

**LUISS** 

Degree Program in Management

*Department of Business and Management*

Course of International Economics

**The Effects of Globalization and Trade  
Openness on Economic Growth: Evidence  
from Panel Data**

Prof. Luca De Benedictis

---

SUPERVISOR

Prof. Gianfranco Di Vaio

---

CO-SUPERVISOR

Daniele Fiori 778591

---

CANDIDATE

Academic Year 2024/2025

## **Abstract**

Economic globalization and its effects play a predominant role in academic literature, and a central and debated question is whether globalization has positive economic outcomes. This study investigates the effect of economic globalization and trade openness on economic growth. It aims to answer the following research questions: does globalization have a significant effect on economic growth? Does trade openness have a significant effect on economic growth? Between globalization and trade openness, which one exerts a stronger influence on economic growth? To answer these questions, this study employs a mixed bibliographic approach, combining a quantitative bibliometric analysis - which will focus the analysis on the effect that globalization and economic integration have on economic growth - with a qualitative examination of the two predominant measures in the literature: the KOF Globalization Index and trade openness. A two-part analysis is then carried out. First, scatter plots visualize the correlations between the KOF Index, trade openness, and economic growth. Second, a panel data analysis based on Dreher's (2002) econometric models, updated by Dreher (2008) and Gozgor (2017), empirically test these relationships. Results indicate no consistent and significant effect of globalization or economic integration on growth, except for a modest positive influence from the economic dimension of the KOF Index and trade openness. No clear evidence suggests that either the KOF Globalization Index or trade openness has a consistently stronger influence on economic growth. This thesis contributes to the academic debate by proposing a novel methodological approach to studying the effect of globalization on economic growth.

## Table of Contents

1. Introduction .....	4
1.1 Academic background .....	4
1.2 Research question .....	5
1.3 Research gap.....	6
2. Literature review.....	8
2.1 Quantitative literature review .....	8
2.1.1 <i>Creation of the data-set of papers</i> .....	9
2.1.2 <i>Analysis of selected papers</i> .....	11
2.2 Qualitative literature review .....	19
2.2.1 <i>The KOF Globalization Index and other composite indexes</i> .....	19
2.2.2 <i>Trade openness</i> .....	22
3. Methodology.....	25
3.1 Correlation Analysis .....	25
3.2 Generalized Method of Moments Regression .....	26
3.2.1 <i>Source and structure of variables</i> .....	28
4. Analysis .....	31
4.1 Globalization, trade openness, and economic growth: Correlation Analysis.....	31
4.2 Globalization, trade openness, and economic growth: GMM Regression .....	37
4.2.1 <i>GDP Growth and globalization</i> .....	38
4.2.2 <i>GDP Growth and economic integration</i> .....	40
4.3 Discussion.....	44
4.3.1 <i>Limitations and future research</i> .....	46
5. Conclusion.....	48
References .....	50
Appendix I: Detailed process of Literature review .....	56
Appendix II: Code of Methodology and Analysis .....	65

# 1. Introduction

Globalization is one of the most debated phenomena of recent decades, with major implications for the economy, politics, and society. It is such a complex phenomenon that scholars are unable to agree on a single, universally accepted definition. In fact, the debate over its definition itself is a subject of discussion in the literature (Al-Rodhan & Stoudmann, 2006; Scholte, 2002). The definition given by Palan et al. (2020) - taken from Norris (2000, p.2) - is an excellent starting point: “[Globalization is] a process that erodes national boundaries, integrates national economies, cultures, technologies and governance, and produces complex relations of mutual interdependence”. Palan et al. also distinguish between globalization and internationalization. The latter, in fact, implies taking the individual nation as the focal point of analysis, while the former “emphasizes the interdependencies of countries that result in deeper trade relations” (Palan et al., 2020, p.1). Therefore, when talking about globalization, a global, supranational point of view is assumed. The topic has been of public interest for several decades, partly because globalization is a cultural as well as an economic issue (Edwards, 2006), and partly because its implications also concern the ideological and emotional sphere of individuals. As Martens and Zywiets (2006, p.1) argue, “more people have an opinion about globalization than a deep understanding of the concept”.

## 1.1 Academic background

In economics, globalization has been particularly associated with its economic implications. In other words, what is most discussed is economic globalization. To define what is meant by economic globalization, the definition employed by Brady et al. (2005, p. 922) is particularly apt: “globalization as the intensification of international economic exchange and the label for the contemporary era of international economic integration. Thus, globalization involves the current economic environment shaping welfare states and the heightening of concrete economic exchanges between countries”. This definition underlines the close link between globalization and economic openness, described by Gräbner et al. (2020, p. 89) as “the most common term for capturing phenomena of increasing international integration in

trade and finance.” Economic openness is therefore a direct consequence of economic globalization, which manifests itself through trade and financial flows between countries worldwide. A central question in the literature is whether globalization and economic openness foster economic growth. In the literature, both globalization and economic openness have been analyzed as potential drivers of economic growth. Composite indices such as the KOF Globalization Index, first introduced by Dreher (2002) are typically employed to capture globalization in its multidimensional form, while economic openness is most frequently proxied by measures of trade openness and, to a lesser extent, financial openness (Gräbner et al., 2020). Trade openness has been the indicator most widely associated with growth, given its direct measurability through import and export data (Fujii, 2017; Squalli & Wilson, 2011), therefore, we will consider trade openness as a proxy for economic openness in our analysis of economic growth.

Several studies demonstrate the positive effects of globalization on economic growth (e.g., Dreher 2002 and 2008; Gozgor, 2017; Gygli et al., 2018), while other studies argue that there is no direct influence, or even that the influence is not positive (e.g., Rodrik & Rodriguez, 1999; Osterloh 2010; Hasan, 2019). With regard to trade openness, some studies have demonstrated its positive effect on economic growth (e.g., Yanikkaya, 2003; Huchet-Bourdon et al., 2017; Yu & Meng, 2023), but others highlight critical aspects (e.g., Rodrik & Rodriguez, 1999).

The lack of consensus shows that the relationship between globalization and growth remains controversial and therefore continues to attract the attention of scholars.

## 1.2 Research question

Considering the academic background outlined in the introduction, this thesis aims to investigate the effects that globalization and trade openness have on economic growth. Specifically, this study aims to answer the following questions:

*Does globalization have a significant effect on economic growth? Does trade openness have a significant effect on economic growth? Between globalization and trade openness, which one exerts a stronger influence on economic growth?*

To address these questions, a literature review is carried out in Chapter 2, divided into two sections: a quantitative section (2.1), in the form of a bibliometric review aimed at main macro-trends present in the literature on economic globalization. Based on these results, which will direct the analysis towards the effect that globalization and economic openness have on economic growth, the qualitative literature review (section 2.2) focuses on exploring two measures identified in the literature - a composite indicator and a simple indicator - identified in the bibliometric review. Chapter 3 is dedicated to methodology. Sections 3.1 and 3.2 illustrate two different methodologies that will be used to compare and test these two indicators of globalization. Chapter 4 is dedicated to the main analysis. It applies the methodology in two distinct sections (4.1 and 4.2), and then a discussion (section 4.3) is carried out. Finally, the last chapter is dedicated to the conclusions that will be drawn reflecting on the work as a whole.

### 1.3 Research gap

It is clear from section 1.1 that the topic under analysis is certainly not new: there are numerous contributions in the literature on the effect of globalization and economic integration on economic growth. It is therefore important to specify what the literature has currently overlooked or neglected and that this thesis can help to provide. The methodology used will be explained in detail in sections 3.1 and 3.2, but to get an idea of the research gap that this dissertation aims to fill, it is necessary to first provide some methodological background. The papers in the literature that address this topic naturally use panel data, since in order to investigate the effect of one variable on another (in this case the effect of globalization and trade openness on economic growth) it is necessary to have observations over several years. There are multiple ways to analyze panel data, but two major ways are consistently used and have a solid background. The first method employs second generation

panel unit root tests which consider cross-sectional dependence, such as the one proposed by Pesaran (2007) and OLS/FE regressions or other regression techniques, such as the Pooled Mean Squares estimation (Pesaran et al., 1999) (e.g. Kilic, 2015; Hasan, 2019). The second method consists of using both OLS/FE and GMM regressions - for addressing endogeneity - without employing any panel unit root test for cross-sectional dependence (e.g. Dreher, 2002; Dreher et al., 2008; Samimi & Jenatabadi, 2014; Gozgor, 2017). It should also be noted that the above does not exhaust the methodological discussion on the subject: there are numerous studies that do not use either of the two methodologies described above, and most studies do not address the problem of endogeneity or do so in a less robust way than through GMM regressions (Potrafke, 2014). The methodology of this thesis adopts the second method for regression analysis.

