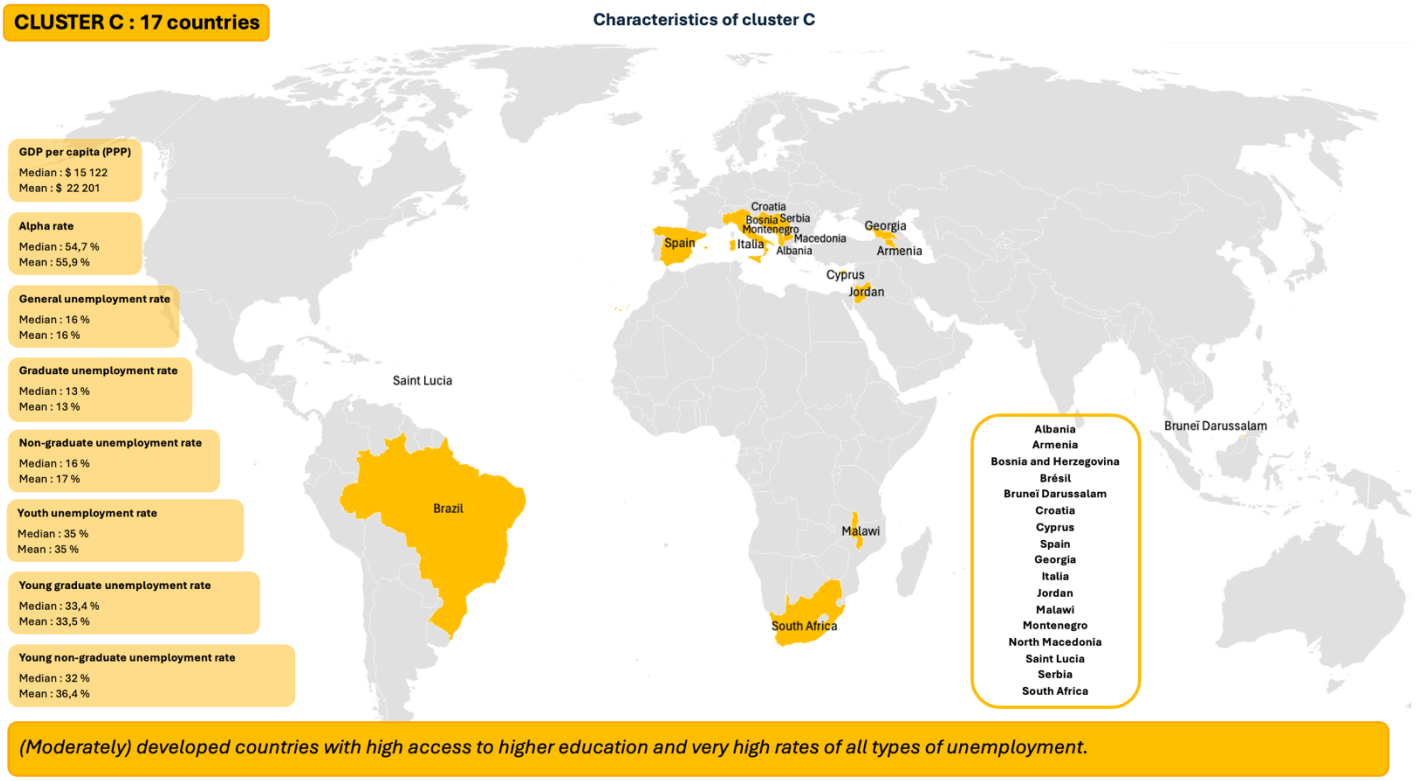


Figure 5-13 World Map, Cluster C

Cluster C: Albania, Armenia, Bosnia and Herzegovina, Brazil, Brunei Darussalam, Croatia, Cyprus, Spain, Georgia, Italy, Jordan, Malawi, Montenegro, North Macedonia, Saint Lucia, Serbia, South Africa.

- The median GDP per capita (PPP) of the countries in this cluster is moderate ($15,122). The mean is $22,201 (GDP per capita (PPP) ranges from $9K (Jordan) to $79K (Brunei Darussalam)).
- The Alpha Rate is relatively high: the median Alpha Rate is 54.3%. The mean is 53% (ranging from 0.9% (Malawi) to 87% (Spain)).
- The general unemployment rate (median = 16%, mean = 16%), the non-graduate unemployment rate (16%, 17%), the youth unemployment rate (34%, 35.7%), and the youth non-graduate unemployment rate (32%, 35.1%) are the highest in comparison with the other clusters.
- The graduate unemployment rate (13%, 13.1%) and the recent graduate unemployment rate (34%, 35.7%) are among the highest.
- It is observed that in this cluster, where both general and youth unemployment rates are high, the effect of a diploma is weak, both in the general population and among young people, with graduates and non-graduates having similar unemployment rates.

	GDP per capita (PPP)	General unemployment rate	Graduate unemployment rate	Non-graduate unemployment rate	Youth unemployment rate	Young graduate unemployment rate	Young non-graduate unemployment rate
Alpha Rate	**r_s=0.414** H_0 not rejected p=0.098	**r_s=-0.463** H_0 not rejected p=0.061	**r_s=-0.355** H_0 not rejected p=0.162	**r_s=-0.365** H_0 not rejected p=0.149	**r_s=-0.478** H_0 not rejected p=0.052	**r_s=-0.571** H_0 rejected p=0.017	**r_s=-0.35** H_0 not rejected p=0.168

Table 5-12 Correlation table, Cluster C

In the case of Cluster C, the Alpha Rate is not linked to GDP per capita (PPP), which is consistent with the observations made at the beginning of Chapter 4 at this level of economic development. The only parameter showing a correlation with the Alpha Rate is the unemployment rate of recent graduates: among these countries where holding a degree is neutral in terms of employment, an increase in the unemployment rate of recent graduates discourages individuals from pursuing higher education, while a decrease encourages them to do so.

Comparison between Cluster B and Cluster C

- A twofold difference is once again observed between the average GDP per capita (PPP) of the two clusters, in favour of Cluster C.
- However, the Alpha Rate in Cluster C increases by 23% compared to Cluster B, exceeding 50%.
- Cluster C shows explosive unemployment rates, particularly in comparison with Cluster B. Indeed, the general unemployment rate quadruples between Cluster B and Cluster C. The unemployment rate of graduates as well as that of recent graduates triples, while that of non-graduates only doubles.
- Compared to the previous comparison between Clusters A and B, we observe a reversal in the trend: in Cluster C, holding a degree appears to be an advantage for employability—except in the case of young people.
- The youth unemployment rate in Cluster C is much higher than in the other observed clusters.

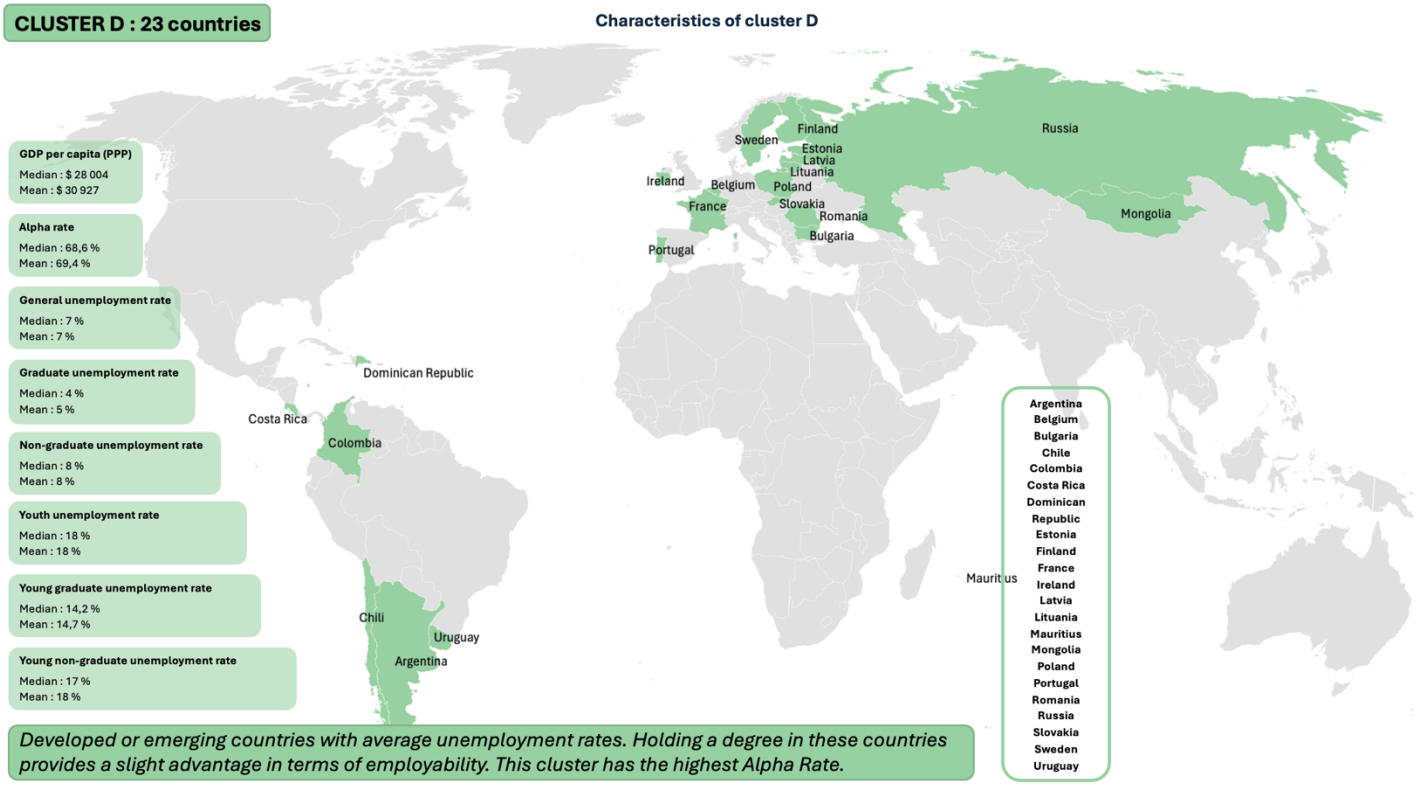


Figure 5-14 World Map, Cluster D

Cluster D: Argentina, Belgium, Bulgaria, Chile, Colombia, Costa Rica, Dominican Republic, Estonia, Finland, France, Ireland, Latvia, Lithuania, Mauritius, Mongolia, Poland, Portugal, Romania, Russia, Slovakia, Sweden, Uruguay.

- The median GDP per capita (PPP) of the countries in this cluster is fairly high ($28K). The mean is $31K (GDP per capita (PPP) ranges from $12.8K (Mongolia) to $74.7K (Iceland)). The median Alpha Rate (68.6%) in Cluster D is the highest compared to those of the other clusters. The mean is 69.4% (ranging from 47% (Mauritius) to 91% (Latvia)).
- The general unemployment rate (7%, 7.2%) and the non-graduate unemployment rate (8.4%, 8.3%) are moderate and close to each other. The youth unemployment rate (18%, 17.8%) and the youth non-graduate unemployment rate (17%, 18.1%) are significant and very similar.
- The graduate unemployment rate (4%, 4.7%) and the recent graduate unemployment rate (14%, 14.7%) are clearly lower than the unemployment rates of non-graduates (8%, 8.3%) and youth non-graduates (17%, 18.1%).
- France belongs to this cluster.

	GDP per capita (PPP)	**General unemployment rate**	**Graduate unemployment rate**	**Non-graduate unemployment rate**	**Youth unemployment rate**	**Young graduate unemployment rate**	**Young non-graduate unemployment rate**
Alpha Rate	**r_s=0.276** H_0 not rejected p=0.203	**r_s=0.11** H_0 not rejected p=0.618	**r_s=-0.296** H_0 not rejected p=0.17	**r_s=0.434** H_0 rejected p=0.039	**r_s=-0.363** H_0 not rejected p=0.089	**r_s=-0.528** H_0 rejected p=0.01	**r_s=-0.27** H_0 not rejected p=0.213

Table 5-13 Correlation table, Cluster D

The following observations can be made regarding the correlations among the variables for the set of countries in cluster D:

- The non-graduate unemployment rate is positively correlated with the Alpha Rate (0.434). This correlation is moderate. In other words, the increase in access to higher education in these countries has a moderate but significant impact on the increase in the non-graduate unemployment rate.
- The correlation between the Alpha Rate and the young graduate unemployment rate is quite strong and negative (-0.528). This suggests that at least one of the following observations is true for these countries:
    - An increase in the young graduate unemployment rate may lead to a decline in student enrolment (a discouraging factor).
    - A decrease in the young graduate unemployment rate may contribute to an increase in university enrolment (an attractive factor).

Among these middle or high-income countries, where the higher education access rate is the highest (average of 69.3%), the absence of a degree particularly harms employment, and this disadvantage increases as access to higher education expands. Conversely, an increase in the young graduate unemployment rate has a deterrent effect on pursuing higher education.

Comparison between Cluster C and Cluster D

- A twofold difference is once again observed between the average GDP per capita (PPP) of the two clusters, in favour of Cluster D.
- The Alpha Rate in Cluster D is 14% higher than that of Cluster C, reaching 69%.
- Unemployment rates in Cluster D are lower than those in Cluster C, particularly the general unemployment rate, which is halved, and the graduate unemployment rate, which is three times lower.
- Cluster D is, however, the first cluster in which non-graduates have higher unemployment rates than graduates, whether young or not. Holding a degree thus appears to be an advantage for the young population (14% for the recent graduate unemployment rate compared to 17% for the youth non-graduate unemployment rate), and even more so for the active population (7% for the general unemployment rate compared to 4% for the graduate unemployment rate).

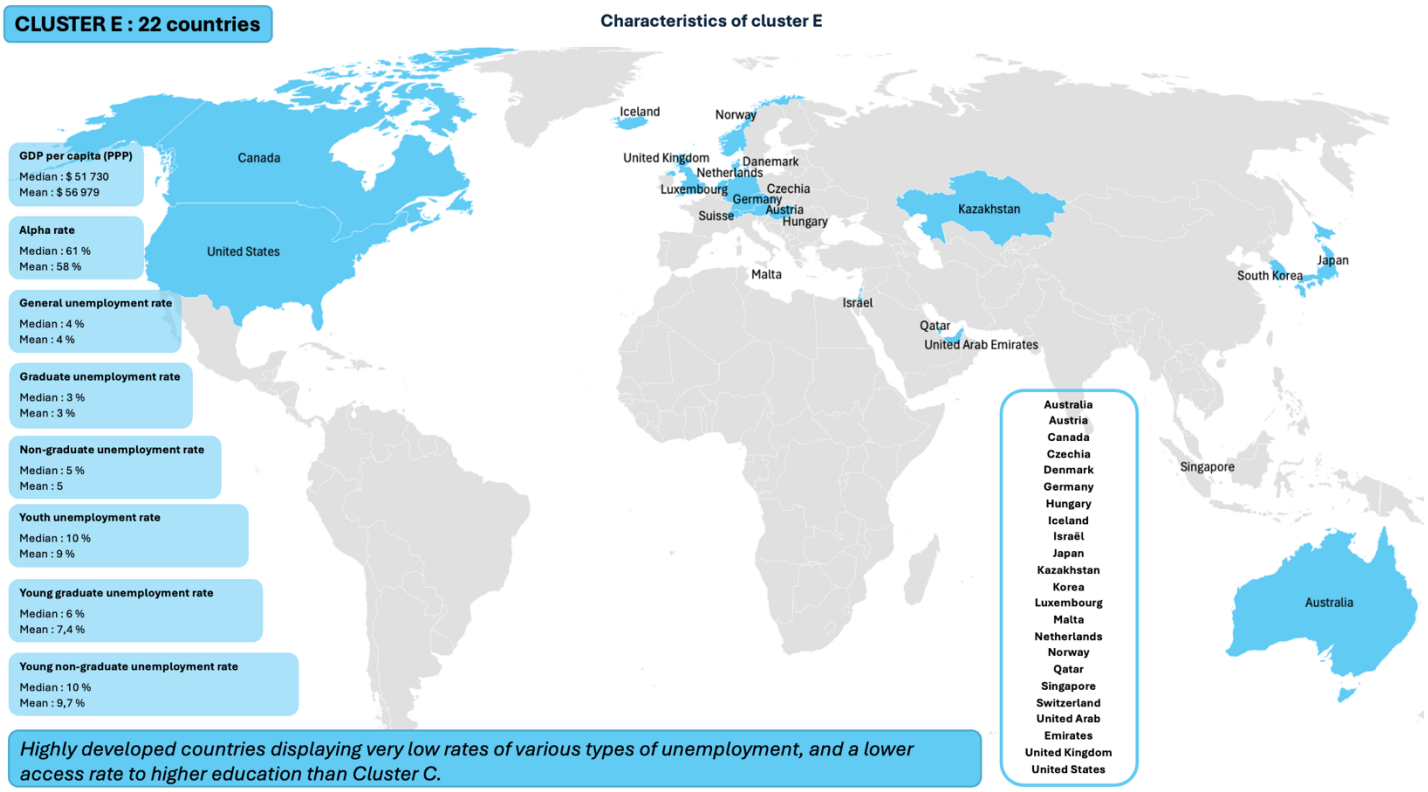


Figure 5-15 World Map, Cluster E

Cluster E: Australia, Austria, Germany, Canada, Czech Republic, South Korea, Denmark, Hungary, Iceland, Israel, Japan, Kazakhstan, Luxembourg, Malta, Norway, Netherlands, Qatar, Singapore, Switzerland, United Arab Emirates, United Kingdom, and United States.

For the countries in Cluster E, it is observed that :

- The median GDP per capita (PPP) is the highest, at $51,730. The mean is $56,979 (GDP per capita (PPP) ranges from $26K (Kazakhstan) to $128K (Qatar).
- The Alpha Rate in Cluster E is high, with a median of 61.2%. The mean is 58% (ranging from 8.8% (Qatar) to 82.4% (Norway)).
- The general unemployment rate (4%, 4.1%), the youth unemployment rate (9.5%, 9.4%), the graduate unemployment rate (3%, 3%), and the recent graduate unemployment rate (6%, 7.5%) are low compared to the other clusters.
- However, the non-graduate unemployment rate (5%, 4.9%) and the youth non-graduate unemployment rate (10%, 9.6%) are high compared to the corresponding rates for graduates.
- It is observed that in this cluster, with the highest level of economic development, unemployment rates are very low, and holding a degree is a positive factor for employability.

	GDP per capita (PPP)	General unemployment rate	Graduate unemployment rate	Non-graduate unemployment rate	Youth unemployment rate	Young graduate unemployment rate	Young non-graduate unemployment rate
Alpha Rate	**r_s=-0.127** H_0 not rejected p=0.573	**r_s=0.142** H_0 not rejected p=0.529	**r_s=-0.04** H_0 not rejected p=0.859	**r_s=0.251** H_0 not rejected p=0.259	**r_s=-0.05** H_0 not rejected p=0.824	**r_s=-0.259** H_0 not rejected p=0.244	**r_s=-0.051** H_0 not rejected p=0.82

Table 5-14 Correlation table, Cluster E

For the countries in Cluster E, we observe that:

- The access rate to higher education is not correlated with GDP per capita (PPP). This indicates that the Alpha Rate does not depend on economic development in this set of countries, and vice versa.
- There is no statistically significant correlation between the Alpha Rate and any of the other unemployment-related variables.

Comparison between Cluster D and Cluster E

- A slightly less than twofold difference is observed between the average GDP per capita (PPP) of the two clusters, in favour of cluster E. Nevertheless, the Alpha Rate in cluster E is more than 7% lower than that of cluster D.
- Unemployment rates in cluster E are lower than those in cluster D, particularly the general unemployment rate, which is halved. In cluster E, a greater gap is observed between the unemployment rate of recent graduates (6%) and that of youth non-graduates (10%).
- Countries in cluster E therefore present a situation in which holding a degree is an even greater advantage than in countries of cluster D, particularly for young people.

CONCLUSION :

In this section, we conducted a clustering method on 89 countries, grouping them into clusters that share similar characteristics in terms of GDP per capita (PPP), Alpha Rate, and various unemployment rates in 2017. After clustering, we performed a correlation analysis between the Alpha Rate and different unemployment rates. The analysis in this chapter suggests that access to higher education exhibits very few correlations with different types of unemployment rates across most clusters.

This approach offers the advantage of clearly identifying behavioral differences across homogeneous clusters. Notably, clusters (A and B) characterized by the lowest GDP per capita (PPP) are the only ones in which a correlation between GDP per capita (PPP) and the Alpha Rate is observed.

The only other significant correlations are observed between the Alpha Rate and the unemployment rate of recent graduates in Cluster C (-0.571) and Cluster D (-0.53), along with a positive correlation between the Alpha Rate and the unemployment rate of non-graduates (0.434) in the case of Cluster D. In both cases, Cluster C and Cluster D represent clusters of countries in which the general unemployment rate or the unemployment rate of non-graduates is higher than that of graduates. This implies an analogy between Cluster C and Cluster D and the intersection of sets X, Y, Z, W previously studied. The absence of correlation between the Alpha Rate and any of the unemployment rates in Cluster D, at a higher level of economic development, indicates that the positive effect of holding a degree exists in wealthy countries only among those with a high or fairly high unemployment rate. While cross-country comparisons should always be approached with caution, this likely reveals a differentiated signalling effect, as the increase in the Alpha Rate in these countries has no impact on general or youth unemployment but gives an advantage to recent graduates and penalises youth non-graduates, thereby reinforcing the polarisation of the labour market.