While sharing the subject matter and methodology with numerous papers in the literature, this thesis also introduces some new features. Firstly, it employs a dual methodology that is unique in literature: the empirical approach based on OLS/FE and GMM regressions is preceded by a comparative analysis based on the use of scatter plots aimed at clearly visualizing the correlation between globalization, trade openness, and economic growth. With regard to empirical analysis through regressions, there are innovations that differ from what has been done before. First of all, among the various regressions carried out, two will be dedicated to the direct influence of trade openness on economic growth. In fact, there is still no paper in the literature that measures the effect of trade openness on economic growth using econometric models such as those employed in this thesis. Furthermore, we utilize updated data ranging from 1970 to 2022 for scatter plots and from 1992 to 2022 for regressions. Moreover, although the methodological approach was updated by Dreher (2008) and Gozgor (2017), the regression equations will follow the approach of Dreher (2002), which has not been faithfully replicated in the literature, which is a pity since the two models based on the effect of globalization on economic growth and the effect of economic integration on economic growth have a solid and consistent structure that is difficult to find in other papers. All these innovations with respect to previous literature constitute the gap in the literature in which this thesis fits.

## 2. Literature review

This chapter presents the literature review. It adopts a dual approach that combines both quantitative and qualitative methods. It is therefore divided into two complementary parts: a quantitative bibliometric review and a traditional qualitative review. The bibliometric review involves creating a dataset of academic articles selected according to the criteria described below, followed by the generation of analytical results such as graphs and tables using R Studio. This phase aims to systematically identify prevalent themes and research trends in academic discourse related to globalization and economic integration.

Based on these findings, the qualitative review delves deeper into the two main measures highlighted by the quantitative analysis: the KOF Globalization Index and trade openness. This section contextualizes and interprets the most influential contributions in literature, integrating the bibliometric results into a discursive analysis.

### 2.1 Quantitative literature review

As previously mentioned, this part of the literature review of the literature will be conducted in the form of a bibliometric review. According to Castillo-Vergara et al. (2018, p. 1), “a bibliometric analysis examines bibliographic material from an objective and quantitative perspective that proves useful in organizing information within a specific field; hence, a bibliometric analysis using keywords allows the analysis of details in the main topics of research within a domain and relationships at the micro level”. The backbone of this work is an R Markdown document created in R Studio, enabling statistical analyses to be integrated directly with textual explanations. The analysis of generated outputs (especially the keywords analysis) form the core of this section, the ultimate objective of which is to provide an overview of the themes that make up the literature on economic globalization. This section will be divided into two sub-parts. The first one (section 2.1.1) illustrates the creation process of the data-set of selected papers which will then be the subject of the analysis. At the end of section 2.1.1 is shown a PRISMA diagram, which stands for Preferred Reporting Items for

Systematic reviews and Meta-Analyses, which is prepared following the 2020 PRISMA guidelines (Page et al., 2021) and will show how the selection of the studies was done. The second sub-part (section 2.1.2) is aimed at commenting on and interpreting all the outputs. We decided to split the process of creating outputs from the analysis and commentary of actual outputs into two distinct sections. This is done in order to make the presentation of the outputs and their discussion more fluent, since each output is linked to the other and doing otherwise would make their consultation awkward and misunderstandable.

The methodology of the bibliometric review is adapted from He et al. (2019), who developed a framework to identify gaps in the literature on construction project success in the context of Construction Engineering and Management (CEM). This framework can be easily adapted to different topics, and in this case, it will be adopted within the concept of economic globalization.

### *2.1.1 Creation of the data-set of papers*

This section will show the process of creating the data-set of papers that will be used to produce the outputs of section 2.1.2. It should be noted that Scopus is the source from which all articles contained in the dataset were exported. It is also important to note that, even though this section was largely created using an R Markdown document in R Studio, all references to code have been omitted and it was decided not to show any technical detail. However, comprehensive technical details and the whole code are available in the dedicated Appendix I for those interested.

The first step of the quantitative literature review consists of the identification of the macro topics that compose the area of interest. The area of interest has already been properly addressed, together with the research question and the research gap. Consequently, four key papers have been identified, constituting the starting point of the quantitative analysis. The papers are the following:

- Robustness of the KOF index of economic globalization, Gozgor, 2017

- Understanding economic openness: A review of existing measures, Gräbner et al., 2020
- The KOF globalization index - revisited, Gygli et al., 2018
- Measuring Fifty Years of Trade globalization, Palan et al., 2020

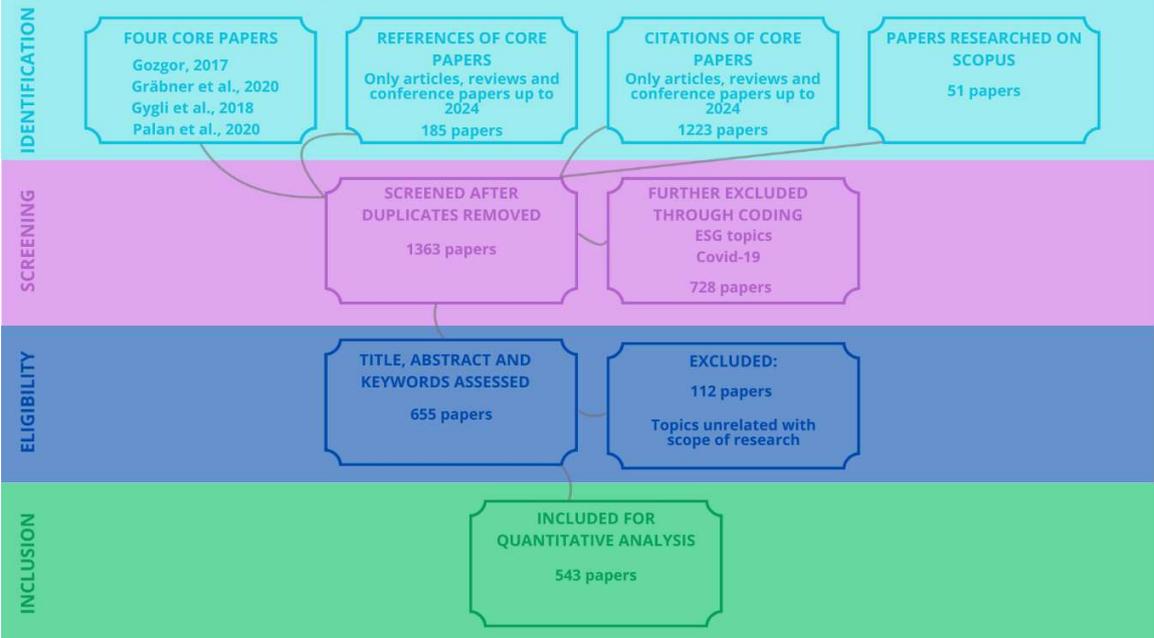
To establish an initial core of papers that will form the complete data-set, all the references for each of the four papers are exported and added to the dataset, which at this point consists of 189 papers. Thereafter, the data-set is enriched with all articles present in Scopus that cite the four articles, so as to include all subsequent literature. Then, a Scopus search on the topics of globalization, economic openness, and economic growth is conducted to include 51 more relevant articles. Additionally, the top 50 papers per number of citations in Scopus when searching “globalization” in the title search are included. The reason for this is to further include the most influential articles on the subject of globalization, in order to ensure that we do not leave out the most influential literature on the subject. Then, the data-set is filtered in order to avoid repetitions and exclude papers that focus on different topics from our research scope. Prevailing unrelated topics were related to environmental, and sustainability matters and to the Covid-19 pandemic.

After collecting all the papers and filtering the data, the data-set now consists of 655 papers. The last step is to manually remove those papers whose scope of research is not suitable for our analysis, but that were nonetheless included in the data-set until now. To do that, an eligibility analysis is conducted, which consists of selecting only those papers considered suitable for analysis based on certain characteristics, in this case their affinity with the theme of the literature review. Detailed information on how the paper eligibility process is carried out can be found in Appendix I. The papers considered ineligible are then removed from the data-set. The final data-set consists of 543 papers.

A PRISMA diagram, which serves to show clearly and transparently how the selection of the studies in the bibliometric review of this thesis was done, is presented, describing its four different phases - identification, screening, eligibility and inclusion - identified by Cipollina

et al. (2024), who in turn relied on the 2020 PRISMA guidelines (Page et al., 2021). The Diagram is presented below.

**Figure 1 - PRISMA Diagram**



*PRISMA Diagram (Cipollina et al., 2024) of identification, screening, eligibility and inclusion stages of academic contributions. The resulting sample is obtained through a search on Scopus.*

*2.1.2 Analysis of selected papers*

The analysis of selected papers begins with the table presented below (Table 1). It shows the top ten number of publications in the dataset both per country and per author. Focusing on the part of the table dedicated to authors, five major authors can be distinguished. The most included author in the data-set is Giray Gozgor, with seventeen publications, followed by Sèna Kimm Gnanon, with eight publications. Gozgor’s work focuses on the theme of globalization. Of particular interest is his analysis of the effect of globalization on economic growth, in particular in his paper “Robustness of the KOF index of economic globalization” (Gozgor, 2017), which is one of the four key articles discussed in section 2.1.1, he conducts an empirical analysis that forms one of the foundations of the analysis carried out in Chapter

4. Gnanon's work focuses on trade openness and the link between trade openness and economic development. An important part of Gnanon's research involves investigating how fiscal policies, reforms, and aid in developing countries modulate the benefits and costs of openness, and how the trade–development link depends on structural factors - such as diversification, real exchange rates - and social factors - such as poverty. Other contributions important to the development of this thesis will naturally be revisited and explored in greater depth in section 2.2, dedicated to the qualitative review of the literature.

**Table 1 – Number of publications**

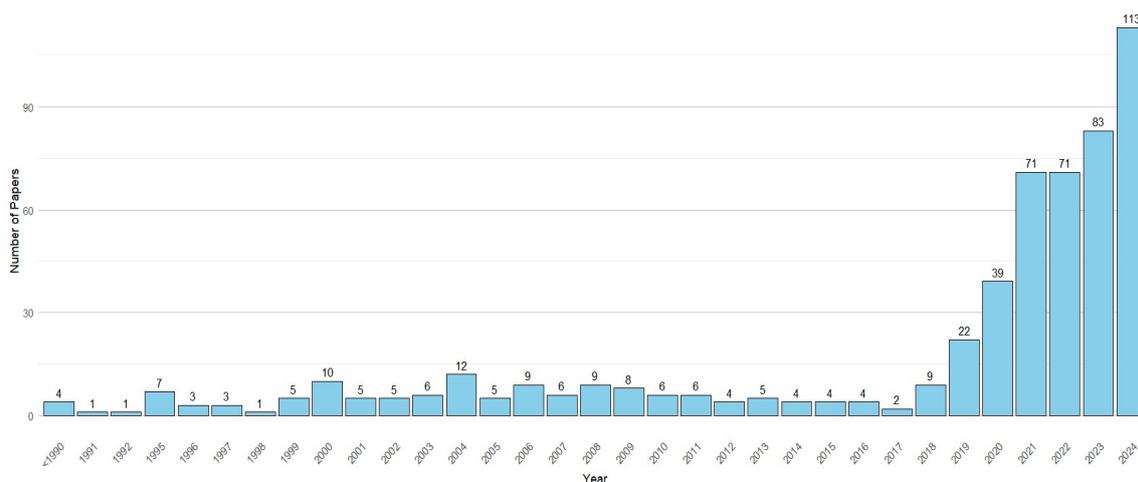
<i>Number of publications per country</i>		<i>Number of publications per author</i>	
Country	N° of publications	Author	N° of publications
United States	88	Gozgor G.	17
United Kingdom	40	Gnanon S. K.	8
Germany	31	Haini H.	7
China	25	Mahalik M. K.	7
Turkey	24	Ofori I. K.	7
Switzerland	19	Pal S.	7
India	15	Rodrik D.	7
Spain	15	Potrafke N.	6
Netherlands	12	Sturm J. E.	6
Australia	11	Camarero M.	5

Regarding the part of the table dedicated to countries, the most represented one is the United States, followed by two European countries - Germany and United Kingdom - and by China.

United States confirms itself the leading country in terms of academic presence on the topic of economic globalization. The graph also shows important presence of European Countries in the top ten, but also new world powers such as China and India. Moreover, rounding out the top ten there's Australia, who positions itself after Spain and Italy.

The next output (Figure 2) is obtained by processing and filtering publication data from our data-set, in order to visualize the number of papers published per year. The output consists of a graph obtained from a data frame named "papers\_per\_year", showing the count of papers per year, from which is derived a bar chart displaying the trend of publications over time.

**Figure 2 - Number of papers per year**

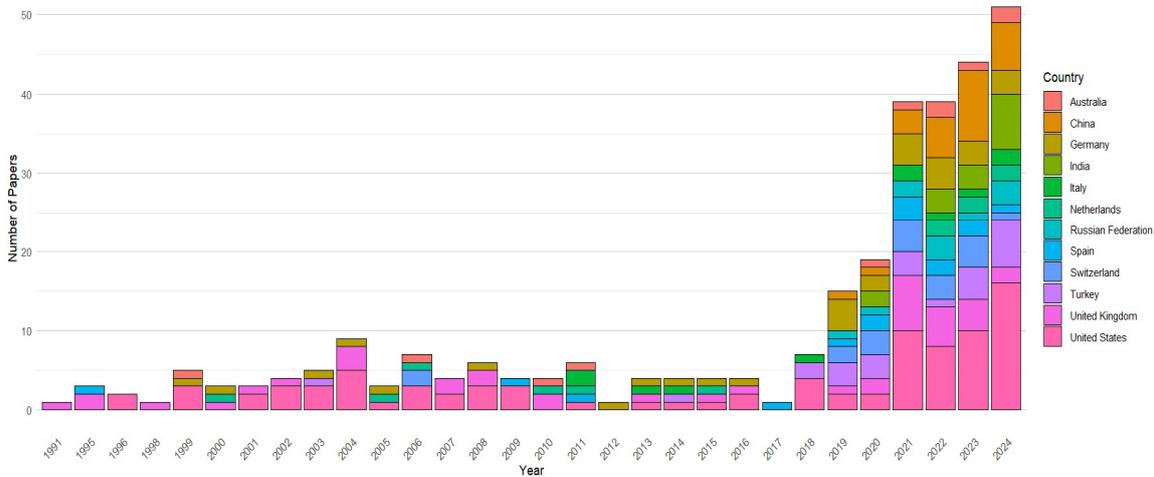


*Results prior to 1990 have been grouped together in a single column, the first one starting from the left.*

There is a great difference between the number of papers published before and after 2019-2020. The increase in number can be explained in two ways. The first explanation is that it is an intrinsic limit of the data-set: we chose to include in the data-set all papers present in Scopus that cited the four core papers, and this inevitably leads to an increase in more recent papers, which, using a different approach, might have been excluded in favor of older studies. Secondly, after 2019 there may have been an increase in academic interest towards economic

globalization and openness, which could be caused by the refinement of methodology, e.g. the recent improvement of the composite indicators of globalization (such as the revised version of the KOF Globalization Index by Gygli et al., 2018).