The map on the following page summarises all of this information relating to the five clusters.

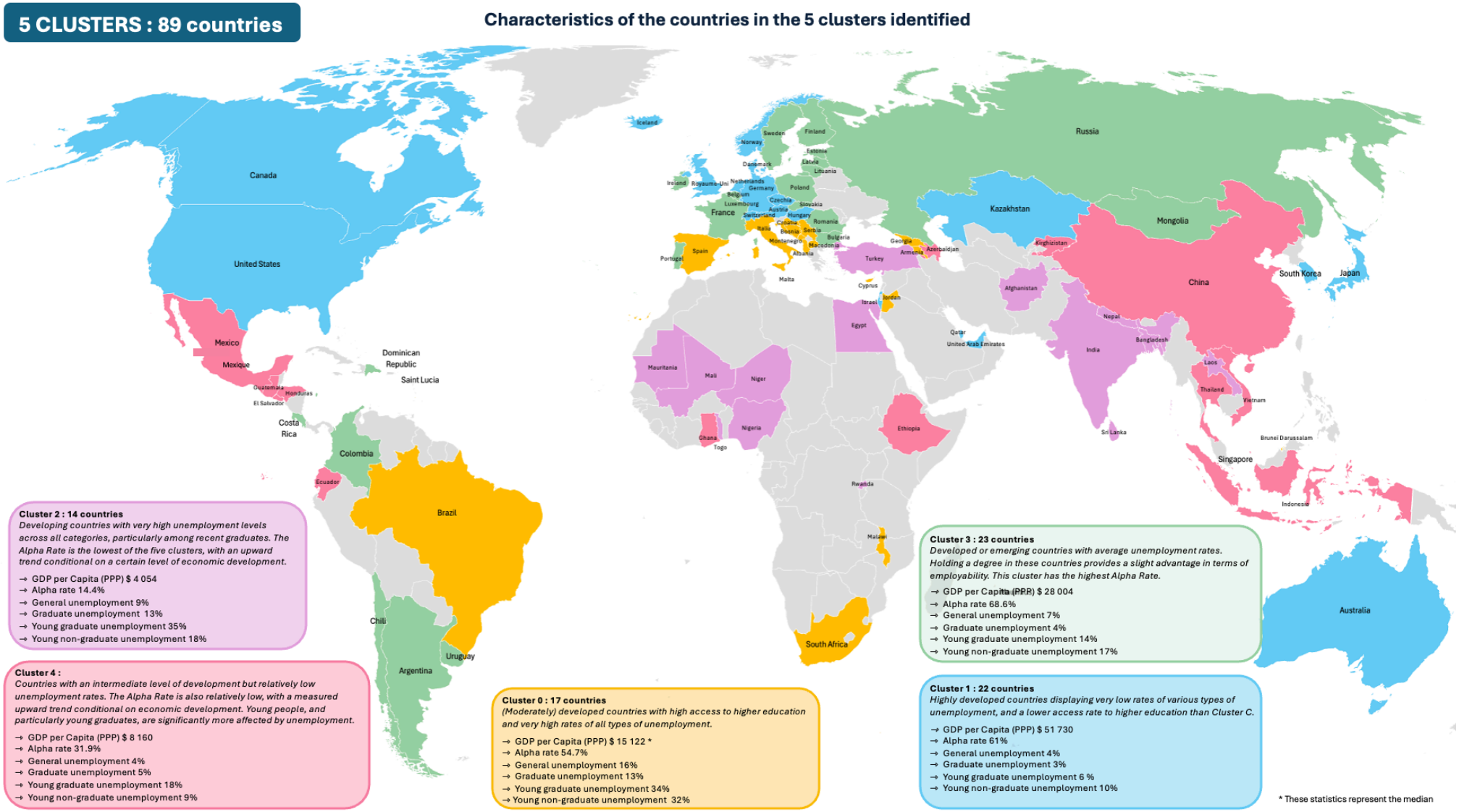


Figure 5-16 World Map: 5 clusters

5.4.3 Evolution over time of unemployment rates in each cluster

In this section, we aimed to observe the evolution over time of unemployment rates across different countries grouped into clusters. Our initial database on unemployment covers the period from 1990 to 2019. However, data are not consistently recorded for each country and each year, particularly for years prior to 2000. Therefore, we restricted our analysis to the period from 2000 to 2017.

As previously discussed, the analysis focuses on 89 countries grouped into five clusters. The clustering was performed based on 2017 data. To track the evolution of a specific unemployment rate within a cluster, we examined the average of this variable within the cluster over a given period. Given the lack of homogeneous data for all countries in a cluster between 2000 and 2017, we selected for each cluster the longest period during which the maximum number of countries could be included in the analysis based on consistent data.

We analyzed general unemployment rate data separately from youth unemployment rate data. In the "Global" category, we considered all countries, regardless of their cluster, over a period that allows for the inclusion of the largest possible number of countries with available unemployment rate data. The following table describes the countries included in the analysis of temporal trends as well as the periods over which the evolutions are observed for the two age categories.

Cluster	Country	Period
Cluster A	Egypt, Mali, Sri Lanka, Turkey (4 out of 14 countries)	2013 - 2017
Cluster B	Ecuador, El Salvador, Honduras, Indonesia, Mexico, Thailand, Vietnam (7 out of 13 countries)	2010 - 2017
Cluster C	Albania, Bosnia and Herzegovina, Croatia, Cyprus, Italy, North Macedonia, Serbia, Spain (8 out of 17 countries)	2009 - 2017
Cluster D	Belgium, Bulgaria, Dominican Republic, Estonia, Finland, France, Ireland, Latvia, Lithuania, Poland, Portugal, Romania, Slovakia, Slovenia, Sweden, Uruguay (16 out of 22 countries)	2000 - 2017
Cluster E	Germany, Australia, Austria, Canada, Czech Republic, Denmark, Hungary, Israel, Luxembourg, Netherlands, Norway, Switzerland, United Kingdom, United States (13 out of 22 countries)	2000 - 2017
Global	Albania, Germany, Australia, Austria, Bosnia and Herzegovina, Belgium, Bulgaria, Canada, Chile, Colombia, Costa Rica, Croatia, Cyprus, Czech Republic, Denmark, Dominican Republic, Ecuador, Egypt, El Salvador, Estonia, Finland, France, Honduras, Hungary, Iceland, Indonesia, Ireland, Israel, Italy, Latvia, Lithuania, Luxembourg, Malta, Mexico, Mongolia, Netherlands, North Macedonia, Norway, Poland, Portugal, Romania, Russia, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Thailand, Turkey, United Kingdom, United States, Uruguay, Vietnam (54 out of 89 countries)	2010-2017

Table 5-15 Countries and periods for general unemployment rate trends

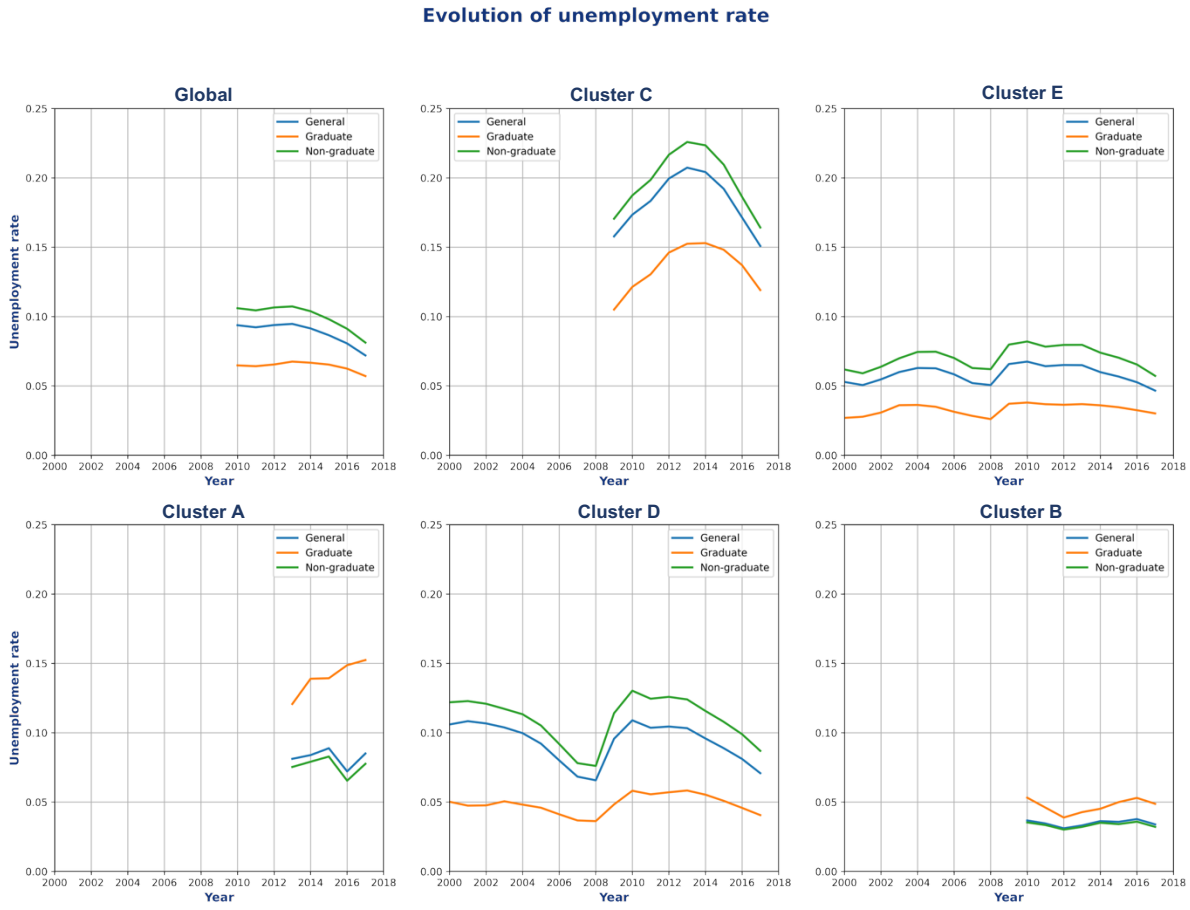


Figure 5-17 Evolution of different types of unemployment rates, all age groups

Figure 5-11 illustrates the evolution of different types of unemployment rates, across all age groups, for the five clusters. As previously mentioned, these trends are observed over different periods due to data limitations. However, we note that the relative positioning of the curves remains consistent over time and that the gaps between the curves of the different unemployment rates across clusters are relatively stable:

- Cluster C, E and D exhibit lower unemployment rates for graduates compared to non-graduates. In contrast, Cluster A and B show higher unemployment rates for graduates than for non-graduates. These relative positions remain stable over time, highlighting the structural nature of these patterns in the respective countries.

When comparing clusters, Cluster C consistently experiences the highest unemployment rates across all categories—general, graduate, and non-graduate unemployment. Conversely, Cluster B exhibits the lowest unemployment rates across all categories, and this trend remains stable over time.

Next, we examine the various types of youth unemployment rates (ages 15-24), specifically the unemployment rates for the youth in general, young graduates, and young non-graduates. As before, we have selected the longest period for which the maximum number of countries can be included in the analysis based on homogeneous data. The following table provides an overview of the countries included in the analysis of youth unemployment rate trends. The number of countries per cluster and the observation periods are more limited compared to general unemployment (graduates and non-graduates), as fewer countries consistently report youth unemployment rates.

Cluster	Country	Period
Cluster A	Egypt, Turkey (2 out of 14 countries)	2009 - 2017
Cluster B	Ecuador, El Salvador, Indonesia, Mexico, Thailand, Vietnam (6 out of 13 countries)	2010 - 2017
Cluster C	Albania, Bosnia and Herzegovina, Croatia, Cyprus, Italy, North Macedonia, Serbia, Spain (8 out of 17 countries)	2009 - 2017
Cluster D	Belgium, Chile, Colombia, Dominican Republic, France, Ireland, Lithuania, Mongolia, Poland, Portugal, Romania, Russia, Slovakia, Slovenia, Sweden, Uruguay (16 out of 22 countries)	2009 - 2017
Cluster E	Germany, Australia, Canada, Czech Republic, Israel, Netherlands, United Kingdom, United States (8 out of 22 countries)	2009 - 2017
Global	Albania, Germany, Australia, Bosnia and Herzegovina, Belgium, Canada, Chile, Colombia, Costa Rica, Croatia, Cyprus, Czech Republic, Dominican Republic, Ecuador, Egypt, El Salvador, France, Indonesia, Ireland, Israel, Italy, Lithuania, Mexico, Mongolia, Netherlands, North Macedonia, Poland, Portugal, Romania, Russia, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Thailand, Turkey, United Kingdom, United States, Uruguay, Vietnam (42 out of 89 countries)	2010 - 2017

Table 5-16 Countries and periods for youth unemployment rate trends

Figure 5-18 Evolution of different types of youth unemployment rates

- The evolution of different types of youth unemployment rates is observed from 2008 to 2017 for all clusters except Cluster B, where data is available from 2010 to 2017. Consequently, the global average is considered for the period 2010 to 2017.
- The global averages (across all countries) for youth unemployment rates of graduates and non-graduates align almost perfectly over the period 2010–2017. These rates tend to decline between 2013 and 2017. However, this overall trend masks significant differences between clusters, underscoring the importance of a cluster-based analysis.
- As observed for general unemployment rates among graduates and non-graduates, Cluster C remains the one with the highest youth unemployment rates. The unemployment rate for young graduates peaked in 2013, while young non-graduates experienced the highest unemployment rates in 2014. Between 2014 and 2017, both curves exhibit a downward trend. The employment trends for graduates and non-graduates reversed in 2013 but remained close over time.
- Cluster E and B exhibit the lowest youth unemployment rates. In Cluster E, graduates are more protected from unemployment than non-graduates, whereas in Cluster B, non-graduates are less affected by unemployment than graduates.
- Cluster A and D have similar levels of general and non-graduate unemployment. However, in Cluster A, graduates are twice as likely to be unemployed as non-graduates, whereas in Cluster D, graduates are slightly better protected from unemployment than non-graduates.

Conclusion on the evolution of unemployment rates over time:

It is striking to observe that within each of the five clusters, the relative positions of graduate and non-graduate unemployment rates remain unchanged over time. These patterns appear to be a constant characteristic of each cluster, which can be summarized as follows (Table 5-17 and Table 5-18).

5.4.4 Synthesis and conclusions

Clusters	Country
A	Afghanistan, Bangladesh, Egypt, India, Laos, Mali, Mauritania, Nepal, Niger, Nigeria, Rwanda, Sri Lanka, Togo, Turkey.
B	Azerbaijan, China, Ecuador, El Salvador, Ethiopia, Ghana, Guatemala, Honduras, Indonesia, Kyrgyzstan, Mexico, Thailand, Vietnam.
C	Albania, Armenia, Bosnia and Herzegovina, Brazil, Brunei Darussalam, Croatia, Cyprus, Spain, Georgia, Italy, Jordan, Malawi, Montenegro, North Macedonia, Saint Lucia, Serbia, South Africa.
D	Argentina, Belgium, Bulgaria, Chile, Colombia, Costa Rica, Dominican Republic, Estonia, Finland, France, Ireland, Latvia, Lithuania, Mauritius, Mongolia, Poland, Portugal, Romania, Russia, Slovakia, Sweden, Uruguay.
E	Australia, Austria, Germany, Canada, Czech Republic, South Korea, Denmark, Hungary, Iceland, Israel, Japan, Kazakhstan, Luxembourg, Malta, Norway, Netherlands, Qatar, Singapore, Switzerland, United Arab Emirates, United Kingdom, United States.

Table 5-17 Clusters and countries

Cluster	GDP per capita (PPP) ($) (median)	Alpha Rate (median, mean)		Young graduate unemployment rate (median, mean)		Young non-graduate unemployment rate (median, mean)		Significant correlations (statistically significant)	General properties
A	4 054	14.4%	18.1 %	35%	41%	18%	18%	(Alpha Rate, GDP per capita (PPP)) = **0.829**	Developing countries with very high unemployment levels across all categories, particularly among recent graduates. The Alpha Rate is the lowest of the five clusters, with an upward trend conditional on a certain level of economic development.
B	8 160	31.9%	33.1%	18%	17%	9 %	8.7%	(Alpha Rate, GDP per capita (PPP)) = **0.621**	Countries with an intermediate level of development but relatively low unemployment rates. The Alpha Rate is also relatively low, with a measured upward trend conditional on economic development. Young people, and particularly young graduates, are significantly more affected by unemployment.
C	15 122	54.3%	53%	34 %	36 %	32%	35%	(Alpha Rate, Young graduate unemployment) **= -0.571**	(Moderately) developed countries with high access to higher education and very high rates of all types of unemployment.
D	28 004	68.6%	69.4%	14%	15%	17%	18%	(Alpha Rate, Non- graduate unemployment) = **0.434** (Alpha Rate, Young graduate unemployment) =**-0.528**	Developed or emerging countries with average unemployment rates. Holding a degree in these countries provides a slight advantage in terms of employability. This cluster has the highest Alpha Rate.
E	51 730	61.2%	58%	6%	7.5%	10%	9.6%	-	Highly developed countries displaying very low rates of various types of unemployment, and a lower access rate to higher education than Cluster D.

Tableau 5-18 Summary of clusters

5.5 Conclusion of chapter 5

Ultimately, regardless of the grouping methods and comparisons of these different countries with one another, the following conclusions emerge:

1. There is, in general, no statistical relationship between higher education access rates and employability, except in two cases: youth graduate employability (negative correlation in cluster C and D) and non-graduate employability (cluster D, which consists of developed countries with high unemployment rates). Furthermore, we have observed multiple possible configurations of unemployment rate positions for each category of countries.

2. The employability of different populations, analyzed through six types of unemployment rates, appears to be a constant that remains stable within each cluster over long or very long periods. These results correspond to enduring socio-economic characteristics that define each cluster.

3. Considering these eight different variables (six unemployment rates, the level of economic development, and the Alpha Rate), countries can be grouped into five distinct categories, within which behaviors are coherent. These countries correspond to very different levels of economic development.

4. It is observed that the wealthiest countries (average GDP per capita (PPP) = $ 57K) with the lowest unemployment rates (cluster E) (average general unemployment rate = 4%) display a significantly lower rate of access to higher education (58%) than those in cluster D (69%), the cluster of fairly wealthy countries to which France belongs (average GDP per capita (PPP) = $ 31K) with higher unemployment rates (average general unemployment rate = 7%). This confirms and refines the findings from the beginning of chapter 4, where it was established that beyond a certain level of economic development, increasing access to higher education has no effect on unemployment.

Given that these characteristics persist over time, these international comparisons highlight how youth employability depends on the ability to integrate workers across all qualification levels, from the least to the most qualified. The challenge also lies in a country's ability to align the level and type of qualifications within its workforce with the needs of its economy.

6 Models and variations

In this part of the analysis, we will examine the possibility of constructing predictive models for different unemployment rates to explore the extent to which variations in the Alpha Rate impact variations in different unemployment rates, when such an impact exists. This section serves as a complement to the main data of our study, which was presented in Chapters 4 and 5; its reading is optional for understanding the overall work.

The previous sections of the analysis indicate monotone relationships between the Alpha Rate and, in some cases, certain unemployment rates. Statistically significant, albeit weak, relationships exist in the 2017 data between the Alpha Rate and the unemployment rate of graduates (r = -0.38) or the youth graduate unemployment rate (r = -0.46). The correlations between the Alpha Rate and other unemployment rates are not statistically significant (general unemployment rate: 0.046, youth unemployment rate: 0.015, non-graduate unemployment rate: 0.135, youth non-graduate unemployment rate: 0.027), meaning that the null hypothesis is not rejected. These observations led us to conclude that, for the 2017 data, variations in the Alpha Rate in 2017 do not impact variations in the general unemployment rate, youth unemployment rate, non-graduate unemployment rate, and youth non-graduate unemployment rate.

One might question whether the fact that all data were compared synchronously (in 2017), despite an average time of four years to obtain a degree, could have introduced a distortion in the observed phenomena. The questions we will now examine are as follows:

Does the variation in the Alpha Rate in 2013 have a significant impact on the youth graduate unemployment rate in 2017, corresponding to the average graduation year of first-cycle students who started in 2013?

NOTE: To be more precise, an increase of p% in the Alpha Rate (Alpha Rate + p%) corresponds to an increase in the number of students in the country by p% of the number of young people aged 18 to 22.

If the answer is affirmative, we will attempt to construct a model to measure the impact of variations in the youth graduate unemployment rate in 2017 as a function of variations in the Alpha Rate in 2013. To achieve this, our 2017 data set will be expanded to include Alpha Rate and graduate unemployment data from 2013, allowing us to observe potential relationships within this new data set.

First, a simple multiple linear regression model will be constructed with the objective of estimating the youth graduate unemployment rate in 2017 based on a set of variables, among which the Alpha Rate in 2013 will necessarily be included.

This model will provide a global picture of the relationship between the youth graduate unemployment rate in 2017 and the Alpha Rate in 2013 across all countries, meaning that no specific information about a particular country will be obtained. To make our analysis of variations more specific in terms of the relationship between the 2013 Alpha Rate and the 2017 youth graduate unemployment rate in a particular country, we will explore the use of ensemblist predictive models (using XGBOOST) and examine their behavior within the previously constructed clusters.