**Figure 3 - Number of papers per year (stacked by country)**



Results prior to 1990 have been grouped together in a single column, the first one starting from the left.

Figure 3 is a variation of Figure 2 and offers us a new perspective compared to the previous graph. It helps us assess which country has contributed most to literature each year in terms of relevant annual publications. The United States is undoubtedly the leading contributor, being the country with the highest number of relevant publications per year in almost every year since 1991. This finding is consistent with what was found previously in Table 1. It is also interesting to note that China and India have increased their presence in the ranking over the years, especially since 2019. To date, their contribution to literature, although still far from the annual publications of the United States, is undoubtedly among the top four countries in the last four years, after decades of almost non-existent contributions.

The next output, Table 2, is divided into two parts. The left part, “keywords ranking”, shows the top 10 keywords that appear most frequently in the selected papers. It comes as no surprise that the keyword most frequently found is by far “globalization”. “Economic growth” ranks second, suggesting how strong the link between globalization and economic growth is in literature. Another frequently used keyword is “trade openness”, ranking third. Two other important keywords are “economic development”, which once again outlines the important relationship in literature between globalization and its positive economic consequences, and “panel data”, which “refers to data sets consisting of multiple observations on each sampling unit” (Baltagi, 1998, p. 311). In the case of econometric papers, it is used in the analysis of data from multiple countries over a continuous period of time. This keyword is particularly relevant, given that in Chapter 4 we will carry out a panel data analysis, and its presence among the most frequent keywords confirms that it is the most popular type of data employed on the subject. Other important keywords are “foreign direct investment”, “international trade”, “financial development” and “economic globalization”, which once again suggest a predominance in the literature of themes relating to economic development, the economic integration of countries into a single global network, the importance of international trade, and the movement of capital within countries. Finally, the keyword “africa” is curiously the only one that refers to a geographical region, indicating a particular interest in literature towards the African continent.

**Table 2 – Keywords frequency**

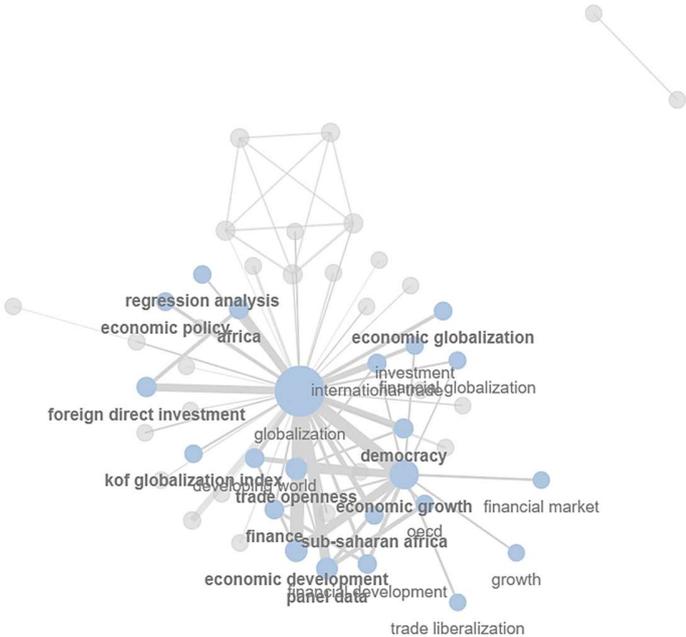
<i>Keywords ranking</i>		<i>Keywords co-occurrences</i>		
Keyword	Frequency	Keyword 1	Keyword 2	Frequency
globalization	248	economic development	globalization	35
economic growth	81	economic growth	globalization	32
trade openness	38	globalization	panel data	26
economic development	36	africa	globalization	22
panel data	33	economic integration	globalization	21
foreign direct investment	31	foreign direct investment	globalization	20
international trade	30	democracy	globalization	19
financial development	27	economic development	economic growth	19
economic globalization	24	globalization	international trade	18
africa	23	globalization	trade openness	18

The right part of the table, “keywords co-occurrences”, shows the ten pairs of keywords that appear most frequently together in the selected papers. “Globalization” is undoubtedly the word most closely linked with other concepts, in line with the left part of the table. With the exception of the keyword “democracy”, the words most closely associated with globalization are the same as those found in the left part of the table. The most frequent associations with globalization are “economic development” and “economic growth”. This confirms once again that the association between globalization and its positive economic consequences remains the central theme of the area of interest. Along the same lines is the association between globalization and “economic integration,” “international trade,” and “trade openness,” demonstrating once again the intertwining in literature between a country's levels of globalization and trade. The association between globalization and “africa” shows once

again how part of the literature is devoted to analyzing this issue in relation to the Africa, while its link with “democracy” shows how the literature is interested in investigating globalization from the perspective of OECD countries and, in general, in relation to the concept of democracy.

Figure 4 shows a keyword co-occurrence network - a very useful graph to visualize thematic relationships in scientific literature based on how keywords are used together in our selected papers. The graph is basically made up of two different elements. The first element are the nodes, which are the blue and grey circles, each representing a keyword. The other element are the arcs, which are the lines connecting the various nodes and representing the co-occurrence between two keywords within the articles. The thickness of the arcs is directly proportional to the frequency with which two keywords appear together. It is also important to note that the size of each node is directly proportional to the frequency with which the keyword emerges from literature, and this effect was intentionally achieved by showing more clearly the relationships of importance between the various keywords based on their frequency in addition to the number of arcs that the node representing each keyword has with the others.

**Figure 4 - Keywords network**



*The blue and grey circles are the nodes, each representing a keyword. The lines connecting the various nodes are the arcs, representing the co-occurrence between two keywords within the articles. The most important nodes, i.e., those mentioned in the discussion, have been highlighted in blue and their labels displayed.*

Once again, the most important keyword is “globalization”, as it is at the center of the network linked with basically all the other keywords. Among the many thematic centers linked to globalization, the strongest link (and therefore the thickest arc) is with economic growth. The graph therefore confirms once again what was stated in Table 2. Other important links related to globalization are “panel data”, “economic development”, “trade openness”, and “foreign direct investment”, once again in line with what was said for the previous tables. It is also interesting to note that trade openness is directly linked to both economic growth and panel data and democracy. This seems to suggest an interest in literature in the influence of trade openness on economic growth, through analyses using panel data. Other important links are with “democracy,” “africa,” and “sub-saharan africa,” which is therefore the African region most investigated by the literature on its relationship with globalization. Special mention should be made of the link between globalization and “kof globalization index.” The latter is in fact the main composite indicator used in literature to measure globalization, first presented by Dreher (2002).

It is intriguing to note that two different ways of measuring globalization in particular emerge from the graph: the KOF Globalization Index and trade openness. To be specific, the KOF index measures a country's globalization as a whole, while trade openness measures the country's level of outward orientation, i.e., its level of economic integration with other countries. How intertwined these two measures actually are, we will find out in the next two sections (sections 2.2.1 and 2.2.2) where they will be analyzed individually, with regards to academic literature.

The conclusions we can draw from this bibliometric review are that the theme of globalization in literature is particularly related to economic growth. In relation to economic growth, its link with trade openness is also important. A whole series of ancillary concepts closely related to these three can also be identified, such as “economic development”,

“international trade”, and “financial development”. In addition, the concept of panel data analysis often emerges, which is the best way to analyze the effect of globalization and trade openness on economic growth. Ultimately, the KOF Globalization Index, as the main measure of globalization, plays a strategically important role in literature on the subject.

## 2.2 Qualitative literature review

### 2.2.1 *The KOF Globalization Index and other composite indexes*

The keywords network of the previous section shows that, among all the nodes linked with globalization, there are two that deserve special attention: they are two important measures of globalization and economic integration. The first one is a composite index, and the second one is a simple indicator. In this section, the composite index, the KOF Globalization Index, is presented, described and compared with other composite indexes of globalization.

The KOF Globalization Index was first introduced by Dreher (2002), who aimed at assessing whether globalization affects growth, an effect which at that moment was supported by empirical studies that used trade flows as proxies. The index covers data of 122 countries collected from 1970 to 2002. Dreher himself in his paper states that the existent literature does not examine “the consequences of globalization on economic growth in greater detail. The effects reported might therefore appear only because other important aspects of globalization are omitted from the regressions” (Dreher, 2002, p. 3). Dreher identifies three main sub-dimensions of globalization: economic globalization, which is defined as “long distance flows of goods, capital and services as well as information and perceptions that accompany market exchanges”; political globalization which is “characterized by a diffusion of government policies”; and social globalization, which is “expressed as the spread of ideas, information, images, and people” (Dreher, 2002, p. 4). Economic globalization is measured through level of trade as a percentage of GDP (trade share), foreign direct investment, portfolio investment, hidden import barriers, mean tariff taxes and taxes on international trade to measure restrictions on capital and trade. Political globalization is assessed using the number of embassies in a country, the number of international organizations to which the

country is a member and the number of UN missions the country has participated in. Social globalization is the dimension of globalization most difficult to measure, and Dreher used 15 variables to measure it, divided into three main categories: data on personal contact, data on information flows, and data on cultural proximity. It is interesting how the author included among the variables some like “number of McDonalds’ restaurants per capita”, “average cost of calls to USA”, as he interprets cultural globalization “as the domination of American cultural products” since “the United States is the pacesetter in much of the global social-cultural realm” (Dreher, 2002, p. 5).

Dreher et al. (2010), compare the KOF Globalization Index, presented in a 2009 updated version, mostly regarding the usage of more recent data, with the main globalization composite indicators developed at the time. One of them is the Maastricht Globalization Index, which was developed independently at around the same time, and follows a different approach in order to measure globalization. In fact, The 2009 KOF Index also introduces a series of methodological improvements. First of all, each variable is transformed into an index with a scale from 1 to 100, while in the old version the scale was from 0 to 10, “where 100 is the maximum value for a specific variable (and hence highest globalization) over the period 1970 to 2006 and 1 is the minimum value” (Dreher et al. 2010, p. 179). Moreover, other methodological improvements included the transformation of the data according to the original distribution, meaning that “results are no longer driven by extreme outlying observations and missing values”, and “weights for calculating the sub-indices are determined using principal components analysis for the entire sample of countries and years”. (Dreher et al., 2010, pp. 179-180). The Maastricht Globalization Index (MGI) was developed by Martens and Zywiec (2006) and further updated by Martens and Raza (2009). It is composed of five domains: the political domain, the economic domain, the socio-cultural domain, the technological domain and the ecological domain. The domain structure resembles the sub-dimension structure of KOF, but the composition criteria for each domain differ from the composition criteria for KOF sub-dimensions: for example, technology in the KOF Index is included within the social dimension of globalization, while other aspects such as the environment are not even included in the KOF index, which considers it “outcome-

based” and therefore not a direct measure of globalization (Dreher et al., 2010). Among the other composite indexes of globalization, the Foreign Policy/Kearney Index of Globalization is the oldest one among those discussed. There are four macro-categories that make up the indicator: globalization in goods and services, financial globalization, globalization of personal contact, Internet connectivity. As can be seen, compared to KOF and MGI, this composite index focuses only on the economic/technological aspect of globalization, which makes it less versatile for analyzing the dynamics of globalization that differ from economic analysis and covers significantly fewer countries than the other two indexes. Moreover, Lockwood (2004, p. 507) questions the robustness of the index, stating that it is “constructed by making some problematic assumptions about the measurement, normalization and weighting of the variables included”.

Gygli et al. (2018), introduced a new version of the KOF Globalization index. The major changes include the separation of trade and financial dimensions of globalization within economic globalization and a clear distinction between de facto and de jure globalization measures. The KOF Index already employed both de facto and de jure measures, but Gygli et al. (2018) introduced a clear distinction between them by dividing each dimension into two distinct sub-dimensions, one de jure and one de facto, making the structure of the composite indicator much clearer. According to the authors, “de facto globalization measures actual international flows and activities, de jure globalization measures policies and conditions that, in principle, enable, facilitate and foster flows and activities” (Gygli et al., 2018, p. 2).

A significant part of the literature uses the KOF Index to assess the influence of globalization on economic growth, such as Dreher (2002 & 2010) and Gygli et al., (2018) themselves, but also many other contributions, with results that are not always consistent (e.g. Osterloh, 2010; Samini & Jenatabadi 2014; Radulović & Kostić, 2024).

### 2.2.2 Trade openness

The second measure that emerges from the bibliometric review is trade openness. It is defined by Fujii (2017, p. 869) “as the degree to which an economy maintains its outward orientation in trade”. The measures most commonly used in literature, as they are substantially simpler and more effective, are the empirical measures of trade openness: import share ( $\frac{M_i}{GDP_i}$ ), export share ( $\frac{X_i}{GDP_i}$ ), and trade share ( $\frac{(M+X)_i}{GDP_i}$ ), where M is the sum of country i’s imports and X is the sum of country i’s exports (Squalli & Wilson, 2011). Among these three, the main measure of actual trade openness is trade share (Harrison, 1995; Lloyd and MacLaren, 2002). These measures have been criticized for being influenced by the size of the country's economy: a country such as the United States, which is a leader in world trade, would paradoxically, due to its very high GDP, show a lower trade openness indicator than other countries that are much less open and less involved in global trade (Squalli and Wilson, 2011). Another problem with this measure is that import and exports at the numerator are in current prices, therefore the results “may diverge because of changes in the exchange rate or other relative price movements” (Lloyd and MacLaren, 2002, p. 70). Furthermore, Federico and Tena-Junguito (2017), state that export share would be a better measure than trade share, as imports include transportation and related costs. Moreover, another important way of capturing trade openness used in the literature is by measuring it through trade policies and trade restrictions rather than through trade volume. The difference between these two concepts is significant, as “the importance of these two dimensions lies in the ability to focus on actual trade flows rather than potential trade flows, as captured by lax or liberal trade as well as other relevant socioeconomic, geographic and demographic factors” (Yanikkaya, 2003, p.3). Therefore, a measure of trade volume can only measure one aspect of trade openness - that of actual trade flows - and will tell nothing about trade policies.

Despite all these complexities in measuring trade openness, doing so through trade share remains the most common measure employed in literature, and it will also be used in this thesis, for a number of good reasons. First of all, the fact that it is the most popular measure implies easy comparability with the results of other studies: using a more complex measure

can lead to comparability problems. Furthermore, a simple empirical indicator reduces the risks of arbitrariness in the definition of “trade openness”. Finally, in a panel data analysis such as this dissertation’s, the absolute value of trade openness for a given country is of little importance: what matters is that the variation in the proxy over time is able to capture an actual consequent variation in trade openness, and fortunately, trade share is able to capture this variation. Therefore, from now on in this dissertation, when referring to trade openness, it is implied that we are referring to trade share.

Trade openness is used in the analysis of various phenomena, from government size (Rodrik, 1998) to environmental issues (Destek & Sinha, 2020), but undoubtedly the most prominent area in which it is used and the most important for the purposes of this thesis is analysis related to economic growth. In fact, literature has long debated the correlation between openness and economic integration and economic growth. Firstly, Dollar (1992) found empirical evidence of the statistically significant relationship between a country’s outward-orientation and economic growth. The author defines outward-orientation as a combination of low level of protection, especially for inputs into the production process, and little variability of the real exchange rate, and it is a measure positively related to the number of tradable goods and inversely related with trade barriers. This paper had a major influence on the subsequent literature and paved the way for other studies that found evidence of the influence of trade openness on economic growth (e.g. Yanikkaya, 2003; Huchet-Bourdon et al., 2017; Yu and Meng, 2023). The empirical evidence on the positive effect of trade openness on economic growth was criticized by Rodriguez and Rodrik (1999). In particular, Rodrik’s academic work focuses on trade and economic growth and is characterized by his skeptical approach to empirical evidence. He argues that empirical evidence, which in the case of the papers included in our dataset concerns the correlation between greater trade openness and more open and favorable trade policies and economic growth (Rodriguez & Rodrik, 1999). He also emphasizes the importance of institutions and the political context in a country's economic destiny and the non-existence of a single solution that leads different countries to the same desired point of economic growth: different examples of success have followed different paths. In short, his contribution distances itself from a part of the literature

that seeks to demonstrate the inevitable and always true correlation between economic growth and factors such as the level of globalization and trade openness (Rodriguez & Rodrik, 1999).