6.1 Choice of variables to estimate the impact of the 2013 Alpha Rate on youth graduate unemployment in 2017

The main objective of this section is to predict variations in the youth graduate unemployment rate in 2017 based on variations in the Alpha Rate in 2013.
We have access to data such as GDP per capita (PPP) (2013, 2015, 2017), the Gini index (2013, 2017), the Alpha Rate (2013, 2017), and six different types of unemployment rates for three years. In total, we have 27 variables. Our model can be presented as follows:

Youth graduate unemployment rate in 2017 = Model (Alpha Rate 2013; other variables).

Given that we will study variations in the youth graduate unemployment rate in 2017 as a function of changes in the 2013 Alpha Rate, the "other variables" must be selected from among the remaining 25 variables.
Annex 8.5 presents the correlations between the 2017 youth graduate unemployment rate in 2017 and all available variables. It shows that many variables are significantly correlated with this unemployment rate.

To construct our model, we must select only those variables that meet both of the following conditions simultaneously:

- They must be correlated with the youth graduate unemployment rate in 2017 (Annex 8.5.1).
- They must not be correlated with the 2013 Alpha Rate (Annex 8.5.2).

This second condition is essential: it allows us to vary the 2013 Alpha Rate without affecting the other variables in the model. As a result, any observed changes in the youth graduate unemployment rate in 2017, following an increase in the 2013 Alpha Rate, will necessarily and exclusively be the direct consequence of variations in the 2013 Alpha Rate.

By cross-referencing the available correlation data in Annex 8.4, we identify a subset of variables that meet both criteria. This subset is referred to as the optimal set.

Ultimately, this optimal set, which satisfies the previously described criteria (all variables in the optimal set are correlated with the youth graduate unemployment rate in 2017, and the 2013 Alpha Rate is not correlated with any other variables in the set), consists of the following variables:

Variables in the optimal set	Correlation between unemployment rate in 2017, unemployment rate of 2013 graduates and Alpha 2013
Alpha 2013	-0.565, p=0.0, H_0 rejected
Graduate unemployment rate 2013	0.636, p=0.0, H_0 rejected
General unemployment rate 2017	0.479, p=0.0, H_0 rejected
Non-graduate unemployment rate 2017	0.337, p=0.001, H_0 rejected
Youth unemployment rate 2017	0.537, p=0.0, H_0 rejected
Young non-graduate unemployment rate 2017	0.475, p=0.0, H_0 rejected

Table 6-1 Optimal set of variables for the models

This optimal set will be our main basis for building models.

6.2 Non-monotonic relationships between variables: Hoeffding's D

Hoeffding's D correlation is a measure of linear, monotonic, and non-monotonic relationships. It takes values between -0.5 and 1. The sign of the Hoeffding's D correlation coefficient has no interpretation.
To interpret this type of relationship between variables, Hoeffding's D is used in conjunction with either Spearman correlation (monotonic relationship) or Pearson correlation (linear relationship). The following rules serve as a basis for interpretation:

- If Pearson correlation is very low and Hoeffding's D correlation is very high, the relationship between the variables is non-linear.
- If Spearman correlation is very low and Hoeffding's D correlation is very high, the relationship between the variables is non-monotonic.
- If both Spearman and Hoeffding's D correlations are low, then the relationship between the variables is completely random.

We used the available Python code (Dev s.d.) to calculate the Hoeffding's D correlation between the variables.

	Alpha	Taux de chômage général	Taux de chômage des non-diplômés	Taux de chômage des jeunes	Taux de chômage des jeunes non-diplômés
Alpha	0.975381	0.002830	0.019044	-0.005382	-0.005278
Taux de chômage général	0.002830	0.975381	0.649263	0.442417	0.457515
Taux de chômage des non-diplômés	0.019044	0.649263	0.975381	0.345923	0.389388
Taux de chômage des jeunes	-0.005382	0.442417	0.345923	0.975381	0.790450
Taux de chômage des jeunes non-diplômés	-0.005278	0.457515	0.389388	0.790450	0.975381

Table 6-2 Hoeffding's D correlations (Higher education access rate and different unemployment rates in 2017)

This table clearly shows that the Hoeffding's D correlations between the Alpha Rate and the other variables are very small, almost insignificant. It should be noted that the Spearman correlations between these variables are also small and statistically insignificant. This leads us to conclude that the relationships between the Alpha Rate and the four different unemployment rates are entirely due to chance, and no non-monotonic relationship exists between these variables.

Based on the 2017 data, we can conclude that an increase or decrease in the Alpha Rate has no impact, whether described by a linear or non-linear, monotonic or non-monotonic relationship, on variations in the four unemployment rates: general unemployment rate, non-graduate unemployment rate, youth unemployment rate, and youth graduate unemployment rate.

6.3 Linear model

The multiple regression model we are constructing aims to predict the youth graduate unemployment rate in 2017 based on the Alpha Rate in 2013. This model must include the Alpha Rate 2013 and variables from the optimal set in table 6.1 (though not necessarily all of them). Since the linear model should not contain variables that are correlated with each other and given that the variables representing different unemployment rates in the optimal set are correlated, the set of variables that can be used as independent variables in this model is reduced to the following: Alpha Rate 2013 and Graduate Unemployment Rate 2013. This set includes variables

that are correlated with the youth graduate unemployment rate in 2017 but not with each other. The following table provides further numerical details:

	Taux de chômage des jeunes diplômés 2017	Alpha 2013	Taux de chômage des diplômés 2013
Taux de chômage des jeunes diplômés 2017	(1.0, p=0.0, Reject H0)	(-0.565, p=0.0, Reject H0)	(0.636, p=0.0, Reject H0)
Alpha 2013	(-0.565, p=0.0, Reject H0)	(1.0, p=0.0, Reject H0)	(-0.176, p=0.104, Fail to reject H0)
Taux de chômage des diplômés 2013	(0.636, p=0.0, Reject H0)	(-0.176, p=0.104, Fail to reject H0)	(1.0, p=0.0, Reject H0)

Table 6-3 Correlations between Alpha Rate 2013 and different unemployment rates in 2017

We obtain the following linear model:

$$Youth\ graduate\ unemployment\ rate\ in\ 2017\ =$$
$$\mathbf{0.23}\ +\ \mathbf{1.66}\ \times\ \boldsymbol{Graduate\ unemployment\ rate}\ \mathbf{2013}\ -\ \mathbf{0.30}\ \times\ \boldsymbol{Alpha\ Rate}\ \mathbf{2013}$$

We evaluated the model's performance and obtained a coefficient of determination (R^2) of 62% across 87 countries (r = 78.8%). This is presented in the following table:

OLS Regression Results

Dep. Variable:	Taux de chômage des jeunes diplômés 2017	**R-squared:**	0.621
Model:	OLS	**Adj. R-squared:**	0.612
Method:	Least Squares	**F-statistic:**	68.81
Date:	Mon, 21 Nov 2022	**Prob (F-statistic):**	2.01e-18
Time:	19:33:59	**Log-Likelihood:**	81.111
No. Observations:	87	**AIC:**	-156.2
Df Residuals:	84	**BIC:**	-148.8
Df Model:	2		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	0.2370	0.028	8.463	0.000	0.181	0.293
Taux de chômage des diplômés 2013	1.6610	0.207	8.025	0.000	1.249	2.073
Alpha 2013	-0.3020	0.043	-7.070	0.000	-0.387	-0.217

Omnibus:	28.384	**Durbin-Watson:**	1.620
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	73.233
Skew:	1.080	**Prob(JB):**	1.25e-16
Kurtosis:	6.941	**Cond. No.**	22.2

Table 6-4 Linear representation between Alpha Rate 2013 and different unemployment rates in 2017

This model indicates that a 1% increase in the Alpha Rate in 2013 leads, on average, to a 0.3% decrease in the youth graduate unemployment rate across different countries in 2017.

This formula is significant for two reasons:

- The coefficient +1.66, the multiplicative factor of the graduate unemployment rate in 2013, represents what could be considered the "handicap" of a young graduate compared to a general graduate in the labor market. This simply means—consistent with well-documented findings—

that young people take longer on average, to find employment than adults, who generally already have work experience and often transition between jobs without experiencing unemployment.

- The -0.30 factor, the multiplicative coefficient of the Alpha Rate 2013, which influences the increase or decrease in the rate of access to higher education, can be explained as follows: four years later, 30% of the additional young people who began their studies in 2013 likely do not enter the labour market immediately—either because they continue their studies or because they wait for a better opportunity to position themselves in the job market.

In both cases, what matters is both the sign of the coefficient and its magnitude (greater than one for the first coefficient, significantly lower for the second). It should be noted that with $R^2 = 0.62$, the correlation is strong and explains 60% of the variations in the data distribution.

6.4 Ensemblist model

In the previous section, we constructed the linear model that estimates variations in the graduate unemployment rate in 2017 based on the Alpha Rate in 2013 and the graduate unemployment rate in 2013 across all countries. According to this model, variations are equal for each country: a 0.3% decrease in the youth graduate unemployment rate for a 1% increase in the Alpha Rate. The ensemblist model that we will use in this chapter aims to estimate these variations separately for each country, which we will then group according to the previously described clusters. To achieve this, we will use the XGBoost-Regression model.

The set of variables presented in section 6.1 will be used to construct the model. Two new variables will be added: GDP per capita (PPP) (2012) and the Gini index (2012). It is important to emphasize that variations in the Alpha Rate of 2013 cannot cause variations in GDP per capita (PPP) in 2012 or the Gini index in 2012. These variables are essential as they reflect a country's economic development.

The data is divided into two sets: TRAIN and TEST, using an 80:20 ratio. After training on the TRAIN data set, the model that performed best on the TEST data set is presented below:

XGBRegressor(base_score=0.5, booster='gbtree', colsample_bylevel=1, colsample_bynode=1, colsample_bytree=0, enable_categorical=False, gamma=0.014586238235480787, gpu_id=-1, importance_type=None, interaction_constraints='', learning_rate=0.3589158743315239, max_delta_step=0, max_depth=7, min_child_weight=3, monotone_constraints='()', n_estimators=100, n_jobs=1, num_parallel_tree=1, predictor='auto', random_state=0, reg_alpha=0.2093846908972189, reg_lambda=0.4056447522900459, scale_pos_weight=1, subsample=1, tree_method='exact', validate_parameters=1, verbosity=None)

This model does not exhibit overfitting, and the coefficient of determination remains relatively high. The model's performance is described in the following table:

Performance Indicators	TRAIN	TEST
R^2	0.87	0.85
Mean square error (MSE)	0.002	0.0034
Root mean square error (RMSE)	0.05	0.058
Mean absolute error (MAE)	0.035	0.04

Table 6-5 Ensemblist model applied to 27 variables (9 over 3 years, § 6.1)

First, it is important to highlight the performance of this model, whose residual error remains very low (4%). We will use this model to predict changes in the youth graduate unemployment rate in 2017 based on increases in the Alpha Rate in 2013. It is worth noting that the Alpha Rate in 2013 is not correlated with the four different unemployment rates used in this model (Annex 8.4) in 2017, nor do variations in the Alpha Rate in 2013 have any impact on these four variables.

Additionally, the Alpha Rate in 2013 is uncorrelated with the graduate unemployment rate in 2013 (Annex 8.4). Although the Alpha Rate in 2013 is correlated with **GDP per capita (PPP)** in 2012 and the Gini index in 2012, its variations have no effect on the variations of these two variables. We can therefore conclude that varying the Alpha Rate in 2013 will not induce changes in the other independent variables of this model.

The variation results are presented for each country, grouped into clusters, assuming that the Alpha Rate has increased by:

- 1%, i.e., New Alpha Rate = Old Alpha Rate (in the country) + 1%
- 5%, i.e., New Alpha Rate = Old Alpha Rate (in the country) + 5%
- 10%, i.e., New Alpha Rate = Old Alpha Rate (in the country) + 10%

It is important to recall that an increase of p% in the Alpha Rate corresponds to an increase in the student population in the given country by p% × number of individuals aged 18 to 22.

The x-axis represents the actual values of the Alpha Rate for 87 countries in 2013.

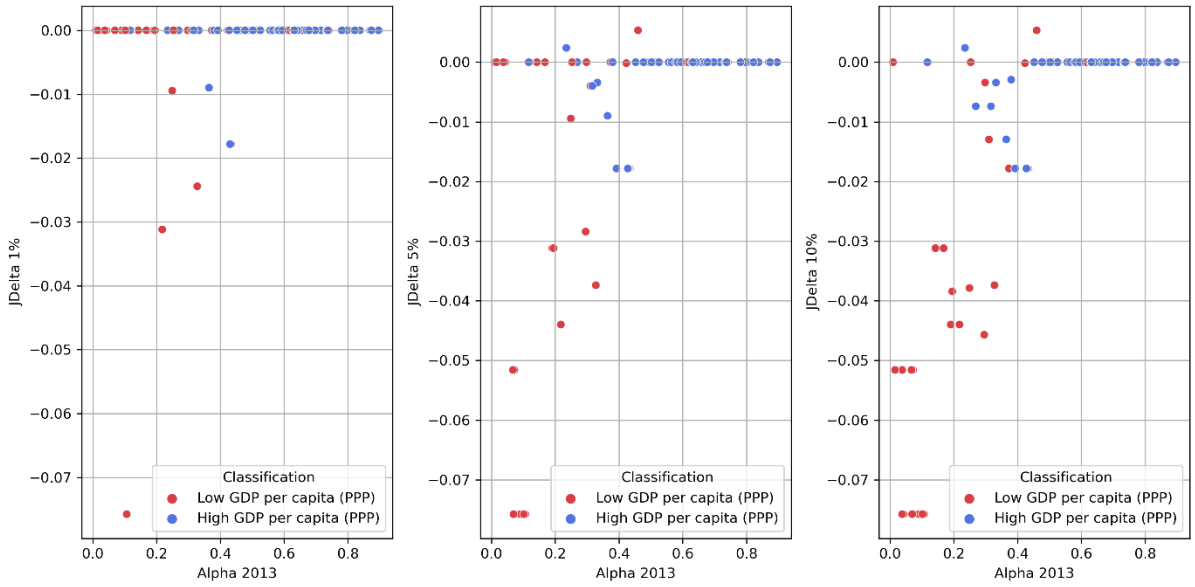


Figure 6-1 Variations in the graduate unemployment rate based on different increases in the Alpha Rate 2013 (1%, 5%, 10%) across all countries

The first figure illustrates that a 1% increase in the Alpha Rate has only a limited impact on the youth graduate unemployment rate in most countries. Except for one country, where a 1% increase in the Alpha Rate results in a nearly 7.5% decrease in youth graduate unemployment, the absolute variations for the remaining countries are below 3%. For countries with an Alpha Rate above 0.5, a 1% increase has no effect on the youth graduate unemployment rate.

The second figure shows that a 5% increase in the Alpha Rate can lead to changes in the youth graduate unemployment rate in more countries than in the previous case. While this variation reaches nearly 7.5% in two countries, the remaining countries experience changes ranging between 1% and 5%.
For countries where the Alpha Rate in 2013 was above 0.5, this 5% increase in the Alpha Rate has no impact on the youth graduate unemployment rate in 2017.

The third figure represents a larger number of countries for which a 10% increase in the Alpha Rate leads to a decrease in the youth graduate unemployment rate.
These countries, which had an Alpha Rate below 0.5 in 2013, exhibit different levels of economic development. For most of them, the absolute variations in the graduate unemployment rate do not exceed 5%; only a limited number of countries show absolute variations between 5% and 7.5%.

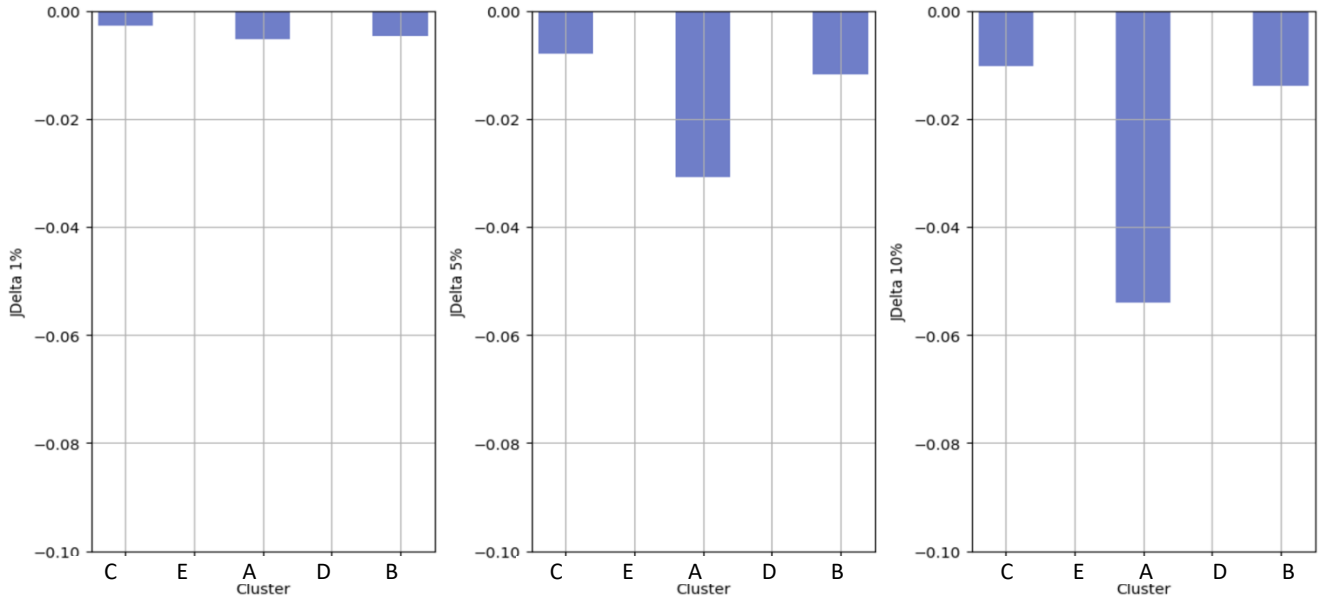


Figure 6-2 Average variations in youth graduate unemployment rate by cluster

- The previous figure illustrates the magnitude of variations within the clusters defined in Chapter IV based on increases in the Alpha Rate. Variations with an average below 1% are observed in cluster C, A and B for a 1% increase in the Alpha Rate.
- A 5% increase in the Alpha Rate leads to variations in the youth graduate unemployment rate with an average below 1% in cluster C and B. Meanwhile, the average variation in cluster A is around 3%.
- A 10% increase in the Alpha Rate introduces variations in the youth graduate unemployment rate in cluster C, A and B. In cluster C, the average of these variations is approximately 1%, while in cluster EB the variations are around 1.5%. Cluster A is by far the most sensitive to changes in the Alpha Rate, with an average variation of 5%.
- In contrast, no variations in youth graduate unemployment rates are observed in cluster E and D, which correspond to wealthy countries (France is in cluster D), when the Alpha Rate varies.

Figure 6-1 highlights the extent of variations in the youth graduate unemployment rate within clusters (defined in Chapter 4) in response to increases of 1%, 5%, and 10% in the Alpha Rate.

These observations are consistent with the previous part of the analysis, where we noted that variations in the youth graduate unemployment rate were only observable in countries with an Alpha Rate below 50% in 2017. Referring to the cluster descriptions provided at the end of Chapter 4; we observed that cluster A had an average Alpha Rate of only 17.3% in 2017. Similarly, cluster B, which exhibits the most significant variations after cluster A, had an average Alpha Rate of 33% in 2017.

Thus, an increase in the Alpha Rate in countries where it is already high (above 50%) will not reduce the youth graduate unemployment rate four years later. However, an increase in the Alpha Rate in countries where this rate is below 50% may lead to a decrease in youth graduate unemployment four years later, as these countries are generally those where the demand for skilled labor is growing.

The following table presents the 35 countries for which at least one of the three increases in the Alpha Rate in 2013 results in variations in the youth graduate unemployment rate in 2017. All of these countries have an Alpha Rate below 46%.

Country (35 countries)	GDP per capita (PPP) 2012	Alpha 2013	Young graduate unemployment rate 2017	Delta YGU for 1 %	Delta YGU for 5 %	Delta YGU for 10 %
Afghanistan	1874	0.04	0.35	0	0	-0.0516
Azerbaijan	17216	0.23	0.1	0	0.0024	0.0024
Bangladesh	3033	0.14	0.36	0	0	-0.0312
Bosnia and Herzegovina	11016	0.43	0.41	-0.0178	-0.0178	-0.0178
Brazil	16111	0.43	0.19	-0.0178	-0.0178	-0.0178
China	12361	0.27	0.19	0	0	-0.0074
Ecuador	11028	0.42	0.13	0	-0.0002	-0.0002
Egypt	10340	0.3	0.54	0	-0.0284	-0.0457
El Salvador	7014	0.3	0.19	0	0	-0.0034
Ethiopia	1367	0.09	0.14	0	-0.0758	-0.0758
Georgia	8694	0.37	0.27	0	0	-0.0178
Honduras	4263	0.22	0.27	-0.0312	-0.044	-0.044
Hungary	24464	0.33	0.07	0	-0.0034	-0.0034
India	5252	0.25	0.36	-0.0095	-0.0095	-0.0379
Indonesia	10009	0.31	0.19	0	-0.004	-0.013
Kyrgyzstan	3232	0.46	0.21	0	0.0053	0.0053
Luxembourg	95590	0.38	0.11	0	0	-0.0029
Madagascar	1414	0.04	0.28	0	0	-0.0758
Malaysia	24279	0.36	0.24	-0.009	-0.009	-0.0129
Mali	1863	0.07	0.84	0	-0.0516	-0.0516
Mexico	17373	0.32	0.13	0	-0.004	-0.0074
Nepal	2333	0.17	0.25	0	0	-0.0312
Niger	902	0.02	0.24	0	0	-0.0516
Nigeria	5698	0.09	0.49	0	-0.0758	-0.0758
North Macedonia	12653	0.39	0.51	0	-0.0178	-0.0178
Pakistan	4398	0.11	0.22	-0.0758	-0.0758	-0.0758

Philippines	6514	0.33	0.18	-0.0244	-0.0374	-0.0374
Rwanda	1651	0.07	0.52	0	-0.0516	-0.0516
Senegal	2903	0.07	0.5	0	-0.0758	-0.0758
Serbia	14629	0.39	0.36	0	-0.0178	-0.0178
South Africa	12815	0.19	0.36	0	-0.0312	-0.044
Sri Lanka	10618	0.19	0.34	0	-0.0312	-0.0384
Tanzania	2539	0.04	0.56	0	0	-0.0758
Togo	1430	0.1	0.3	0	-0.0758	-0.0758
Turkey	22269	0.43	0.34	0	-0.0178	-0.0178

Table 6-6 Variations in youth graduate unemployment rate in 2017 due to changes in the Alpha Rate in 2013, for the 35 countries where Alpha Rate variations in 2013 induce variations in youth graduate unemployment

Reading the table: If the Alpha Rate increases by 1%, 5%, or 10% in Mexico in 2013, the youth graduate unemployment rate in Mexico will decrease by 0%, 0.4%, or 0.74%, respectively, in 2017.