Trade openness is also analyzed from different perspectives other than assessing economic growth. Another way of doing it is, for example, by focusing the analysis on a single nation, as Gaber and Elhassan (2020), Kaba and Renard (2022), Behanzin and Konte (2021) do, or on a group of countries sharing the same geographical and/or political background - such as Gnanon (2020a), (2020b), and (2021) does, or by macroeconomic parameters as Oloyede et al. (2021) do. There is also a focus on its relation with cultural globalization: while there is a certain affinity between a measure of trade and a measure of cultural globalization, both of them being a consequence of progressive internationalization of the globe, it is uncertain how strong their correlation will be, although the positive correlation between culture and trade has already been proved in academic literature (Coyne and Williamson, 2009; Coyne and Williamson, 2012).

## 3. Methodology

As already mentioned in the introductory chapter, the methodological section of this thesis is divided into two distinct parts. The first, section 3.1, concerns the methodology used in the comparative analysis through scatter plots (section 4.1) and the second, section 3.2, concerns the methodology used in the analysis through regressions (section 4.2). Furthermore, the quantitative analysis was carried out entirely in the R Markdown document in R Studio. In line with the bibliometric review in Chapter 2, it was decided to omit any reference to the code and technical details regarding the functioning of R Studio. The code used can be consulted in its entirety in Appendix II.

### 3.1 Correlation Analysis

For the purpose of answering the research questions, a correlation analysis based on scatter plots will firstly be conducted. The scatter plot illustrates the descriptive linear relationship between two variables over a determined time interval - in this case the KOF Index and one of its sub-dimensions and trade openness - and thus do not show causal indications but only visual patterns. The time interval starts from 1970 and ends in 2022. There have been identified three different time periods within the time interval: The first runs from 1970 to 1989, the second from 1989 to 2006 and the third from 2006 to 2022. Dividing the output in three different periods helps to obtain less noise and a more stable nature from the data compared to annual data, as each point represents a country in a certain time interval merged instead of a single year.

Regarding the variables to be included in the scatter plot analysis, we measure trade openness through the most used proxy in the literature, being the sum of exports and imports of a given country in terms of its GDP (trade share):  $\frac{(X + M)_i}{GDP_i}$ , first employed by Grassman (1980).

We measure globalization through the KOF Globalization Index, in the revised version of Gygli et al. (2018). We measure economic growth through the GDP per capita growth rate. In

the analysis, in section 4.1, a series of scatter plots concerning trade openness, KOF Index and per capita GDP growth are shown and discussed. A series of scatter plots is also shown and discussed, where observations are divided both by time period and level of country income according to three income brackets, to see whether the relationships between the variables change as the level of development of the countries changes.

## 3.2 Generalized Method of Moments Regression

As for the regressions, as previously anticipated, the approach used by Dreher (2002) and updated by Dreher (2008) and Gozgor (2017) will be followed, even though the version of the KOF Index utilized is the revised version of Gygli et al. (2018). This methodological approach was chosen because, although the KOF has been updated several times since its first publication, the econometric models used by Dreher (2002) regarding the relationship between globalization and economic integration and economic growth appear particularly suitable for the purpose of this thesis. The approach adopted by Dreher (2002) is the following: he first introduces the KOF Globalization Index. Once the index is constructed, Dreher estimates its effect on GDP per capita growth using OLS regressions with Fixed Effects (FE) and estimates its effect on the logarithm of GDP growth using generalized method of moments (GMM) regressions in differences - proposed by Arellano and Bond (1991) - to account for growth dynamics and endogeneity risk. An OLS regression represents the simplest starting point, where the effect of the independent variables is assumed to be homogeneous across countries and constant over time. However, this approach does not control unobserved heterogeneity across countries. The Fixed Effects regression improves the OLS one by including controls for country-specific fixed effects, capturing time-invariant characteristics such as culture, geography or institutions. However, the FE regression is not suitable in the presence of autoregressive dynamics, as the presence of a lagged dependent variable in the panel introduces correlation with error. If a regressor is correlated with the error, it means that the estimates are biased. GMM regressions, based on Arellano and Bond (1991), were developed specifically to address the problem of endogeneity. If there is reverse causality or if there are omitted variables that are correlated with both the regressors and the

error term, the GMM uses lagged values of the variables as valid instruments to obtain consistent estimates. This regression technique is particularly suitable in contexts where economic growth depends on its own past dynamics (e.g. convergence) and some of the explanatory variables are potentially endogenous. In addition, GMM allows for the treatment of unbalanced data and exploits the panel structure to construct instruments. The combination of OLS with Fixed Effects and GMM regressions is done in order to strengthen the robustness of the results and to compare estimates as assumptions on data characteristics change. Moreover, as stated in section 1.3, a methodological choice was made that distanced this thesis from works that employ second-generation panel unit root tests that consider cross-sectional dependence, such as those using Pesaran's approach (2007). In fact, it should be noted that cross-sectional dependence was not addressed in this thesis, in line with Dreher (2002 & 2008) and Gozgor (2017), who favor GMM to address endogeneity rather than spatial dependence.

Moreover, although the methodology used is indeed Dreher's (2002), and the regressions will be carried out as closely as possible to the econometric models described in his study, we employ some updates introduced by Dreher (2008) and Gozgor (2017), who used a different version of GMM regressions. In fact, while Dreher first used GMM in standard differences, based on Arellano and Bond (1991), his second work along with Gozgor's work use System GMM, based on the model of Arellano and Bover (1995) and Blundell and Bond (1998), which adds level equations to those in differences, creating a system of equations and therefore two sets of moment conditions, making the estimator more efficient and robust. System GMMs also deal with endogeneity and serial autocorrelation more appropriately. It is also important to mention that, regarding OLS/FE regressions, there is a problem of endogeneity. As Dreher (2006) mentioned in his paper, research shows that some variables are influenced by economic growth, for example, fertility may be influenced by growth (Barro and Lee, 1994). Moreover, the low  $R^2$  for all OLS/FE regressions (Tables 3 and 4) indicate that these regressions have a limited ability to explain the variability of the dependent variable.

The regressions will be tested using an approach as similar as possible to Dreher's (2002, pp. 24-25), therefore constructing two different econometric models, one confronting per capita GDP growth and globalization (section 4.2.1), and one confronting per capita GDP growth and economic integration (section 4.2.2). The first model will employ three OLS/FE regressions and two GMM regressions, while the second model will employ five OLS/FE regressions and four GMM regressions. With regard to validity tests for GMM regressions, Tables 3 and 4 will include the two main validity tests, in line with what is stated by Roodman (2009): namely the Hansen test to verify the joint validity of the instruments used in the models, and the Arellano-Bond test for second-order autocorrelation (AR2), to verify that the model errors are not autocorrelated to the second degree (e.g. the current error is correlated with the error from two periods ago). Subsequently, the results obtained will be discussed in section 4.3.

### *3.2.1 Source and structure of variables*

Dreher's (2002) time interval went from 1970 to 2000, and he used five-years averages. In this thesis, the time period will be updated using data that goes from 1992 to 2022, and we will use five-years averages. As for the variables that will be used for regressions, not all of the ones used by Dreher provide a sufficient number of observations or are up to date as of 2022, therefore some of these variables will not be used in the analysis. In particular, rule of law and growth rate of terms of trade will not be included among the variables, as the public available data does not have enough observation or it is not updated to this day and, in the case of growth of terms of trade, the great majority of observations will result in a variance almost near zero. Moreover, regarding the "democracy" variable employed by Dreher, it consists of the Polity-IV-indicator of democracy. To introduce an equivalent variable, it should have used the Polity-V-indicator of democracy, but the latest available data are as of 2018, therefore it has been decided to use data as of 2022 and exclude the variable from the model. It has also been decided to add the variable trade openness, which was not used by Dreher in the model, because from our analysis it emerges as a valid proxy of globalization.

The exclusions help avoid distortions due to series that are too short or almost constant, strengthening the internal validity of the analysis.

As for the source of the variables used, here's how each variable was sourced and the role that it has in the regressions:

- **Log GDPpc End:** Our dependent variable for GMM regressions, is the logarithm of GDP per capita at the End of the period. Following Dreher's approach, we used the logarithm of GDP per capita at the beginning of each five years period as key regressor for the OLS/ FE regressions and the logarithm of GDP per capita at the end of the period as dependent variable for GMM regressions. We sourced GDP from the Penn World Table directly in R through the "pwt10.0" package and manually divided GDP by population and computed the logarithm.
- **GDPpc Growth:** Dependent variable of OLS/FE regressions following Dreher's approach. Real GDP per capita growth rate percentage as a five-years average. We sourced it from the World Bank package directly in R.
- **Log GDPpc Initial:** A control variable, it is the logarithm of GDP per capita at the beginning of the period. For the GMM, we consider instead the logarithm of GDP per capita at the end of the previous period.
- **KOFGI (and its sub-dimensions):** It's the revised version of the KOF Globalization index. Following Dreher's approach, in the first model, we also included three sub-components – economic, social and political sub-dimensions of globalization – and in the second model we included the Economic de facto and trade restriction sub-component of the index. We downloaded the latest revised version, which comprises of all sub-dimensions, as an excel file from the KOF Swiss Economic Institute and manually loaded in R as an object.
- **Investment (as % of GDP):** Fixed investment as a percentage of GDP. We sourced it from the Penn World Table directly in R through the "pwt10.0" package.

- **Gov Consumption (as % of GDP):** Share of public consumption as a percentage of GDP. We sourced it from the Penn World Table directly in R through the “pwt10.0” package.
- **Inflation (as % of GDP):** Percentage inflation rate based on the consumer price index. We sourced it from the World Bank package directly in R.
- **Secondary school enrollment:** Gross enrolment rate in secondary school as a percentage of school-age population. We sourced it from the World Bank package directly in R.
- **Log Fertility:** Logarithm of fertility rate. We sourced it from the World Bank package directly in R.
- **Log Life Expectancy:** Logarithm of life expectancy at birth. We sourced it from the World Bank package directly in R.
- **Pop Growth:** Population growth rate. We sourced it from the World Bank package directly in R.
- **Trade Openness:** Sum of exports and imports of goods and services, as percentage of GDP. We downloaded the excel file from the World Bank website and then manually loaded it into the R environment as an object.
- **PR and CL:** Political Rights and Civil Liberties Indexes. Since the latest available data of the indexes starts from 2014, we downloaded the excel file from the Freedom House website and downloaded the previous version that goes from 1973 to 2013 from the old Freedom House website version found through the Wayback Machine.
- **Dummies time periods:** we excluded them from results in Tables 3 and 4.
- **Restriction (Developed, Developing, Emerging):** Proxies from trade restrictions. We have created three different variables based on a breakdown of income brackets in the various countries. This aspect will be explored in more detail in section 4.2.

## 4. Analysis

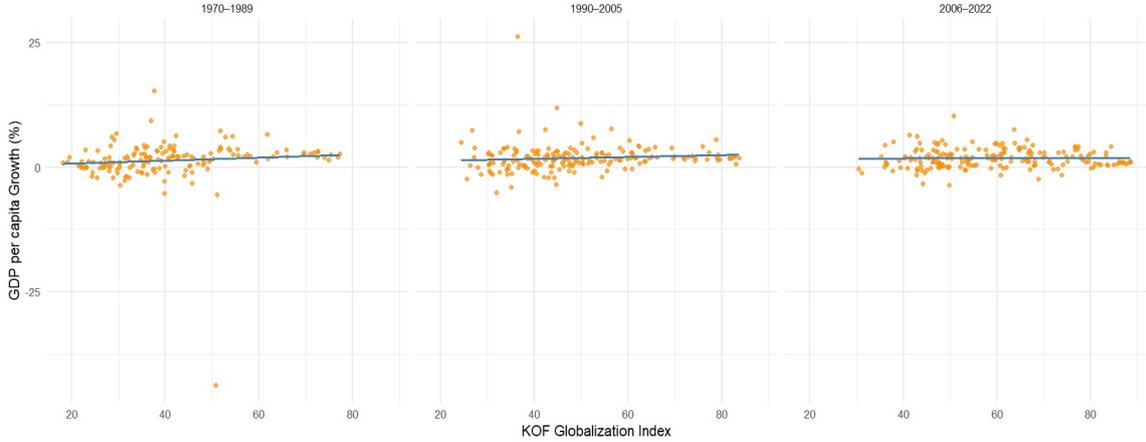
### 4.1 Globalization, trade openness, and economic growth: Correlation Analysis

The two main measures of globalization that emerge from literature have already been identified in previous sections. Now it is time to analyze them in relation to economic growth, measured by GDP per capita growth rate. It has already been seen how the two measures are intertwined: a measure of trade openness is also contained in the latest version of the KOF Index, within the trade component of the economic sub-dimension of globalization. The aim of this chapter is therefore to compare these two measures of globalization in relation to economic growth through a series of scatter plots. Moreover, even though it does not answer the research questions, KOF Index and trade openness will also be directly compared to each other, to see how the relationship changes over time.

The first graph (Figure 5) – a scatter plot measuring the correlation between the KOF Globalization Index and per capita GDP growth – shows that there is essentially no correlation between the two variables. Furthermore, in all three periods, the position of the dots on the y-axis remains essentially unchanged, meaning that per capita GDP growth has not undergone significant changes over the course of 50 years. On the other hand, the KOF values appear to be shifted more to the right in the last period (2006-2022), suggesting a conceivable increase in the index over the last 16 years.

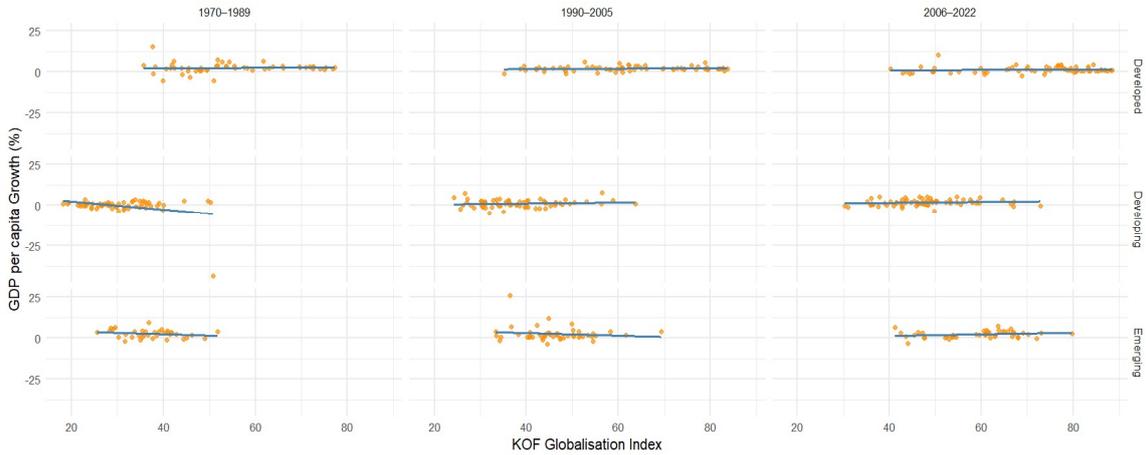
The second series of scatter plots (Figure 6) introduces a new element: the distinction is not only between three periods, but also between countries in different income brackets. Therefore, there are nine scatter plots instead of three, as the countries are now divided into three clusters by income level: developed, developing and emerging. The series of graphs essentially confirm what was shown in Figure 5: a lack of correlation between the two measures in all nine graphs. Indeed, there appears to be a weak negative correlation between the two variables in the first period (1970–1989) for developing countries, but it is so weak that it legitimately raises doubts about the validity of this result.