Let us verify whether a relationship exists between the Alpha Rate in 2013 and the different variations (Deltas) in the youth graduate unemployment rate. The following table presents the correlations between these variables. A strong correlation is observed between the Alpha Rate in 2013 and the Delta when the Alpha Rate increases by 10%.

	Delta YGU for 1 %	**Delta YGU for 5 %**	**Delta YGU for 10 %**
Alpha 2013	(**0.101**, p=0.58, H_0 not rejected)	(**0.352**, p=0.048, H_0 rejected)	(**0.786**, p=0.0, H_0 rejected)

The following table presents the simple linear regression model developed to establish the relationship between these two variables **for 35 countries.**

OLS Regression Results

Dep. Variable:	JDelta 10%	**R-squared:**	0.688
Model:	OLS	**Adj. R-squared:**	0.679
Method:	Least Squares	**F-statistic:**	72.76
Date:	Mon, 19 Dec 2022	**Prob (F-statistic):**	7.41e-10
Time:	15:52:08	**Log-Likelihood:**	97.795
No. Observations:	35	**AIC:**	-191.6
Df Residuals:	33	**BIC:**	-188.5
Df Model:	1		
Covariance Type:	nonrobust		

	coef	**std err**	**t**	**P>|t|**	**[0.025**	**0.975]**
const	-0.0717	0.005	-14.068	0.000	-0.082	-0.061
Alpha 2013	0.1571	0.018	8.530	0.000	0.120	0.195

Omnibus:	1.941	**Durbin-Watson:**	1.671
Prob(Omnibus):	0.379	**Jarque-Bera (JB):**	1.602
Skew:	0.371	**Prob(JB):**	0.449
Kurtosis:	2.260	**Cond. No.**	7.56

Table 6-7 Correlation between the decrease in youth graduate unemployment rate in 2017 and the increase in Alpha Rate in 2013, for the 35 concerned countries

The linear regression model presented in the table above indicates that the coefficient of determination is relatively high, at 69.8%. We obtain the following equation:

$$\boldsymbol{Delta\ YGU\ (for\ a\ 10\ \%\ increase\ in\ Alpha)_{2017}} = \text{-}0.0717 + 0.1571 \times Alpha\ Rate\ (real)_{2013}$$

In other words, for the 35 countries with an Alpha Rate below 46%, the reduction in youth graduate unemployment four years after 10% increase in the Alpha Rate is calculated according to the previous formula. In countries where the Alpha Rate is below 46%, the youth graduate unemployment rate will decrease if the Alpha Rate increases by 10%; this decline is more significant when the initial Alpha Rate is lower.

From a sociological perspective, a 10% increase in the Alpha Rate can be perceived as a strong governmental commitment to promoting higher education, thus serving as a signal to employers. It may also be interpreted as a strategy to keep students in universities longer to prevent them from being counted in unemployment statistics.

7 General Conclusion

Our study aimed to examine the relationship between young people's access to higher education and employment opportunities through an international comparative approach encompassing 140 countries. This approach extends beyond the national or local perspectives traditionally used to analyse this issue. To this end, we worked with eight variables: GDP per capita (PPP), access to higher education, and various unemployment rates for each studied country (total, graduates, non-graduates, youth, young graduates, and young non-graduates). Our analysis sought to answer these five research questions:

- How is access to higher education related to economic development?
- Can we establish correlations between access to higher education and different types of unemployment rates? Are these correlations consistent across countries at different stages of economic development?
- How do they evolve across countries at different stages of economic development?
- Can we identify clusters of countries exhibiting similar patterns and explain their behaviors based on shared socio-economic factors?
- Finally, does the strength of correlations, particularly between access to higher education and youth graduate unemployment, enable to develop of predictive models for youth graduate unemployment rates?

Our study revealed, first and foremost, that in countries with a low GDP per capita (PPP) (<$15,000), economic development and access to higher education are correlated. **However, beyond this threshold, economic development and access to higher education become independent variables. Consequently, public policies aiming to expand access to higher education—especially general education—under the assumption that such expansion will systematically improve employment access are not well-founded.** The proportion of young individuals entering the workforce does not increase with a rise in the proportion of graduates, except in limited cases concerning wealthy countries that struggle to provide employment opportunities for less-educated youth. France is among these countries. **Increasing access to higher education enhances the employability of graduates while reducing that of non-graduates, with no significant effect on overall unemployment—likely due to the signaling effect associated with holding a degree.**

Furthermore, increasing higher education access in countries where such access is already high (above 50%) has no measurable impact on reducing the unemployment rate of young graduates after four years of education. We have verified that this observation is not an artifact of clustering effects, whereby antagonistic effects between clusters of countries with homogeneous behaviours internally to clusters could distort the overall data set and bias the analysis. Such is not the case.

Moreover, we demonstrated that in countries where a degree provides a competitive advantage in the labor market, only 20% of the observed reduction in unemployment can be attributed to an increase in higher education attainment. Consequently, over 80% of the decline in unemployment associated with obtaining a diploma is due to other factors, regardless of the level of economic development of the countries studied.

We acknowledge that these findings contradict widely held political narratives, particularly those that emerged following economic crises triggered by oil shocks, famously advocating that "increasing graduation rates will reduce unemployment." We do not suggest that such narratives stem from a deliberate attempt to obscure governments' inability to address unemployment, particularly among young people. Our results merely indicate that these narratives lack empirical support and that their repetition has embedded misconceptions into the collective consciousness. This follows the logic of preconceptions (in Durkheim's sense), as discussed by Esther Duflo[41] in the context of poverty.

We also do not claim that the impact of public policies aimed at expanding access to higher education on employment levels has been overestimated. Our findings simply suggest that, once an "intermediate" level of economic development is reached, reductions in unemployment are predominantly driven by factors other than the access to higher education. It is crucial to distinguish access to higher education from the level of professional skills.

Philippe Aghion, in his extensive research on the influence of innovation on growth[42], shows that a higher degree of innovation is observed alongside greater social mobility, even though social inequalities simultaneously increase. In line with this intuition, it would be interesting to extend our work by considering, beyond higher education access rates, the variations in social mobility rates towards higher education across countries, data that we did not have access to. Furthermore, in relation to these studies that have established the link between growth and innovation, one might be surprised that a higher level of qualifications does not contribute to a higher degree of innovation. A response to this legitimate critique may stem from the fact that our comparisons between countries focus on the TOTAL rate of access to higher education across all fields of academic study. However, it is observed that the distribution of students across disciplines varies considerably from country to country, with some of the same fields accounting for 10% of students in certain countries and 40% in others. While it is likely that there is a link between the disciplines studied (on average) and innovation (often understood in its technological dimension), it is possible that the lack of a relationship between GDP per capita (PPP) and access to higher education, beyond $15,000 per capita (PPP), could be partly due to variations in the distribution of students across disciplines from one country to another, a topic that could extend this research.

The innovative aspect of our study lies primarily in the density of data analyzed, which covered a set of countries representing over 90% of the global population and youth, for which we individually processed higher education access rates for each country.

Our findings suggest that beyond the level of education, it is the relevance of skills to the labour market and the country's stage of economic development that primarily determine employability.

[41] Esther Duflo, "La pauvreté est multidimensionnelle", 2022, Collège de France, https://www.college-de-france.fr/fr/actualites/la-pauvrete-est-multidimensionnelle

[42] Philippe Aghion, "Innovation and Growth from a Schumpeterian Perspective", 2018, https://www.jstor.org/stable/26596237; Philippe Aghion, Gilbert Cette, Élie Cohen & Jean Pisani-Ferry, "Les leviers de la croissance française", p185, https://www.cae-eco.fr/staticfiles/pdf/072.pdf

8 Appendix

8.1 Graduate and young graduate unemployment rates (by genre) in countries where a degree does not facilitate employement access

Country	Graduate unemployment rate (Female)	Graduate unemployment rate (Male)	Delta abs(F-M)	Delta rel(F/M)
Afghanistan	0.306	0.122	0.184	2.512
Albania	0.140	0.133	0.007	1.054
Armenia	0.179	0.178	0.002	1.009
Bangladesh	0.211	0.080	0.130	2.630
Chile	0.074	0.068	0.006	1.091
Colombia	0.125	0.091	0.034	1.370
Ecuador	0.069	0.046	0.023	1.492
Egypt	0.311	0.147	0.164	2.113
El Salvador	0.059	0.050	0.009	1.178
Georgia	0.148	0.164	0.016	0.902
Ghana	0.051	0.051	0.000	0.996
Guatemala	0.042	0.022	0.020	1.918
Honduras	0.111	0.047	0.064	2.345
Indonesia	0.046	0.041	0.004	1.107
Jordan	0.339	0.170	0.170	2.000
Korea	0.044	0.042	0.002	1.048
Mali	0.403	0.161	0.242	2.501
Mauritania	0.260	0.137	0.122	1.893
Mauritius	0.086	0.062	0.024	1.386
Mexico	0.045	0.042	0.004	1.087
Mongolia	0.071	0.069	0.002	1.022
Qatar	0.011	0.001	0.010	10.156
Rwanda	0.240	0.165	0.075	1.452
Sri Lanka	0.101	0.055	0.046	1.835
Thailand	0.017	0.018	0.000	0.980
Togo	0.247	0.126	0.121	1.960
United Arab Emirates	0.095	0.019	0.076	4.986
Vietnam	0.045	0.035	0.009	1.264

Country	Unemployment rate among young graduates (Female)	Unemployment rate among young graduates (Male)	Delta abs (F-M)	Delta rel (F/M)
Afghanistan	0.470	0.300	0.170	1.566
Albania	0.313	0.356	0.044	0.877
Armenia	0.411	0.340	0.071	1.208
Bangladesh	0.395	0.336	0.059	1.174
Chile	0.212	0.160	0.052	1.326
Colombia	0.231	0.186	0.044	1.239
Croatia	0.357	0.300	0.057	1.190
Ecuador	0.142	0.121	0.021	1.174
Egypt	0.558	0.513	0.045	1.087
Indonesia	0.179	0.204	0.025	0.876
Israel	0.091	0.037	0.054	2.488
Jordan	0.618	0.506	0.112	1.221
Kyrgyzstan	0.272	0.165	0.107	1.646
Lao People's Democratic Republic	0.223	0.367	0.143	0.609
Latvia	0.169	0.230	0.061	0.736
Mauritius	0.261	0.269	0.008	0.969
Mexico	0.131	0.122	0.009	1.076
Mongolia	0.296	0.184	0.111	1.604
Nepal	0.210	0.280	0.071	0.748
North Macedonia	0.516	0.486	0.030	1.062
Qatar	0.034	0.005	0.029	6.627
Russia	0.188	0.173	0.015	1.085
Rwanda	0.512	0.519	0.007	0.987
Serbia	0.373	0.332	0.042	1.126
Singapore	0.216	0.109	0.107	1.982
Slovakia	0.311	0.223	0.088	1.395
Sri Lanka	0.360	0.303	0.057	1.187
Thailand	0.146	0.173	0.027	0.845
United Arab Emirates	0.233	0.085	0.148	2.736
Vietnam	0.176	0.179	0.003	0.983

8.2 List of countries (Chapter 4)

8.2.1 GDP per Capita (PPP) and Alpha Rate 2017 (Chapter 4.2.1)

140 Countries: Afghanistan, Albania, Algeria, Angola, Argentina, Armenia, Aruba, Australia, Austria, Azerbaijan, Bahrain, Bangladesh, Belarus, Belgium, Belize, Benin, Bosnia and Herzegovina, Botswana, Brazil, Brunei Darussalam, Bulgaria, Burkina Faso, Burundi, Cabo Verde, Cambodia, Cameroon, Canada, Chad, Chile, China, Colombia, Congo, Costa Rica, Cote d'Ivoire, Croatia, Cuba, Cyprus, Czechia, Denmark, Dominican Republic, Ecuador, Egypt, El Salvador, Eritrea, Estonia, Ethiopia, Finland, France, Georgia, Germany, Ghana, Grenada, Guatemala, Guinea, Honduras, Hungary, Iceland, India, Indonesia, Iran, Ireland, Israel, Italy, Jamaica, Japan, Jordan, Kazakhstan, Kenya, Korea, Kuwait, Kyrgyzstan, Lao People's Democratic Republic, Latvia, Lebanon, Lesotho, Lithuania, Luxembourg, Madagascar, Malawi, Malaysia, Maldives, Mali, Malta, Mauritania, Mauritius, Mexico, Mongolia, Montenegro, Morocco, Mozambique, Namibia, Nepal,

Netherlands, New Zealand, Niger, Nigeria, North Macedonia, Norway, Oman, Pakistan, Panama, Peru, Philippines, Poland, Portugal, Qatar, Republic of Moldova, Romania, Russia, Rwanda, Saint Lucia, Saudi Arabia, Senegal, Serbia, Seychelles, Singapore, Slovakia, Slovenia, South Africa, Spain, Sri Lanka, Sudan, Sweden, Switzerland, Syrian Arab Republic, Tajikistan, Tanzania, Thailand, Togo, Tunisia, Turkey, Uganda, Ukraine, United Arab Emirates, United Kingdom, United States, Uruguay, Uzbekistan, Vietnam, Zimbabwe.

8.2.2 General unemployment rate and Alpha Rate 2017 (Chapter 4.5.1)

91 Countries: Afghanistan, Albania, Argentina, Armenia, Australia, Austria, Azerbaijan, Bangladesh, Belgium, Bosnia and Herzegovina, Brazil, Brunei Darussalam, Bulgaria, Cambodia, Canada, Chile, China, Colombia, Costa Rica, Croatia, Cyprus, Czechia, Denmark, Dominican Republic, Ecuador, Egypt, El Salvador, Estonia, Ethiopia, Finland, France, Georgia, Germany, Ghana, Guatemala, Honduras, Hungary, Iceland, India, Indonesia, Ireland, Israel, Italy, Japan, Jordan, Kazakhstan, Korea, Kyrgyzstan, Lao People's Democratic Republic, Latvia, Lithuania, Luxembourg, Malawi, Mali, Malta, Mauritania, Mauritius, Mexico, Mongolia, Montenegro, Nepal, Netherlands, Niger, Nigeria, North Macedonia, Norway, Poland, Portugal, Qatar, Romania, Russia, Rwanda, Saint Lucia, Serbia, Singapore, Slovakia, Slovenia, South Africa, Spain, Sri Lanka, Sweden, Switzerland, Thailand, Togo, Turkey, Uganda, United Arab Emirates, United Kingdom, United States, Uruguay, Vietnam.

8.2.3 Youth unemployment rate and Alpha Rate 2017 (Chapter 4.5.2)

92 Countries: Afghanistan, Albania, Argentina, Armenia, Australia, Austria, Azerbaijan, Bangladesh, Belgium, Bosnia and Herzegovina, Brazil, Brunei Darussalam, Bulgaria, Cambodia, Canada, Chile, China, Colombia, Costa Rica, Croatia, Cyprus, Czechia, Denmark, Dominican Republic, Ecuador, Egypt, El Salvador, Estonia, Ethiopia, Finland, France, Georgia, Germany, Ghana, Guatemala, Honduras, Hungary, Iceland, India, Indonesia, Ireland, Israel, Italy, Japan, Jordan, Kazakhstan, Korea, Kyrgyzstan, Lao People's Democratic Republic, Latvia, Lithuania, Luxembourg, Malawi, Mali, Malta, Mauritania, Mauritius, Mexico, Mongolia, Montenegro, Nepal, Netherlands, New Zealand, Niger, Nigeria, North Macedonia, Norway, Poland, Portugal, Qatar, Romania, Russia, Rwanda, Saint Lucia, Serbia, Singapore, Slovakia, Slovenia, South Africa, Spain, Sri Lanka, Sweden, Switzerland, Thailand, Togo, Turkey, Uganda, United Arab Emirates, United Kingdom, United States, Uruguay, Vietnam.

8.2.4 Graduate unemployment rate and Alpha Rate 2017 (Chapter 4.5.3)

88 Countries: Afghanistan, Albania, Argentina, Armenia, Australia, Austria, Azerbaijan, Bangladesh, Belgium, Bosnia and Herzegovina, Brazil, Brunei Darussalam, Bulgaria, Canada, Chile, Colombia, Costa Rica, Croatia, Cyprus, Czechia, Denmark, Dominican Republic, Ecuador, Egypt, El Salvador, Estonia, Ethiopia, Finland, France, Georgia, Germany, Ghana, Guatemala, Honduras, Hungary, Iceland, India, Indonesia, Ireland, Israel, Italy, Japan, Jordan, Kazakhstan, Korea, Kyrgyzstan, Lao People's Democratic Republic, Latvia, Lithuania, Luxembourg, Malawi, Mali, Malta, Mauritania, Mauritius, Mexico, Mongolia, Montenegro, Nepal, Netherlands, Niger, Nigeria, North Macedonia, Norway, Poland, Portugal, Qatar, Romania, Russia, Rwanda, Saint Lucia, Serbia, Singapore, Slovakia, Slovenia, South Africa, Spain, Sri Lanka, Sweden, Switzerland, Thailand, Togo, Turkey, United Arab Emirates, United Kingdom, United States, Uruguay, Vietnam.

8.2.5 Young graduate unemployment rate and Alpha Rate 2017 (Chapter 4.5.4)

82 Countries: Afghanistan, Albania, Argentina, Armenia, Australia, Austria, Azerbaijan, Bangladesh, Belgium, Bosnia and Herzegovina, Brazil, Brunei Darussalam, Bulgaria, Canada, Chile, China, Colombia, Costa Rica, Croatia, Cyprus, Czechia, Denmark, Dominican Republic, Ecuador, Egypt, El Salvador, Estonia, Finland, France, Georgia, Germany, Ghana, Guatemala, Honduras, Hungary, Indonesia, Ireland, Israel, Italy, Japan, Jordan, Kazakhstan, Korea, Kyrgyzstan, Lao People's Democratic Republic, Latvia, Lithuania, Luxembourg, Malawi, Mali, Mauritius, Mexico, Mongolia, Montenegro, Nepal, Netherlands, New Zealand, North Macedonia, Poland, Portugal, Qatar, Romania, Russia, Rwanda, Saint Lucia, Serbia, Singapore, Slovakia, Slovenia, South Africa, Spain, Sri Lanka, Sweden, Switzerland, Thailand, Togo, Turkey, United Arab Emirates, United Kingdom, United States, Uruguay, Vietnam.

8.2.6 Non-graduate unemployment rate and Alpha Rate 2017 (Chapter 4.5.5)

90 Countries: Afghanistan, Albania, Argentina, Armenia, Australia, Austria, Azerbaijan, Bangladesh, Belgium, Bosnia and Herzegovina, Brazil, Brunei Darussalam, Bulgaria, Cambodia, Canada, Chile, Colombia, Costa Rica, Croatia, Cyprus, Czechia, Denmark, Dominican Republic, Ecuador, Egypt, El Salvador, Estonia, Ethiopia, Finland, France, Georgia, Germany, Ghana, Guatemala, Honduras, Hungary, Iceland, India, Indonesia, Ireland, Israel, Italy, Japan, Jordan, Kazakhstan, Korea, Kyrgyzstan, Lao People's Democratic Republic, Latvia, Lithuania, Luxembourg, Malawi, Mali, Malta, Mauritania, Mauritius, Mexico, Mongolia, Montenegro, Nepal, Netherlands, Niger, Nigeria, North Macedonia, Norway, Poland, Portugal, Qatar, Romania, Russia, Rwanda, Saint Lucia, Serbia, Singapore, Slovakia, Slovenia, South Africa, Spain, Sri Lanka, Sweden, Switzerland, Thailand, Togo, Turkey, Uganda, United Arab Emirates, United Kingdom, United States, Uruguay, Vietnam.

8.2.7 Young non-graduate unemployment rate and Alpha Rate 2017 (Chapter 4.5.6)

88 Countries: Afghanistan, Albania, Argentina, Armenia, Australia, Austria, Azerbaijan, Bangladesh, Belgium, Bosnia and Herzegovina, Brazil, Brunei Darussalam, Bulgaria, Cambodia, Canada, Chile, Colombia, Costa Rica, Croatia, Cyprus, Czechia, Denmark, Dominican Republic, Ecuador, Egypt, El Salvador, Estonia, Finland, France, Georgia, Germany, Ghana, Guatemala, Honduras, Hungary, Iceland, Indonesia, Ireland, Israel, Italy, Japan, Jordan, Kazakhstan, Korea, Kyrgyzstan, Lao People's Democratic Republic, Latvia, Lithuania, Luxembourg, Malawi, Mali, Malta, Mauritania, Mauritius, Mexico, Mongolia, Montenegro, Nepal, Netherlands, New Zealand, Niger, North Macedonia, Norway, Poland, Portugal, Qatar, Romania, Russia, Rwanda, Saint Lucia, Serbia, Singapore, Slovakia, Slovenia, South Africa, Spain, Sri Lanka, Sweden, Switzerland, Thailand, Togo, Turkey, Uganda, United Arab Emirates, United Kingdom, United States, Uruguay, Vietnam.

8.3 Groups A, B, C (Chapter 5.1)

Group A consists of countries for which all variables are available for the year 2017 (GDP per capita (PPP), Alpha Rate, general unemployment rate, unemployment rate of graduates, unemployment rate of non-graduates, youth unemployment rate, youth unemployment rate of non-graduates), except for the youth unemployment rate of graduates. To train our model, we will use data from the Complete 2017 Database (80 countries) for the year 2017.

The data of 80 countries will be randomly split into two groups: 90% (72 countries) and 10% (8 countries) of the group A, ensuring that both groups maintain similar distributions across all variables. The first group will serve as the TRAIN SET, while the second group will be designated as the TEST SET. Our model will be trained using the TRAIN SET. During this training phase, we will apply the Cross-Validation method (cv=3, 3-fold) to the TRAIN SET to ensure that the model captures the most general hidden patterns within the data while mitigating the risk of the overfitting effect (i.e., a scenario where the model performs well on the seen data in TRAIN SET but poorly on unseen data from the TEST SET).

Cross-Validation is a technique that partitions the TRAIN SET into N subsets (N=3 in our case) of equal size if the TRAIN SET sample size is divisible by N. Otherwise, it creates N-1 subsets of equal size and one subset containing the remainder of the division. The model is then trained iteratively on each subset, adjusting its parameters after each iteration. This process enhances the model's performance by preventing overfitting.

After a hyperparameter tuning session, we obtained the following model (M1):

XGBRegressor(base_score=0.5, booster='gbtree', colsample_bylevel=1, colsample_bynode=1, colsample_bytree=0.3, gamma=0, gpu_id=-1, importance_type='gain', interaction_constraints=' ', learning_rate=0.30, max_delta_step=0, max_depth=2, min_child_weight=6, monotone_constraints='()', n_estimators=30, n_jobs=12, num_parallel_tree=1, objective='reg:squarederror', random_state=0, reg_alpha=0, reg_lambda=1, scale_pos_weight=1, subsample=1, tree_method='exact', validate_parameters=1, verbosity=None).

The performance of the regression model described above is shown in the following table:

Performance indicators	TRAIN	TEST
R^2	0.952	0.943
Mean square error (MSE)	0.001	0.0012
Root mean square error (RMSE)	0.033	0.034
Mean absolute error (MAE)	0.021	0.024

Table 8-1 Model M1 performance indicators

The following figure demonstrates that our model does not suffer from overfitting and that its performance gradually improves on both seen data (Train Set) and unseen data (Validation Set).

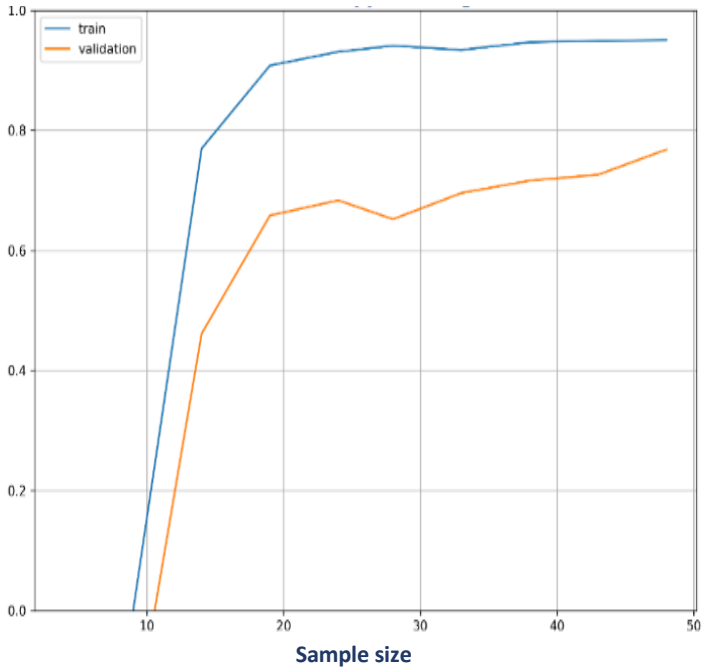


Figure 8-1 Learning curve of model M1

This model provides the following predictions for the youth graduate unemployment rate in 2017 for countries in Group A:

Country	Young graduate unemployment rate
Iceland	0.0583
Malta	0.0777
Mauritania	0.5683
Niger	0.2404
Norway	0.0588

Table 8-2 Predictions by Model M1 for Group A

Group B consists of countries for which the Youth Graduate Unemployment Rate and Youth Non-Graduate Unemployment Rate were unknown in 2017. Once again, we apply the XGBRegressor to estimate the missing values.

The process follows these steps:

- First, we created the model (M2), where the independent variables are: GDP per capita (PPP), Alpha Rate, general unemployment rate, graduate unemployment rate, non-graduate unemployment rate, and youth unemployment rate. The youth graduate unemployment rate was chosen as the target variable.
- Then, we created the model (M3), where the independent variables are: GDP per capita (PPP), Alpha Rate, general unemployment rate, graduate unemployment rate, non-graduate unemployment rate, youth unemployment rate, and youth graduate unemployment rate. The youth non-graduate unemployment rate was chosen as the target variable.

Unlike model (M1), these two models leverage an expanded data set to train. Instead of relying solely on the Complete 2017 Database, we incorporate data from the Complete 2013 Database and the Complete 2015

Database. This expansion is feasible because Alpha Rate data is available for 2013 and 2015. However, not all countries in the Complete 2017 Database are present in the 2013 and 2015 data sets for unemployment rates. By integrating these data sets, we obtain a comprehensive data set with 221 fully populated rows, referred to as the Complete Database. This data set enhances the predictive performance of Models M2 and M3.

The TRAIN to TEST ratio remains 90% vs. 10%. For Cross-Validation (N=3), we trained the model three times. After hyperparameter tuning, we defined as follows the Model (M2):

XGBRegressor (base_score=0.5, booster='gbtree', colsample_bylevel=1, colsample_bynode=1, colsample_bytree=0.5, gamma=0, gpu_id=-1, importance_type='gain', interaction_constraints=' ', learning_rate=0.2, max_delta_step=0, max_depth=5, min_child_weight=6, missing=nan, monotone_constraints='()', n_estimators=20, n_jobs=12, num_parallel_tree=1, objective='reg:squarederror', random_state=0, reg_alpha=0, reg_lambda=1, scale_pos_weight=1, subsample=1, tree_method='exact', validate_parameters=1, verbosity=None**)**

The performance of the regression model (M2) is described in the following table:

Performance indicators	TRAIN	TEST
R^2	0.941	0.913
Mean square error (MSE)	0.0012	0.0015
Root mean square error (RMSE)	0.036	0.039
Mean absolute error (MAE)	0.025	0.029

Table 8-3 Model M2 performance indicators

The model learning curve figure shows that the model learns on both training and validation data. This is shown in the figure below:

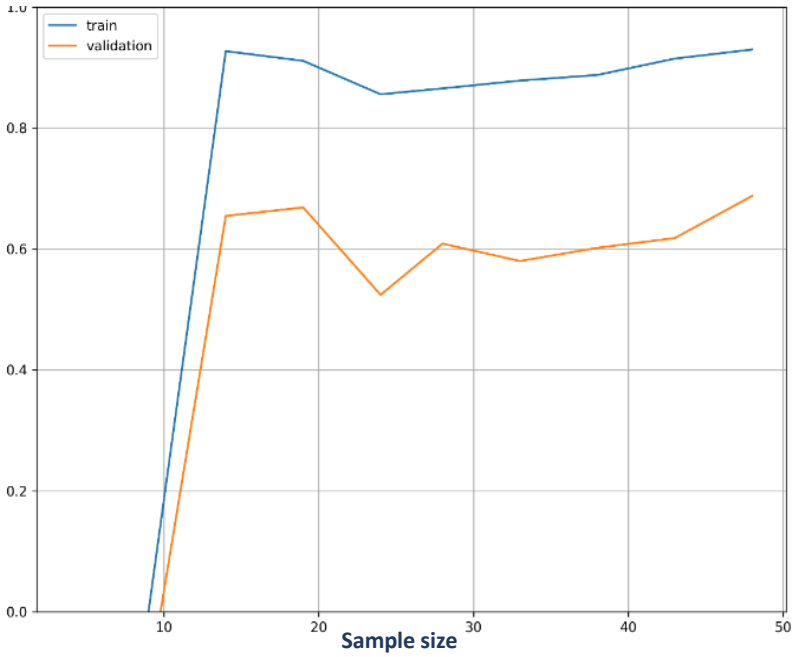


Figure 8-2 Learning curve of model M2

For model (M3), we applied the same TRAIN-TEST ratio (90% vs. 10%) and used Cross-Validation (N=3). After hyperparameter tuning, model **(M3)** is defined as follows:

XGBRegressor(base_score=0.5, booster='gbtree', colsample_bylevel=1, colsample_bynode=1, colsample_bytree=0.8, gamma=0, gpu_id=-1, importance_type='gain', interaction_constraints=' ', learning_rate=0.1, max_delta_step=0, max_depth=5, min_child_weight=1, missing=nan, monotone_constraints='()', n_estimators=100, n_jobs=12, num_parallel_tree=1, objective='reg:squarederror', random_state=0, reg_alpha=0, reg_lambda=1, scale_pos_weight=1, subsample=1, tree_method='exact', validate_parameters=1, verbosity=None **)**

The performance of regression model (M3) is summarized in the following table:

Performance indicators	TRAIN	TEST
R^2	0.999	0.989
Mean square error (MSE)	0.00001	0.0001
Root mean square error (RMSE)	0.0012	0.014
Mean absolute error (MAE)	0.001	0.01

Table 8-4 Performance indicators of Model M3

Below is the learning curve of the model (M3):

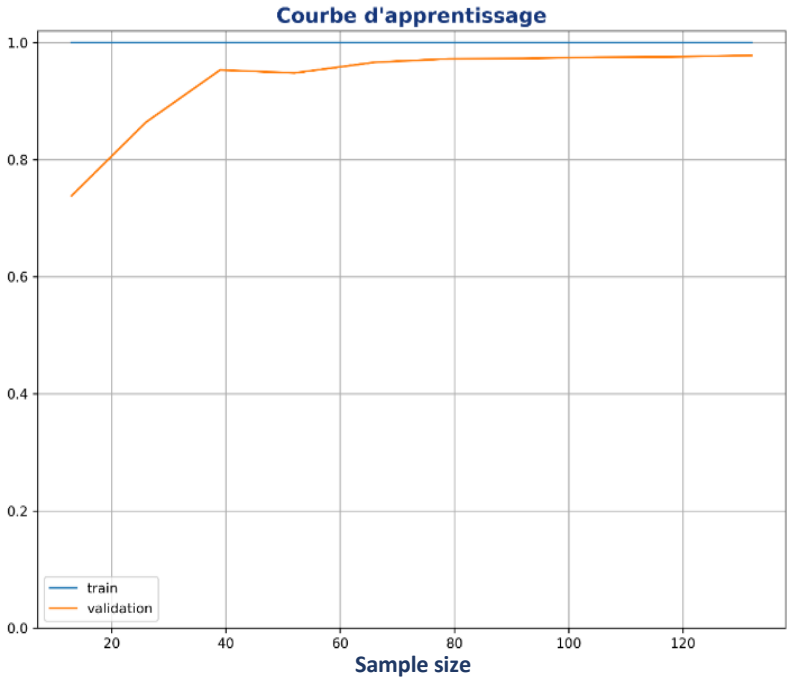


Figure 8-3 Learning curve of model M3

By applying models (M2) and (M3), we estimated the missing values for countries in Group B:

Country	Unemployment rate among young graduates	Unemployment rate among young non-graduates
Ethiopia	0.137	0.034
India	0.358	0.203
Nigeria	0.492	0.166

Table 8-5 Predictions by models M2 and M3 for Group B

Group C includes China. This is a particular case because, for this country, the graduate unemployment rate, the non-graduate unemployment rate, and the youth non-graduate unemployment rate are unknown. This

means that we need to construct two different models to estimate the missing values for the graduate unemployment rate and the non-graduate unemployment rate. To predict the third missing value, we will use model (M3) previously built.

As in the case of Group B, to train our two models, we will use the Complete Database, which gathers data from the years 2013, 2015, and 2017. The same TRAIN-TEST ratio, 90% versus 10%, will be used for these two training processes. For each of the two models, a CROSS-VALIDATION technique with N = 5, which has been found to give the best result, will be applied.

The model predicting the graduate unemployment rate, model **(M4),** takes as independent variables: GDP per capita (PPP), Alpha Rate, general unemployment rate, youth unemployment rate, and youth graduate unemployment rate. **Below is the concerned model:**

XGBRegressor(base_score=0.5, booster='gbtree', colsample_bylevel=1, colsample_bynode=1, colsample_bytree=0.8, gamma=0, gpu_id=-1, importance_type='gain', interaction_constraints='', learning_rate=0.1, max_delta_step=0, max_depth=5, min_child_weight=12, monotone_constraints='()', n_estimators=100, n_jobs=12, num_parallel_tree=1, objective='reg:squarederror', random_state=0, reg_alpha=0, reg_lambda=1, scale_pos_weight=1, subsample=1, tree_method='exact', validate_parameters=1, verbosity=None **)**

The performance of the regression model (M4) is described in the following table:

Performance indicators	TRAIN	TEST
R^2	0.968	0.884
Mean square error (MSE)	0.0012	0.0015
Root mean square error (RMSE)	0.00009	0.0002
Mean absolute error (MAE)	0.006	0.011

Table 8-6 Performance indicators of model M4

The learning curve of the model (figure 8-4) shows that it learns well on both the training data and the validation data.

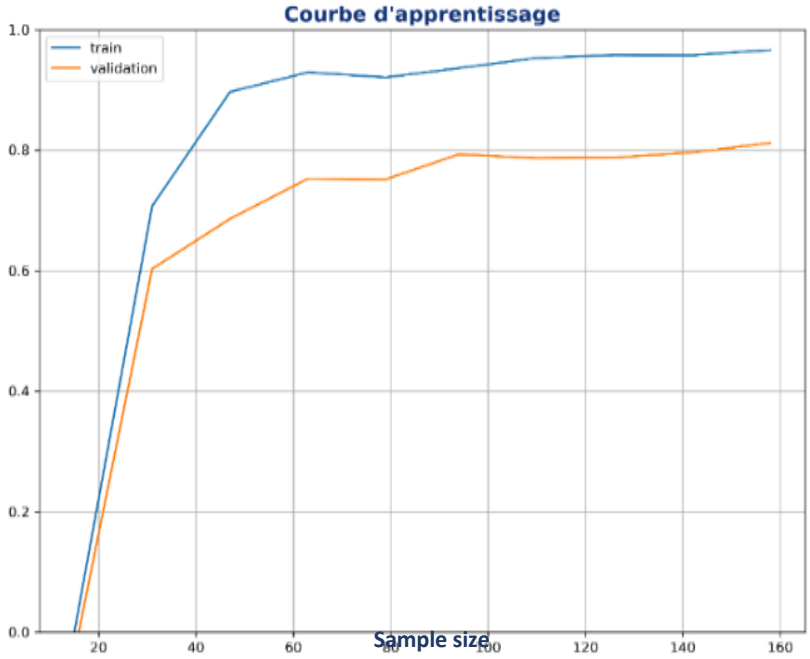


Figure 8-4 Learning curve of model M4

Model (M5), which predicts the non-graduate unemployment rate, uses the following independent variables: GDP per capita (PPP), Alpha Rate, general unemployment rate, youth unemployment rate, and youth graduate unemployment rate. This model is presented below:

XGBRegressor(base_score=0.5, booster='gbtree', colsample_bylevel=1, colsample_bynode=1, colsample_bytree=0.8, gamma=0, gpu_id=-1, importance_type='gain', interaction_constraints='', learning_rate=0.1, max_delta_step=0, max_depth=8, min_child_weight=1, monotone_constraints='()', n_estimators=100, n_jobs=12, num_parallel_tree=1, objective='reg:squarederror', random_state=0, reg_alpha=0, reg_lambda=1, scale_pos_weight=1, subsample=1, tree_method='exact', validate_parameters=1, verbosity=None)

The performance of M5 on the TRAIN and TEST data is described in the following table:

Performance Indicators	TRAIN	TEST
R^2	0.999	0.976
Mean square error (MSE)	0.0000004	0.0001
Root mean square error (RMSE)	0.00006	0.011
Mean absolute error (MAE)	0.0005	0.008

Table 8-7 Performance indicators of model M5

The following learning curve (Figure 8-5) shows the good performance of our model on both the training and validation data sets.

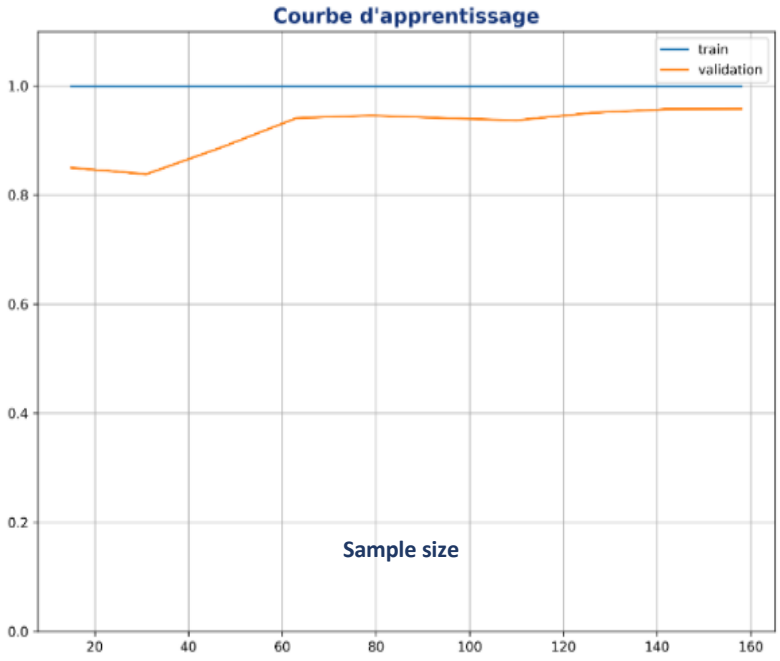


Figure 8-5 Learning curve of model M5

By using models M3, M4, and M5, we can predict the missing values for youth non-graduate unemployment, graduate unemployment, and youth graduate unemployment in China 2017. This results in the following table:

Country	Graduate unemployment rate	Non-graduate unemployment rate	Young non-graduate unemployment rate
China	0.057	0.034	0.096

Table 8-8 Predictions by models M3, M4, and M5 for Group C

As we have seen, the XGBoost procedure is a very powerful method, leading to models based on XGBRegressor having very high coefficients of determination on the one hand and predicting missing values with very low residual errors on the other. Furthermore, the total set of 80 countries for which all data is available has now been expanded by 9 new countries (group A, group B, group C). To provide an overall view of the sample size used in the following analysis, we have measured the total population for the 89 countries covered in 2017. The total population covered by these 89 countries is approximately 5.88 10^9, compared to the global population of 7.5110^9 in 2017. In other words, these 89 countries represent 78.3% of the world's total population. It is not possible to provide figures on the number of unemployed individuals in different categories, as we do not have data on the active population for each of the 89 countries.

8.4 Data sets X, Y, Z, W

8.4.1 Data set X: countries with a high general unemployment rate

Data set X includes 44 countries among the 89. It represents countries where the general unemployment rate is higher than the median unemployment rate across all 89 countries. The following figures illustrate the relationships between the Alpha Rate and the unemployment rate of graduates and young graduates in the countries within data set X.

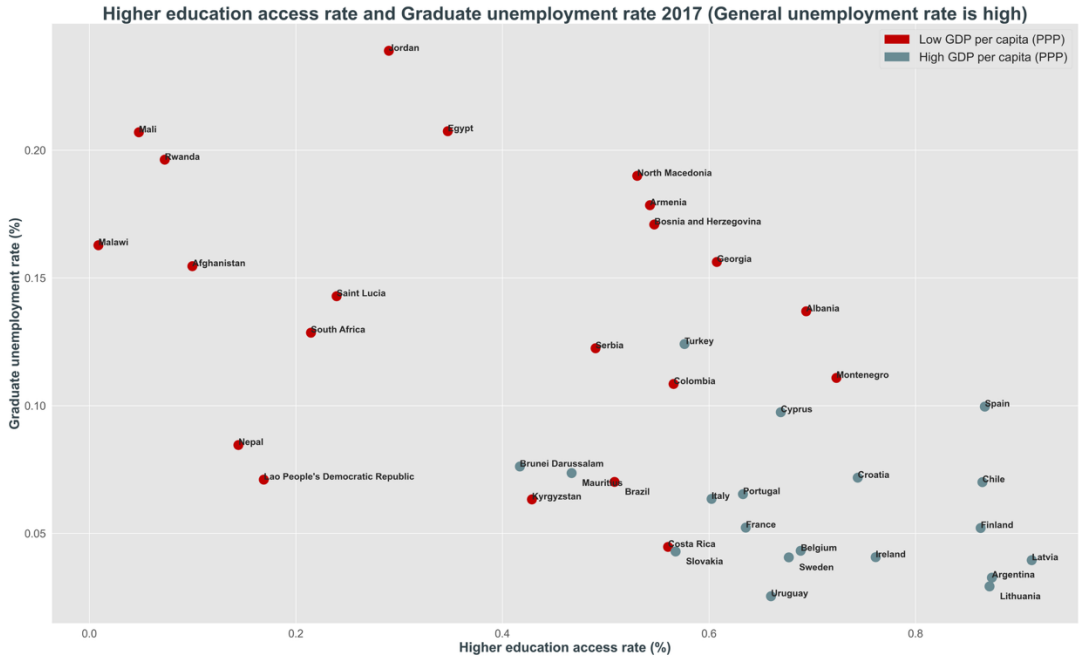


Figure 8-6 Higher education access rate vs graduate unemployment rate, X

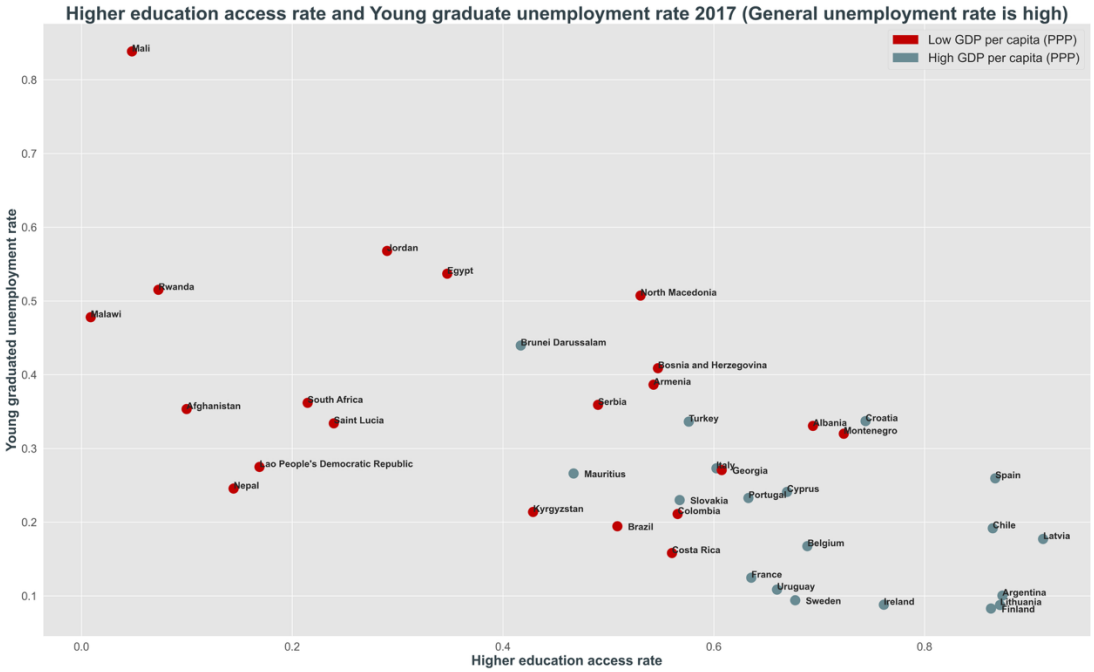


Figure 8-7 Higher education access rate vs young graduate unemployment rate, X

In both figures, the points are concentrated around a roughly horizontal line, implying strong correlations between the variables. The correlations between all variables in set X are presented in the following table.

	Alpha 2017	GDP ppp	Taux de chômage général	Taux de chômage des diplômés	Taux de chômage des non-diplômés	Taux de chômage des jeunes	Taux de chômage des jeunes diplômés	Taux de chômage des jeunes non-diplômés
Alpha 2017	(1.0, p=0.0, Reject H0)	(0.76, p=0.0, Reject H0)	(-0.324, p=0.032, Reject H0)	(-0.607, p=0.0, Reject H0)	(-0.154, p=0.318, Fail to reject H0)	(-0.165, p=0.286, Fail to reject H0)	(-0.694, p=0.0, Reject H0)	(-0.109, p=0.483, Fail to reject H0)
GDP ppp	(0.76, p=0.0, Reject H0)	(1.0, p=0.0, Reject H0)	(-0.387, p=0.009, Reject H0)	(-0.619, p=0.0, Reject H0)	(-0.207, p=0.177, Fail to reject H0)	(-0.106, p=0.494, Fail to reject H0)	(-0.6, p=0.0, Reject H0)	(-0.041, p=0.789, Fail to reject H0)
Taux de chômage général	(-0.324, p=0.032, Reject H0)	(-0.387, p=0.009, Reject H0)	(1.0, p=0.0, Reject H0)	(0.686, p=0.0, Reject H0)	(0.863, p=0.0, Reject H0)	(0.835, p=0.0, Reject H0)	(0.646, p=0.0, Reject H0)	(0.805, p=0.0, Reject H0)
Taux de chômage des diplômés	(-0.607, p=0.0, Reject H0)	(-0.619, p=0.0, Reject H0)	(0.686, p=0.0, Reject H0)	(1.0, p=0.0, Reject H0)	(0.552, p=0.0, Reject H0)	(0.488, p=0.001, Reject H0)	(0.901, p=0.0, Reject H0)	(0.432, p=0.003, Reject H0)
Taux de chômage des non-diplômés	(-0.154, p=0.318, Fail to reject H0)	(-0.207, p=0.177, Fail to reject H0)	(0.863, p=0.0, Reject H0)	(0.552, p=0.0, Reject H0)	(1.0, p=0.0, Reject H0)	(0.67, p=0.0, Reject H0)	(0.489, p=0.001, Reject H0)	(0.704, p=0.0, Reject H0)
Taux de chômage des jeunes	(-0.165, p=0.286, Fail to reject H0)	(-0.106, p=0.494, Fail to reject H0)	(0.835, p=0.0, Reject H0)	(0.488, p=0.001, Reject H0)	(0.67, p=0.0, Reject H0)	(1.0, p=0.0, Reject H0)	(0.517, p=0.0, Reject H0)	(0.981, p=0.0, Reject H0)
Taux de chômage des jeunes diplômés	(-0.694, p=0.0, Reject H0)	(-0.6, p=0.0, Reject H0)	(0.646, p=0.0, Reject H0)	(0.901, p=0.0, Reject H0)	(0.489, p=0.001, Reject H0)	(0.517, p=0.0, Reject H0)	(1.0, p=0.0, Reject H0)	(0.456, p=0.002, Reject H0)
Taux de chômage des jeunes non-diplômés	(-0.109, p=0.483, Fail to reject H0)	(-0.041, p=0.789, Fail to reject H0)	(0.805, p=0.0, Reject H0)	(0.432, p=0.003, Reject H0)	(0.704, p=0.0, Reject H0)	(0.981, p=0.0, Reject H0)	(0.456, p=0.002, Reject H0)	(1.0, p=0.0, Reject H0)

Table 8-9 Correlation table, set X

- A strong correlation (0.76) is observed between the Alpha Rate and GDP per capita (PPP) across countries with a high general unemployment rate.
- Additionally, there are strong correlations between the Alpha Rate and the graduate unemployment rate (-0.607) as well as between the Alpha Rate and the young graduate unemployment rate (-0.694). Across all countries, these correlations were (-0.38) and (-0.46), respectively.
- This suggests that the effect of increased access to education on reducing the graduate and young graduate unemployment rates is more significant in countries with high average unemployment rates.
- However, in these countries with a high average unemployment rate, there is no significant correlation between access to higher education and the unemployment rate of non-graduates or the general youth unemployment rate.

8.4.2 Data set Y: countries with a high unemployment rate among non-graduates

Data set Y includes 44 countries among the initial 89. It represents countries where the non-graduate unemployment rate is high, exceeding the median non-graduate unemployment rate across all 89 countries. The following figures illustrate the relationships between the Alpha Rate and the graduate and young graduate unemployment rates in data set Y.

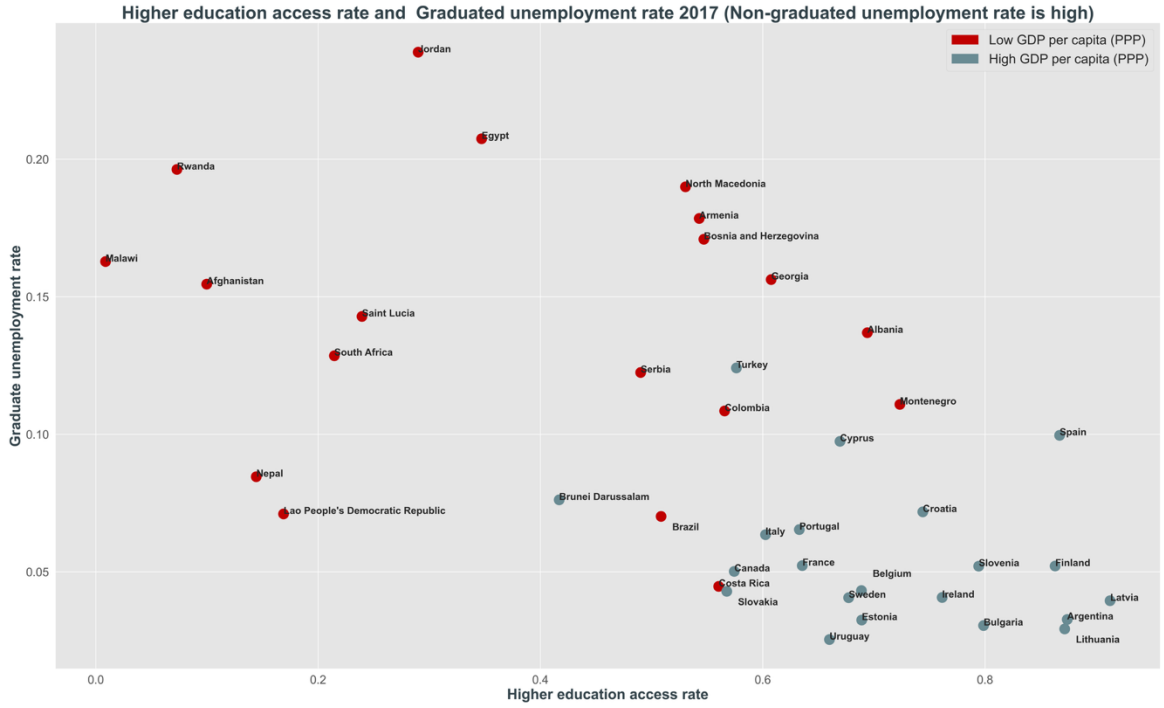


Figure 8-8 Higher education access rate vs graduate unemployment rate, Y

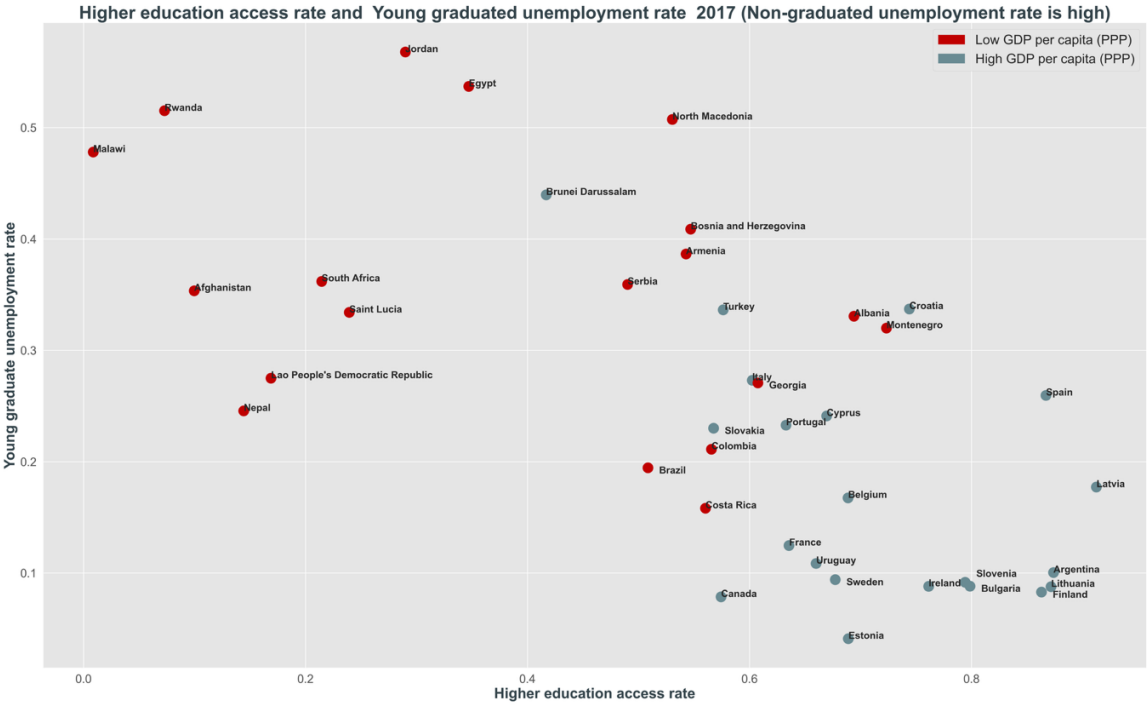


Figure 8-9 Figure 8 9 Higher education access rate vs young graduate unemployment rate, Y

The figures display the countries included in data set Y. Furthermore, the distribution of points on both figures implies strong correlations between the Alpha Rate and the graduate and young graduate unemployment rates. The correlations are provided in the following table:

	Alpha 2017	GDP ppp	Taux de chômage général	Taux de chômage des diplômés	Taux de chômage des non-diplômés	Taux de chômage des jeunes	Taux de chômage des jeunes diplômés	Taux de chômage des jeunes non-diplômés
Alpha 2017	(1.0, p=0.0, Reject H0)	(0.741, p=0.0, Reject H0)	(-0.369, p=0.013, Reject H0)	(-0.634, p=0.0, Reject H0)	(-0.187, p=0.22, Fail to reject H0)	(-0.211, p=0.163, Fail to reject H0)	(-0.689, p=0.0, Reject H0)	(-0.169, p=0.268, Fail to reject H0)
GDP ppp	(0.741, p=0.0, Reject H0)	(1.0, p=0.0, Reject H0)	(-0.424, p=0.004, Reject H0)	(-0.617, p=0.0, Reject H0)	(-0.231, p=0.126, Fail to reject H0)	(-0.183, p=0.228, Fail to reject H0)	(-0.619, p=0.0, Reject H0)	(-0.128, p=0.403, Fail to reject H0)
Taux de chômage général	(-0.369, p=0.013, Reject H0)	(-0.424, p=0.004, Reject H0)	(1.0, p=0.0, Reject H0)	(0.736, p=0.0, Reject H0)	(0.866, p=0.0, Reject H0)	(0.886, p=0.0, Reject H0)	(0.706, p=0.0, Reject H0)	(0.859, p=0.0, Reject H0)
Taux de chômage des diplômés	(-0.634, p=0.0, Reject H0)	(-0.617, p=0.0, Reject H0)	(0.736, p=0.0, Reject H0)	(1.0, p=0.0, Reject H0)	(0.603, p=0.0, Reject H0)	(0.547, p=0.0, Reject H0)	(0.9, p=0.0, Reject H0)	(0.503, p=0.0, Reject H0)
Taux de chômage des non-diplômés	(-0.187, p=0.22, Fail to reject H0)	(-0.231, p=0.126, Fail to reject H0)	(0.866, p=0.0, Reject H0)	(0.603, p=0.0, Reject H0)	(1.0, p=0.0, Reject H0)	(0.736, p=0.0, Reject H0)	(0.546, p=0.0, Reject H0)	(0.769, p=0.0, Reject H0)
Taux de chômage des jeunes	(-0.211, p=0.163, Fail to reject H0)	(-0.183, p=0.228, Fail to reject H0)	(0.886, p=0.0, Reject H0)	(0.547, p=0.0, Reject H0)	(0.736, p=0.0, Reject H0)	(1.0, p=0.0, Reject H0)	(0.591, p=0.0, Reject H0)	(0.985, p=0.0, Reject H0)
Taux de chômage des jeunes diplômés	(-0.689, p=0.0, Reject H0)	(-0.619, p=0.0, Reject H0)	(0.706, p=0.0, Reject H0)	(0.9, p=0.0, Reject H0)	(0.546, p=0.0, Reject H0)	(0.591, p=0.0, Reject H0)	(1.0, p=0.0, Reject H0)	(0.542, p=0.0, Reject H0)
Taux de chômage des jeunes non-diplômés	(-0.169, p=0.268, Fail to reject H0)	(-0.128, p=0.403, Fail to reject H0)	(0.859, p=0.0, Reject H0)	(0.503, p=0.0, Reject H0)	(0.769, p=0.0, Reject H0)	(0.985, p=0.0, Reject H0)	(0.542, p=0.0, Reject H0)	(1.0, p=0.0, Reject H0)

Table 8-10 Correlation table, data set Y

- As in the case of set X, strong negative correlations can be observed between the Alpha Rate and the graduate and young graduate unemployment rates across the countries in set Y.
- The correlation between the Alpha Rate and the graduate unemployment rate is notably negative (-0.634), as is the correlation between the Alpha Rate and the young graduate unemployment rate (-0.689).
- However, for countries where the non-graduate unemployment rate is high, there is no significant correlation between access to higher education and the unemployment rates of non-graduates, youth, or young non-graduates.

8.4.3 Data set Z: countries with a high youth unemployment rate

Data set Z includes 45 countries among the initial 89. It represents countries where the youth unemployment rate is high, exceeding the median youth unemployment rate across all 89 countries. The following figures illustrate the relationships between the Alpha Rate and the graduate and young graduate unemployment rates in data set Z.

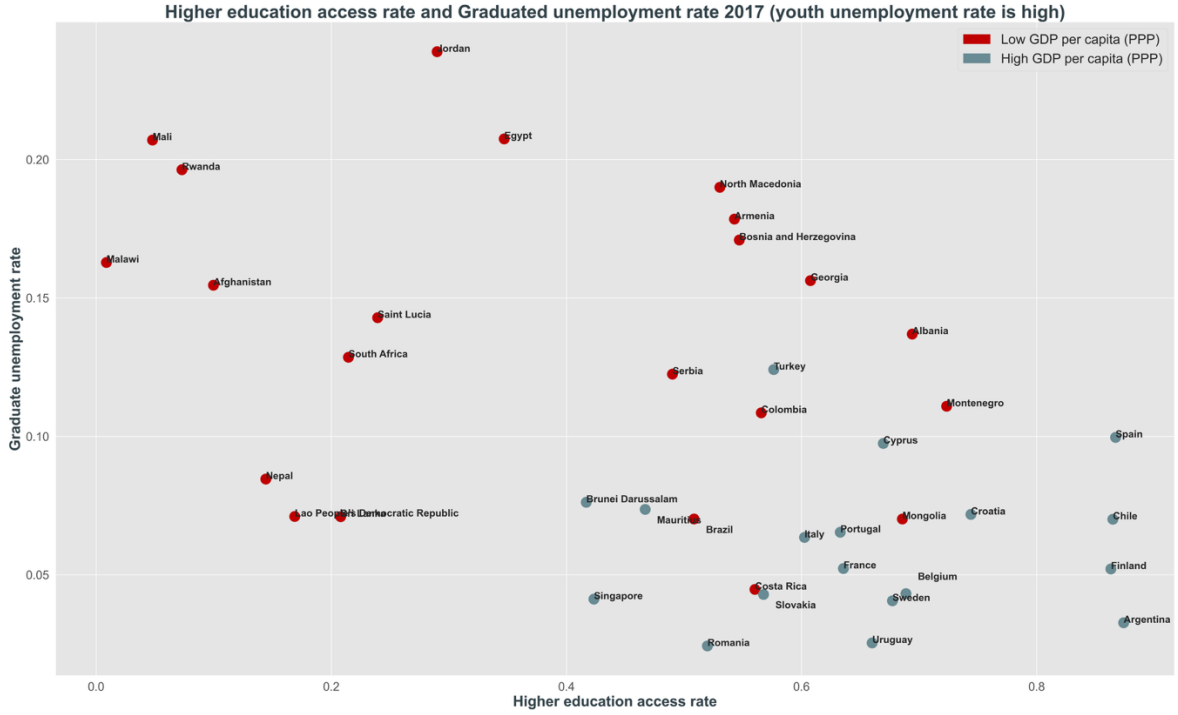


Figure 8-10 Higher education access rate vs graduate unemployment rate, Z

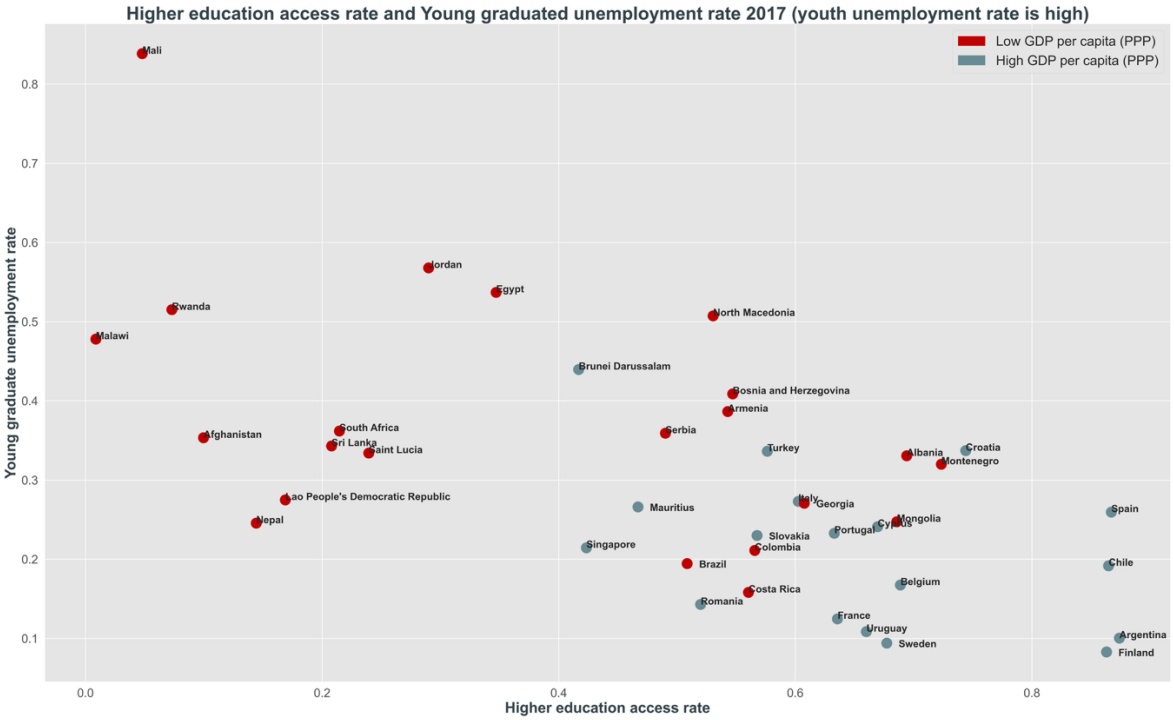


Figure 8-11 Higher education access rate vs young graduate unemployment rate, Z

The first figure shows a moderate correlation between the variables, while the second figure shows a stronger concentration of points along a line, indicating a stronger correlation. The following results are obtained:

	Alpha 2017	GDP ppp	Taux de chômage général	Taux de chômage des diplômés	Taux de chômage des non-diplômés	Taux de chômage des jeunes	Taux de chômage des jeunes diplômés	Taux de chômage des jeunes non-diplômés
Alpha 2017	(1.0, p=0.0, Reject H0)	(0.671, p=0.0, Reject H0)	(-0.145, p=0.348, Fail to reject H0)	(-0.476, p=0.001, Reject H0)	(0.016, p=0.918, Fail to reject H0)	(-0.044, p=0.777, Fail to reject H0)	(-0.604, p=0.0, Reject H0)	(0.026, p=0.867, Fail to reject H0)
GDP ppp	(0.671, p=0.0, Reject H0)	(1.0, p=0.0, Reject H0)	(-0.324, p=0.032, Reject H0)	(-0.634, p=0.0, Reject H0)	(-0.148, p=0.339, Fail to reject H0)	(-0.093, p=0.547, Fail to reject H0)	(-0.587, p=0.0, Reject H0)	(-0.02, p=0.898, Fail to reject H0)
Taux de chômage général	(-0.145, p=0.348, Fail to reject H0)	(-0.324, p=0.032, Reject H0)	(1.0, p=0.0, Reject H0)	(0.677, p=0.0, Reject H0)	(0.954, p=0.0, Reject H0)	(0.791, p=0.0, Reject H0)	(0.537, p=0.0, Reject H0)	(0.786, p=0.0, Reject H0)
Taux de chômage des diplômés	(-0.476, p=0.001, Reject H0)	(-0.634, p=0.0, Reject H0)	(0.677, p=0.0, Reject H0)	(1.0, p=0.0, Reject H0)	(0.514, p=0.0, Reject H0)	(0.538, p=0.0, Reject H0)	(0.883, p=0.0, Reject H0)	(0.443, p=0.003, Reject H0)
Taux de chômage des non-diplômés	(0.016, p=0.918, Fail to reject H0)	(-0.148, p=0.339, Fail to reject H0)	(0.954, p=0.0, Reject H0)	(0.514, p=0.0, Reject H0)	(1.0, p=0.0, Reject H0)	(0.774, p=0.0, Reject H0)	(0.387, p=0.009, Reject H0)	(0.812, p=0.0, Reject H0)
Taux de chômage des jeunes	(-0.044, p=0.777, Fail to reject H0)	(-0.093, p=0.547, Fail to reject H0)	(0.791, p=0.0, Reject H0)	(0.538, p=0.0, Reject H0)	(0.774, p=0.0, Reject H0)	(1.0, p=0.0, Reject H0)	(0.487, p=0.001, Reject H0)	(0.969, p=0.0, Reject H0)
Taux de chômage des jeunes diplômés	(-0.604, p=0.0, Reject H0)	(-0.587, p=0.0, Reject H0)	(0.537, p=0.0, Reject H0)	(0.883, p=0.0, Reject H0)	(0.387, p=0.009, Reject H0)	(0.487, p=0.001, Reject H0)	(1.0, p=0.0, Reject H0)	(0.379, p=0.011, Reject H0)
Taux de chômage des jeunes non-diplômés	(0.026, p=0.867, Fail to reject H0)	(-0.02, p=0.898, Fail to reject H0)	(0.786, p=0.0, Reject H0)	(0.443, p=0.003, Reject H0)	(0.812, p=0.0, Reject H0)	(0.969, p=0.0, Reject H0)	(0.379, p=0.011, Reject H0)	(1.0, p=0.0, Reject H0)

Table 8-11 Correlation table, data set Z

- We can observe a strong correlation between the Alpha Rate and GDP per capita (PPP).
- The correlation between the Alpha Rate and the graduate unemployment rate is -0.476, while the correlation between the Alpha Rate and the young graduate unemployment rate is notably negative (-0.604).
- However, for countries where youth unemployment is high, there is no significant correlation between access to higher education and the unemployment rates of non-graduates, youth, or young non-graduates.

8.4.4 Data set W: countries with a high unemployment rate among young non-graduates

Data set W includes 45 countries among the initial 89. It represents countries where the young non-graduate unemployment rate is high, exceeding the median young non-graduate unemployment rate across all 89 countries. The following figures illustrate the relationships between the Alpha Rate and the graduate and young graduate unemployment rates in data set W.

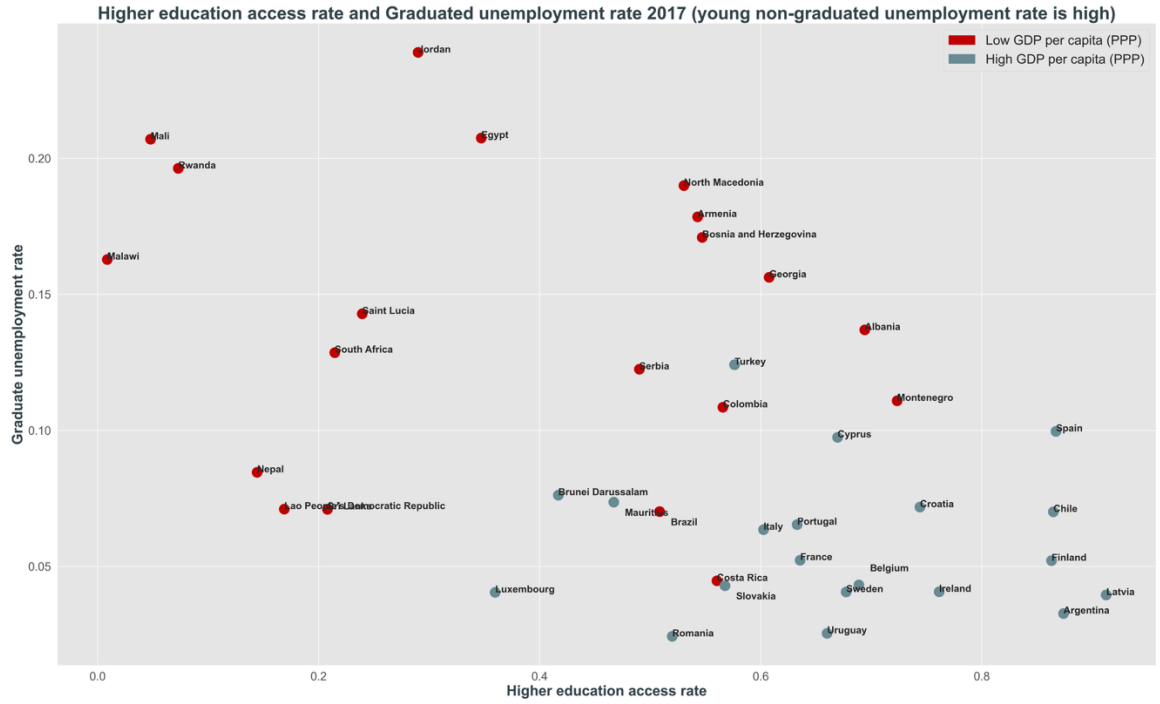


Figure 8-12 Higher education access rate vs graduate unemployment rate, W

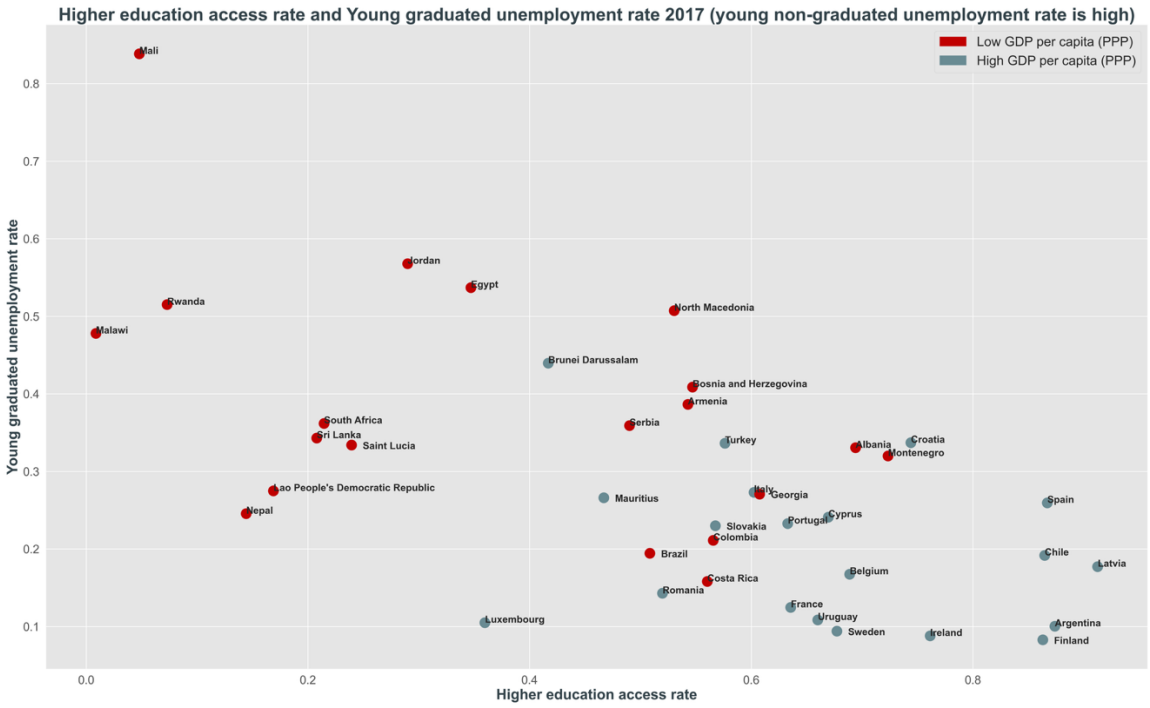


Figure 8-13 Higher education access rate vs young graduate unemployment rate, W

The correlations between the variables in data set W are presented in the following table:

	Alpha 2017	GDP ppp	Taux de chômage général	Taux de chômage des diplômés	Taux de chômage des non-diplômés	Taux de chômage des jeunes	Taux de chômage des jeunes diplômés	Taux de chômage des jeunes non-diplômés
Alpha 2017	(1.0, p=0.0, Reject H0)	(0.706, p=0.0, Reject H0)	(-0.127, p=0.412, Fail to reject H0)	(-0.498, p=0.001, Reject H0)	(-0.032, p=0.835, Fail to reject H0)	(-0.044, p=0.775, Fail to reject H0)	(-0.62, p=0.0, Reject H0)	(0.008, p=0.959, Fail to reject H0)
GDP ppp	(0.706, p=0.0, Reject H0)	(1.0, p=0.0, Reject H0)	(-0.35, p=0.02, Reject H0)	(-0.66, p=0.0, Reject H0)	(-0.23, p=0.133, Fail to reject H0)	(-0.17, p=0.27, Fail to reject H0)	(-0.647, p=0.0, Reject H0)	(-0.11, p=0.477, Fail to reject H0)
Taux de chômage général	(-0.127, p=0.412, Fail to reject H0)	(-0.35, p=0.02, Reject H0)	(1.0, p=0.0, Reject H0)	(0.642, p=0.0, Reject H0)	(0.896, p=0.0, Reject H0)	(0.809, p=0.0, Reject H0)	(0.541, p=0.0, Reject H0)	(0.8, p=0.0, Reject H0)
Taux de chômage des diplômés	(-0.498, p=0.001, Reject H0)	(-0.66, p=0.0, Reject H0)	(0.642, p=0.0, Reject H0)	(1.0, p=0.0, Reject H0)	(0.556, p=0.0, Reject H0)	(0.497, p=0.001, Reject H0)	(0.891, p=0.0, Reject H0)	(0.431, p=0.004, Reject H0)
Taux de chômage des non-diplômés	(-0.032, p=0.835, Fail to reject H0)	(-0.23, p=0.133, Fail to reject H0)	(0.896, p=0.0, Reject H0)	(0.556, p=0.0, Reject H0)	(1.0, p=0.0, Reject H0)	(0.672, p=0.0, Reject H0)	(0.442, p=0.003, Reject H0)	(0.72, p=0.0, Reject H0)
Taux de chômage des jeunes	(-0.044, p=0.775, Fail to reject H0)	(-0.17, p=0.27, Fail to reject H0)	(0.809, p=0.0, Reject H0)	(0.497, p=0.001, Reject H0)	(0.672, p=0.0, Reject H0)	(1.0, p=0.0, Reject H0)	(0.48, p=0.001, Reject H0)	(0.974, p=0.0, Reject H0)
Taux de chômage des jeunes diplômés	(-0.62, p=0.0, Reject H0)	(-0.647, p=0.0, Reject H0)	(0.541, p=0.0, Reject H0)	(0.891, p=0.0, Reject H0)	(0.442, p=0.003, Reject H0)	(0.48, p=0.001, Reject H0)	(1.0, p=0.0, Reject H0)	(0.406, p=0.006, Reject H0)
Taux de chômage des jeunes non-diplômés	(0.008, p=0.959, Fail to reject H0)	(-0.11, p=0.477, Fail to reject H0)	(0.8, p=0.0, Reject H0)	(0.431, p=0.004, Reject H0)	(0.72, p=0.0, Reject H0)	(0.974, p=0.0, Reject H0)	(0.406, p=0.006, Reject H0)	(1.0, p=0.0, Reject H0)

Table 8-12 Correlation table, data set W

- We can observe a strong correlation between the Alpha Rate and GDP per capita (PPP).
- A moderate correlation is observed between the Alpha Rate and the graduate unemployment rate (-0.498).
- The correlation between the Alpha Rate and the young graduate unemployment rate is notably negative (-0.62).
- Among these countries, there is no correlation between the Alpha Rate and the general unemployment rate, the youth unemployment rate, or the young non-graduate unemployment rate.

8.4.5 Definition of variance

Definition: Given a statistical series of a real variable $(x_1, x_2, \ldots, x_n)$, with a calculated mean $\bar{x}$, the variance is the mean of the squared deviations from this mean:

$$Var\ (X) = \frac{1}{n}\sum_1^n\ (x_i - \bar{x})^2.$$

Variance is a measure of dispersion of values, meaning it is always positive, equals zero only when all terms in the statistical series have the same value, and increases as the values become more spread out.

8.4.6 Cluster charts

Country	Alpha 2017	GDP per capita (PPP)	General unemplo-yment rate	Graduate unemplo-yment rate	Non-graduate unemployment rate	Youth unemployment rate	Young graduate unemployment rate	Young non-graduate unemployment rate	Cluster
Afghanistan	0.100	1964	0.11	0.15	0.11	0.18	0.35	0.16	A
Bangladesh	0.187	4054	0.04	0.11	0.04	0.13	0.36	0.12	A
Egypt	0.347	12138	0.12	0.21	0.10	0.30	0.54	0.25	A
India	0.278	7222	0.05	0.13	0.05	0.23	0.36	0.20	A
Lao People's Democratic Republic	0.169	7038	0.09	0.07	0.10	0.18	0.27	0.17	A
Mali	0.048	2253	0.07	0.21	0.07	0.20	0.84	0.19	A
Mauritania	0.058	4195	0.10	0.17	0.10	0.21	0.57	0.21	A
Nepal	0.144	2787	0.11	0.08	0.12	0.21	0.25	0.21	A
Niger	0.033	1015	0.08	0.05	0.08	0.17	0.24	0.17	A
Nigeria	0.131	5876	0.08	0.36	0.61	0.14	0.49	0.17	A
Rwanda	0.073	2074	0.17	0.20	0.17	0.23	0.52	0.22	A
Sri Lanka	0.208	12861	0.04	0.07	0.04	0.18	0.34	0.18	A
Togo	0.131	1666	0.04	0.14	0.03	0.10	0.30	0.09	A
Turkey	0.576	27510	0.11	0.12	0.10	0.21	0.34	0.18	A
Median	0.138	4124	0.09	0.14	0.10	0.19	0.36	0.18	A
Mean	0.177	6618	0.09	0.15	0.12	0.19	0.41	0.18	A

** Statistics related to Nigeria, considered non-reliable, weren't taken into account for the median and mean calculations*

Country	Alpha 2017	GDP per capita (PPP)	General unemplo-yment rate	Graduate unemplo-yment rate	Non-graduate unemplo-yment rate	Youth unemplo-yment rate	Young graduate unemployment rate	Young non-graduate unemployment rate	Cluster
Azerbaijan	0.319	17542	0.05	0.05	0.05	0.13	0.10	0.14	B
China	0.391	16750	0.04	0.06	0.03	0.11	0.19	0.10	B
Ecuador	0.501	11551	0.04	0.06	0.03	0.08	0.13	0.08	B
El Salvador	0.313	7875	0.04	0.05	0.04	0.10	0.19	0.10	B
Ethiopia	0.077	2026	0.02	0.04	0.02	0.03	0.14	0.03	B
Ghana	0.160	5296	0.04	0.05	0.04	0.09	0.19	0.09	B
Guatemala	0.288	8160	0.02	0.03	0.02	0.05	0.09	0.05	B
Honduras	0.247	4995	0.06	0.08	0.05	0.11	0.27	0.11	B
Indonesia	0.361	12363	0.04	0.05	0.04	0.15	0.19	0.14	B
Kyrgyzstan	0.429	3735	0.07	0.06	0.07	0.15	0.21	0.14	B
Mexico	0.406	19292	0.03	0.04	0.03	0.07	0.13	0.06	B
Thailand	0.505	18107	0.01	0.02	0.01	0.04	0.15	0.03	B
Vietnam	0.303	6854	0.02	0.04	0.02	0.07	0.18	0.06	B
Median	0.319	8160	0.04	0.05	0.03	0.09	0.18	0.09	B
Mean	0.331	10350	0.04	0.05	0.03	0.09	0.17	0.09	B

Country	Alpha Rate 2017	GDP per capita (PPP)	General unemplo-yment rate	Graduate unemplo-yment rate	Non-graduate unemployment rate	Youth unemployment rate	Young graduate unemployment rate	Young non-graduate unemployment rate	Cluster
Albania	0.694	12719	0.14	0.14	0.14	0.32	0.33	0.31	C
Armenia	0.543	9582	0.18	0.18	0.18	0.38	0.39	0.38	C
Bosnia and Herzegovina	0.547	12946	0.21	0.17	0.21	0.46	0.41	0.46	C
Brazil	0.509	15635	0.13	0.07	0.14	0.29	0.19	0.30	C
Brunei Darussalam	0.417	78873	0.09	0.08	0.10	0.29	0.44	0.28	C
Croatia	0.744	25526	0.11	0.07	0.13	0.27	0.34	0.27	C
Cyprus	0.669	37003	0.11	0.10	0.12	0.25	0.24	0.25	C
Georgia	0.608	10669	0.14	0.16	0.13	0.29	0.27	0.29	C
Italy	0.602	39630	0.11	0.06	0.13	0.35	0.27	0.35	C
Jordan	0.290	9196	0.18	0.24	0.16	0.35	0.57	0.30	C
Malawi	0.009	1186	0.29	0.16	0.29	0.41	0.48	0.40	C
Montenegro	0.723	18604	0.16	0.11	0.18	0.32	0.32	0.32	C
North Macedonia	0.530	15122	0.22	0.19	0.24	0.47	0.51	0.46	C
Saint Lucia	0.239	13986	0.19	0.14	0.22	0.46	0.33	0.50	C
Serbia	0.490	15897	0.13	0.12	0.14	0.32	0.36	0.31	C
South Africa	0.215	13461	0.27	0.13	0.30	0.53	0.36	0.55	C
Spain	0.867	38651	0.17	0.10	0.22	0.39	0.26	0.43	C
Median	0.543	15122	0.16	0.13	0.16	0.35	0.34	0.32	C
Mean	0.530	22201	0.16	0.13	0.17	0.35	0.36	0.35	C

Country	Alpha 2017	GDP per capita (PPP)	General unemplo-yment rate	Graduate unemplo-yment rate	Non-graduate unemployment rate	Youth unemployment rate	Young graduate unemployment rate	Young non-graduate unemployment rate	Cluster
Argentina	0.874	20815	0.08	0.03	0.10	0.23	0.10	0.23	D
Belgium	0.689	48034	0.07	0.04	0.09	0.19	0.17	0.21	D
Bulgaria	0.799	21367	0.06	0.03	0.08	0.13	0.09	0.13	D
Chile	0.865	24651	0.07	0.07	0.07	0.17	0.19	0.17	D
Colombia	0.566	14437	0.09	0.11	0.08	0.18	0.21	0.16	D
Costa Rica	0.560	17003	0.08	0.04	0.09	0.21	0.16	0.21	D
Dominican Republic	0.591	16482	0.06	0.05	0.06	0.13	0.15	0.14	D
Estonia	0.689	32652	0.06	0.03	0.07	0.12	0.04	0.13	D
Finland	0.863	45585	0.09	0.05	0.11	0.20	0.08	0.21	D
France	0.636	44074	0.09	0.05	0.12	0.22	0.12	0.26	D
Ireland	0.761	74704	0.07	0.04	0.09	0.14	0.09	0.16	D
Latvia	0.912	28004	0.09	0.04	0.11	0.17	0.18	0.17	D
Lithuania	0.872	32815	0.07	0.03	0.10	0.13	0.09	0.14	D
Mauritius	0.467	22340	0.07	0.07	0.07	0.24	0.27	0.23	D
Mongolia	0.686	12731	0.06	0.07	0.06	0.18	0.25	0.15	D
Poland	0.717	29820	0.05	0.02	0.06	0.15	0.11	0.15	D
Portugal	0.633	31688	0.09	0.07	0.10	0.24	0.23	0.24	D
Romania	0.520	25655	0.05	0.02	0.06	0.18	0.14	0.19	D
Russia	0.557	26618	0.05	0.05	0.05	0.16	0.18	0.15	D

Country	Alpha 2017	GDP per capita (PPP)	General unemplo-yment rate	Graduate unemplo-yment rate	Non-graduate unemployment rate	Youth unemployment rate	Young graduate unemployment rate	Young non-graduate unemployment rate	Cluster
Slovakia	0.568	32371	0.08	0.04	0.09	0.19	0.23	0.19	D
Slovenia	0.794	35432	0.07	0.05	0.07	0.11	0.09	0.11	D
Sweden	0.677	51573	0.07	0.04	0.08	0.18	0.09	0.19	D
Uruguay	0.660	22469	0.08	0.03	0.09	0.25	0.11	0.25	D
Median	0.686	28004	0.07	0.04	0.08	0.18	0.14	0.17	D
Mean	0.694	30927	0.07	0.05	0.08	0.18	0.15	0.18	D

Country	Alpha 2017	GDP per capita (PPP)	General unemplo-yment rate	Graduate unemplo-yment rate	Non-graduate unemployment rate	Youth unemployment rate	Young graduate unemployment rate	Young non-graduate unemployment rate	Cluster
Australia	0.660	49970	0.05	0.03	0.07	0.12	0.06	0.13	E
Austria	0.702	51921	0.05	0.03	0.07	0.10	0.06	0.11	E
Canada	0.574	47391	0.06	0.05	0.09	0.12	0.08	0.13	E
Czechia	0.604	38020	0.03	0.02	0.03	0.08	0.06	0.08	E
Denmark	0.751	52463	0.06	0.05	0.07	0.12	0.11	0.12	E
Germany	0.645	51539	0.04	0.02	0.04	0.07	0.04	0.07	E
Hungary	0.488	29213	0.04	0.02	0.05	0.11	0.07	0.11	E
Iceland	0.831	54488	0.03	0.02	0.03	0.08	0.06	0.08	E
Israel	0.622	37715	0.04	0.03	0.05	0.07	0.07	0.07	E
Japan	0.620	42414	0.03	0.02	0.03	0.06	0.04	0.08	E
Kazakhstan	0.564	26385	0.05	0.03	0.06	0.04	0.03	0.04	E
Korea	0.785	40077	0.04	0.04	0.03	0.10	0.10	0.10	E
Luxembourg	0.360	105395	0.06	0.04	0.06	0.15	0.11	0.16	E
Malta	0.495	41736	0.04	0.02	0.05	0.11	0.08	0.13	E
Netherlands	0.767	54242	0.05	0.03	0.06	0.09	0.05	0.09	E
Norway	0.824	67152	0.04	0.02	0.05	0.10	0.06	0.11	E
Qatar	0.088	128125	0.00	0.00	0.00	0.01	0.02	0.00	E
Singapore	0.423	95260	0.04	0.04	0.04	0.17	0.21	0.13	E
Switzerland	0.523	64269	0.05	0.04	0.06	0.08	0.06	0.08	E
United Arab Emirates	0.150	71186	0.02	0.04	0.02	0.08	0.14	0.07	E
United Kingdom	0.488	44611	0.04	0.03	0.05	0.12	0.08	0.13	E
United States	0.799	59964	0.04	0.03	0.06	0.09	0.05	0.10	E
Median	0.612	51730	0.04	0.03	0.05	0.10	0.06	0.10	E
Mean	0.580	56979	0.04	0.03	0.05	0.09	0.07	0.10	E

8.5 Correlations

8.5.1 Young graduate unemployment rate 2017 and other variables

```
Country                                          (0.097, p=0.374, Fail to reject H0)
Alpha 2017                                        (-0.494, p=0.0, Reject H0)
GDP 2017                                          (-0.713, p=0.0, Reject H0)
GDP scaled 2017                                   (-0.713, p=0.0, Reject H0)
Gini 2017                                          (0.386, p=0.0, Reject H0)
Classification 2017                               (-0.611, p=0.0, Reject H0)
Taux de chômage général 2017                       (0.479, p=0.0, Reject H0)
Taux de chômage des diplômés 2017                  (0.875, p=0.0, Reject H0)
Taux de chômage des non-diplômés 2017             (0.337, p=0.001, Reject H0)
Taux de chômage des jeunes 2017                    (0.537, p=0.0, Reject H0)
Taux de chômage des jeunes diplômés 2017             (1.0, p=0.0, Reject H0)
Taux de chômage des jeunes non-diplômés 2017       (0.475, p=0.0, Reject H0)
Alpha 2015                                        (-0.539, p=0.0, Reject H0)
GDP 2015                                          (-0.716, p=0.0, Reject H0)
GDP scaled 2015                                   (-0.716, p=0.0, Reject H0)
Gini 2015                                          (0.381, p=0.0, Reject H0)
Classification 2015                               (-0.611, p=0.0, Reject H0)
Taux de chômage général 2015                       (0.32, p=0.002, Reject H0)
Taux de chômage des diplômés 2015                  (0.824, p=0.0, Reject H0)
Taux de chômage des non-diplômés 2015             (0.236, p=0.028, Reject H0)
Taux de chômage des jeunes 2015                    (0.386, p=0.0, Reject H0)
Taux de chômage des jeunes diplômés 2015           (0.901, p=0.0, Reject H0)
Taux de chômage des jeunes non-diplômés 2015      (0.307, p=0.004, Reject H0)
Alpha 2013                                        (-0.565, p=0.0, Reject H0)
GDP 2013                                          (-0.716, p=0.0, Reject H0)
GDP scaled 2013                                   (-0.716, p=0.0, Reject H0)
Gini 2013                                         (0.362, p=0.001, Reject H0)
Classification 2013                               (-0.652, p=0.0, Reject H0)
Taux de chômage général 2013                    (0.155, p=0.152, Fail to reject H0)
Taux de chômage des diplômés 2013                  (0.636, p=0.0, Reject H0)
Taux de chômage des non-diplômés 2013           (0.048, p=0.658, Fail to reject H0)
Taux de chômage des jeunes 2013                 (0.161, p=0.137, Fail to reject H0)
Taux de chômage des jeunes diplômés 2013           (0.729, p=0.0, Reject H0)
Taux de chômage des jeunes non-diplômés 2013    (0.113, p=0.296, Fail to reject H0)
HCI                                               (-0.657, p=0.0, Reject H0)
Idh 2017                                          (-0.709, p=0.0, Reject H0)
Idh 2015                                          (-0.711, p=0.0, Reject H0)
Idh 2013                                          (-0.715, p=0.0, Reject H0)
Name: Taux de chômage des jeunes diplômés 2017, dtype: object
```

8.5.2 Alpha Rate 2013 and Other variables

```
Country                                          (-0.122, p=0.26, Fail to reject H0)
Alpha 2017                                        (0.962, p=0.0, Reject H0)
GDP 2017                                          (0.745, p=0.0, Reject H0)
GDP scaled 2017                                   (0.745, p=0.0, Reject H0)
Gini 2017                                        (-0.424, p=0.0, Reject H0)
Classification 2017                               (0.681, p=0.0, Reject H0)
Taux de chômage général 2017                     (0.088, p=0.415, Fail to reject H0)
Taux de chômage des diplômés 2017                 (-0.471, p=0.0, Reject H0)
Taux de chômage des non-diplômés 2017            (0.203, p=0.059, Fail to reject H0)
Taux de chômage des jeunes 2017                  (0.069, p=0.524, Fail to reject H0)
Taux de chômage des jeunes diplômés 2017          (-0.565, p=0.0, Reject H0)
Taux de chômage des jeunes non-diplômés 2017      (0.108, p=0.32, Fail to reject H0)
Alpha 2015                                         (0.98, p=0.0, Reject H0)
GDP 2015                                          (0.746, p=0.0, Reject H0)
GDP scaled 2015                                   (0.746, p=0.0, Reject H0)
Gini 2015                                        (-0.408, p=0.0, Reject H0)
Classification 2015                               (0.681, p=0.0, Reject H0)
Taux de chômage général 2015                     (0.229, p=0.033, Reject H0)
Taux de chômage des diplômés 2015                   (-0.4, p=0.0, Reject H0)
Taux de chômage des non-diplômés 2015            (0.292, p=0.006, Reject H0)
Taux de chômage des jeunes 2015                  (0.203, p=0.059, Fail to reject H0)
Taux de chômage des jeunes diplômés 2015          (-0.464, p=0.0, Reject H0)
Taux de chômage des jeunes non-diplômés 2015     (0.261, p=0.015, Reject H0)
Alpha 2013                                           (1.0, p=0.0, Reject H0)
GDP 2013                                          (0.748, p=0.0, Reject H0)
GDP scaled 2013                                   (0.748, p=0.0, Reject H0)
Gini 2013                                        (-0.409, p=0.0, Reject H0)
Classification 2013                               (0.632, p=0.0, Reject H0)
Taux de chômage général 2013                      (0.381, p=0.0, Reject H0)
Taux de chômage des diplômés 2013                (-0.176, p=0.104, Fail to reject H0)
Taux de chômage des non-diplômés 2013               (0.46, p=0.0, Reject H0)
Taux de chômage des jeunes 2013                   (0.407, p=0.0, Reject H0)
Taux de chômage des jeunes diplômés 2013         (-0.27, p=0.012, Reject H0)
Taux de chômage des jeunes non-diplômés 2013        (0.429, p=0.0, Reject H0)
HCI                                               (0.801, p=0.0, Reject H0)
Idh 2017                                           (0.82, p=0.0, Reject H0)
Idh 2015                                          (0.818, p=0.0, Reject H0)
Idh 2013                                          (0.823, p=0.0, Reject H0)
Name: Alpha 2013, dtype: object
```

9 References

1. Clustering, defined on page 7

2. Pierre Tapie, "Future Educational Directions & Challenges in Asia Pacific", Keynote Speech, Asia Pacific Deans Summitt, Séoul (Korea), August 28th 2014 ; Pierre Tapie, "Internationalization and the student body", Keynote Speech, Canadian Federation of business School Deans, Toronto, October 17th, 2014, ; Pierre Tapie, "Higher Education Demographics and Economic New Frontiers", Keynote speech, Higher Education Summitt, 19 Octobre 2015 ; Pierre Tapie, "Singapour : Higher Education and international Mobility Trends beyond Europe", Keynote Speech, Conference TIME, October 2019

3. Gary S.Becker, "Investment in Human Capital: A Theoretical Analysis", 1962, https://www.nber.org/system/files/chapters/c13571/c13571.pdf; Gary S.Becker, "Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education(3rd ed.)", 1993, Chicago: University of Chicago Press

4. Harald Schomburg, "Higher Education and Graduate Employement in Germany", 2000, https://www.jstor.org/stable/1503705

5. Jose-Gines Mora, José Garcia-Montalvo & Adela Garcia-Aracil, "Higher Education and Graduate Employment in Spain", 2000, https://www.jstor.org/stable/1503709

6. Esther Duflo, "The Medium Run Effects of Educational Expansion: Evidence from a Large School Construction Program in Indonesia," 2004, Journal of Development Economics, Vol. 74, 163-197, https://www.sciencedirect.com/science/article/abs/pii/S0304387803001846

7. Marie-Pierre Moreau & Carole Leathwood, "Graduates' employment and the discourse of employability: a critical analysis", 2006, European Journal of Education, https://www.tandfonline.com/doi/abs/10.1080/13639080600867083

8. Julia Varga, "Why to get a 2nd diploma? Is it life-long learning or the outcome of state intervention in educational choices?", 2006, Budapest Working Paper on the Labour Market, BWP, Institute of Economics, Hungarian Academy of Sciences, https://vmek.oszk.hu/06300/06311/06311.pdf

9. Thomas Plumper & Christina Schneider, "Too Much to Die, Too Little to Live Unemployment, Higher Education Policies and University Budgets in Germany", 2007, Journal of European Public Policy, https://www.researchgate.net/publication/228177374_Too_Much_to_Die_Too_Little_to_Live_Unemployment_Higher_Education_Policies_and_University_Budgets_in_Germany

10. Rosa Dias & Dorrit Posel, "Unemployment, Education and Skills Constraints in Post-Apartheid South Africa", 2007, University of Cape Town, Development Policy Research Unit, https://open.uct.ac.za/server/api/core/bitstreams/ae8a7d7a-dd15-4ae2-932f-c35a4e27073f/content

11. Imanol Nunez. & Ilias Livanos, "Higher education and unemployment in Europe: an analysis of the academic subject and national effects", 2009, Higher Education, https://www.jstor.org/stable/40602413

12. Lisa Kahn, "The long-term labor market consequences of graduating from college in a bad economy", 2010, Labour Economics, https://www.sciencedirect.com/science/article/abs/pii/S0927537109001018

13. Ekrem Erdem & Can Tansel Tugcu, "Higher Education and Unemployment: A co-integration and causality analysis of the case of Turkey", 2012, European Journal of Education, https://www.jstor.org/stable/23272491

14. Brett Lockard & Michael Wolf, "Employment outlook : 2010-2020 ; Occupational employment projections to 2020", 2012, 135 Monthly Lab. Rev. 84, https://heinonline.org/HOL/LandingPage?handle=hein.journals/month135&div=10&id=&page=

15. Marla McDaniel & Daniel Kuehn, "What Does a High School Diploma Get You? Employment, Race, and the Transition to Adulthood, The review of black political economy", 2013, Vol. 40, Issue 4, https://doi.org/10.1007/s12114-012-9147-1

16. Damon Clark & Paco Martorell, "The signaling Value of a High School Diploma", 2014, Journal of Political Economy Vol. 122 Number 2, https://www.journals.uchicago.edu/doi/abs/10.1086/675238

17. Ilga Lavrinovicha, Olga Lavrinenko & Janis S. Teivans-Treinovskisont, "Influence of education on unemployment rate and incomes of residents", 2015, Procedia - Social and Behavioral Sciences, https://www.researchgate.net/profile/Olga-Lavrinenko-2/publication/277651400_Influence_of_Education_on_Unemployment_Rate_and_Incomes_of_Residents

18. Diana Barbu, "The relation between unemployment and college enrolment and success outcomes", 2015, Electronic Theses, The Graduate School, Florida State University

19. Jennifer Delaney & William R.Doyle, "State spending on higher education: Testing the balance wheel over time", 2011, Journal of education finance, https://www.researchgate.net/publication/236709515

20. Thierry Kamionka & Xavier Vu Ngoc, "Trajectoire des jeunes sur le marché du travail, quartier d'origine et diplôme : une modélisation dynamique", 2015, Working Papers, Center for Research in Economics and Statistics

21. Mohd Sahandri Gani Bin Hamzah, Saifuddin Kumar Bin Abdulla & Mazura Mastura Binti Muhammad, "The Evaluation of Employment Marketability Connectivity Skills Within Polytechnic Engineering Diploma Students in Malaysia", 2016, US-China Education Review A, Vol. 6, No. 4, p. 230-243

22. Sharanjit Uppal, "Young men and women without a high school diploma", 2017, Statistics Canada, https://files.eric.ed.gov/fulltext/ED585313.pdf

23. Nigusse Weldemariam Reda & Mulugeta Tsegai Gebre-Eyesus, "Graduate unemployment in Ethiopia: the "Red Flag" and Its Implications", 2019, International Journal of African Higher Education,

24. Deni Mazrekaj, Kristof De Witte & Sarah Vanteenkiste, "Labour Market and consequences of a high school diploma", 2018, Applied Economics, Vol. 51, 2019, Issue 21, https://www.tandfonline.com/doi/abs/10.1080/00036846.2018.1543939

25. Dilrabo Jonbekova, "The diploma disease in Central Asia : student's views about purpose of university education in Kazakhstan and Tajikistan", 2019, Studies in Higher Education, Vol 45, Issue 6, p. 1183-1196, https://www.tandfonline.com/doi/abs/10.1080/03075079.2019.1628199

26. Ming Cheng, Olalekan Adekola, Jo Clarisse Albia & Sanfa Cai, "Employability in higher education: a review of key stakeholder's perspectives", 2021, https://www.emerald.com/insight/content/doi/10.1108/heed-03-2021-0025/full/html

27. Katarzyna Cieslik, Anna Barford & Bhaskar Vira, "Young people not in Employment, Education or Training (NEET) in Sub-Saharan Africa: Sustainable Development Target 8.6 missed and reset", 2021, Journal of Youth Studies, 25(8), 1126–1147, https://doi.org/10.1080/13676261.2021.1939287

28. P. Varsha Pramod & Remya Ramachandran, "Youth employment for inclusive growth: a review and research agenda in global perspective with special reference to India", 2023, J Glob Entrepr Res 13, https://doi.org/10.1007/s40497

29. Noreddine Oumansour & Youb Al Edrissi, "Microeconmetric evaluation of youth employment policies: empirical evidence for Morocco", 2023, https://revues.imist.ma/index.php/JISELSC/article/view/40715

30. Mohamed Niaré & Ousmane Mariko, "Unemployment in the WAEMU Countries: A Cross-Sectional Data Approach [Le chômage dans les pays de l'UEMOA : Une approche par données transversales]", 2023, https://ideas.repec.org/p/hal/journl/halshs-04313205.html

31. Nesrine Djellouli & Kahina Ait Hatrit, "La problématique du chômage et de l'emploi des titulaires du diplôme de doctorat en Algérie", 2023, https://asjp.cerist.dz/en/downArticle/160/19/2/231359

32. Esther Duflo, Pascaline Dupas, Elizabeth Spelke & Mark P. Walsh, "Intergenerational Impacts of Secondary Education: Experimental Evidence from Ghana", 2024, NBER Working Paper No. 32742 Ch 2.2, https://www.nber.org/system/files/working_papers/w32742/w32742.pdf

33. Diana Barbu, "The relation between unemployment and college enrolment and success outcomes", 2015, Electronic Theses, The Graduate School, Florida State University; Jennifer A.Delaney & William R.Doyle, "State spending on higher education: Testing the balance wheel over time", 2011, Journal of education finance

34. Note on Alpha Rate, defined on page 20

35. Note on China, page 54

36. Note on China, page 58

37. Tianqi Chen & Carlos Guestrin, "XGBoost: A Scalable Tree Boosting System", 2016, In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785–794), https://dl.acm.org/doi/pdf/10.1145/2939672.2939785

38. Details in appendix

39. Definition of variance page 75

40. Explanation of groups on page 77

41. Esther Duflo, "La pauvreté est multidimensionnelle", 2022, Collège de France, https://www.college-de-france.fr/fr/actualites/la-pauvrete-est-multidimensionnelle

42. Philippe Aghion, "Innovation and Growth from a Schumpeterian Perspective", 2018, https://www.jstor.org/stable/26596237 ; Philippe Aghion, Gilbert Cette, Élie Cohen & Jean Pisani-Ferry "Les leviers de la croissance française", p185, https://www.cae-eco.fr/staticfiles/pdf/072.pdf

10 Figures and tables

10.1 Figures

10.2 Tables