**Figure 5 - KOF Globalization Index and GDP per capita growth scatter plots**



The scatter plots are divided into three periods: 1970–1989, 1990–2005, and 2006–2022. Each point represents a country in a given period. The line shows the estimated central trend for each time interval.

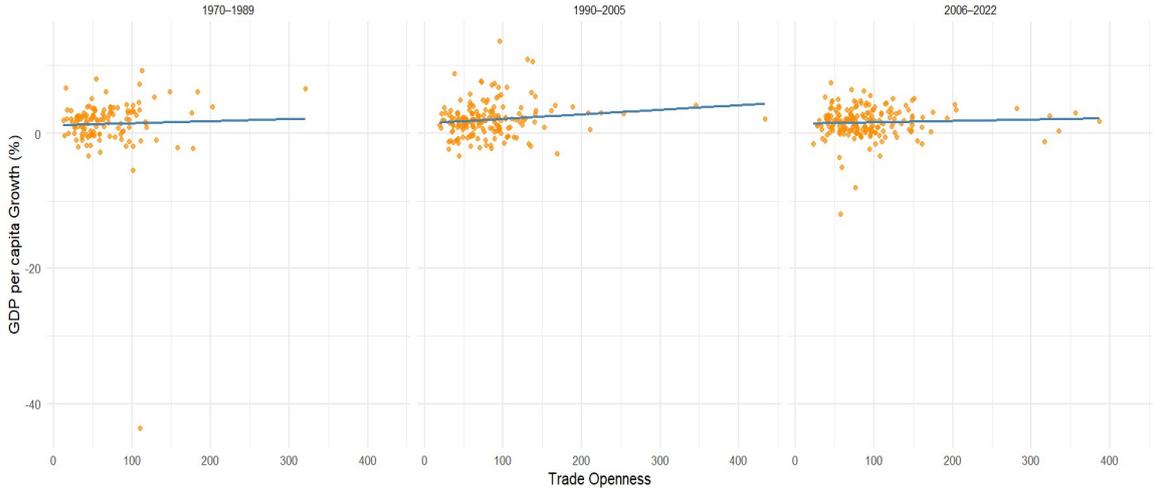
**Figure 6 - KOF Globalization Index and GDP per capita growth scatter plots by income group**



The scatter plots are divided into three periods (1970–1989, 1990–2005, and 2006–2022) and three country groups based on economic development (developed, developing, emerging). Each point represents a country in a given period. The line shows the estimated central trend for each time interval.

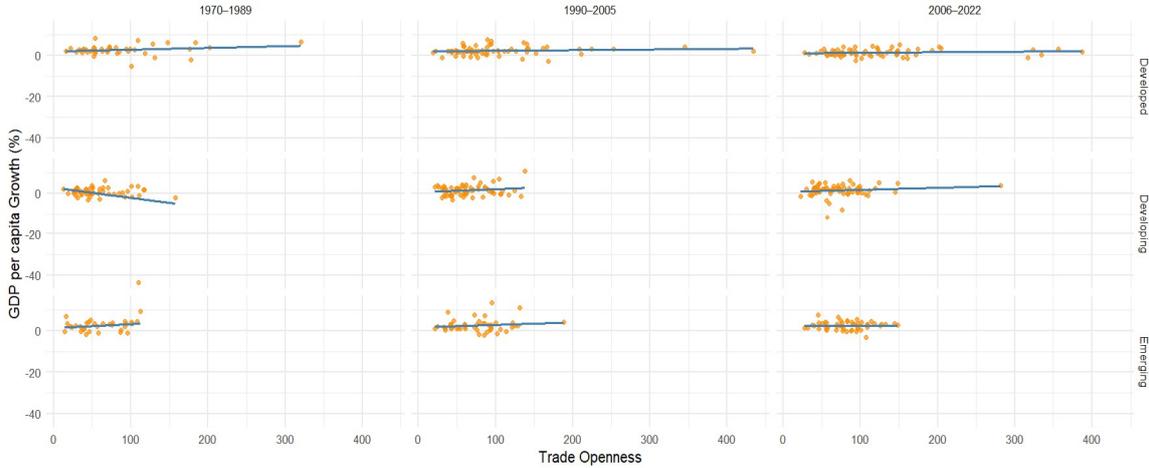
The next series of graphs shown in Figure 7 illustrates the level of correlation between trade openness and per capita GDP growth in the three historical periods already identified. The scatter plots show essentially the same results as Figures 5 and 6: there is no significant correlation between the variables. In Figure 8 - the results do not change: there remains a weak negative correlation in the first period (1970-1989) between the two variables in developing countries, but the correlation is so weak and close to zero that it can be safely said that in no period and in no specific group of countries does there appear to be a significant correlation between the two variables.

**Figure 7 - Trade Openness and GDP per capita growth scatter plots**



*The scatter plots are divided into three periods: 1970–1989, 1990–2005, and 2006–2022. Each point represents a country in a given period. The line shows the estimated central trend for each time interval.*

**Figure 8 - Trade Openness and GDP per capita growth scatter plots by income group**



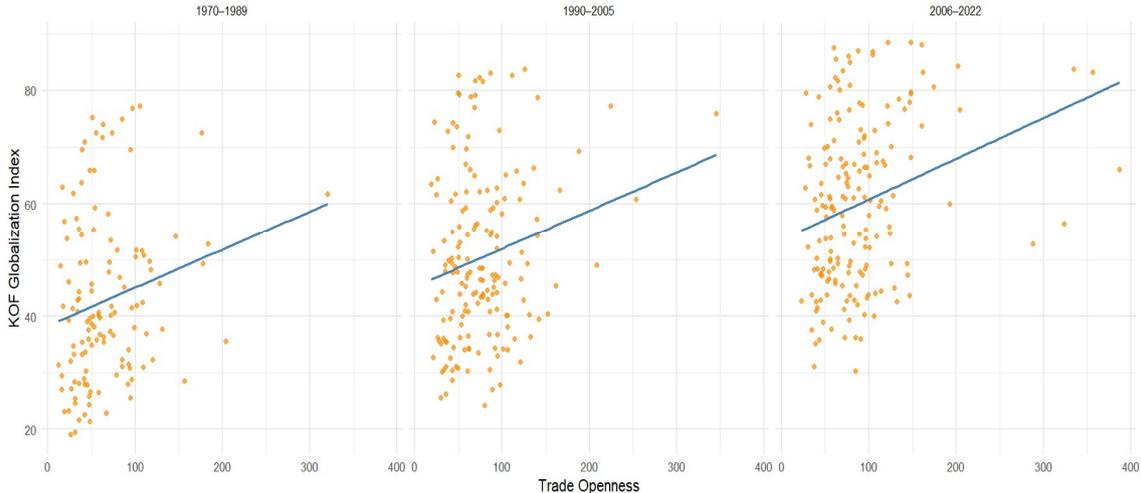
*The scatter plots are divided into three periods (1970–1989, 1990–2005, and 2006–2022) and three country groups based on economic development (developed, developing, emerging). Each point represents a country in a given period. The line shows the estimated central trend for each time interval.*

Figure 9 illustrates the correlation between the KOF Index and trade openness. As can be seen, the Figure shows a clear positive relationship between the two variables in all three periods. While the intensity of the correlation remains virtually unchanged, it can be seen that the average KOF score increases progressively between periods - and this is more noticeable now that the KOF variable is measured on the y-axis. whereas this increase in KOF over the years was more subtle and less visible in the previous scatter plots, when the variable was measured on the x-axis, suggesting that the globalization index in various countries has increased progressively over the decades, in line with the findings in the graphs above. There does not appear to be a similar shift to the right or left of the points on the graph, suggesting that over the decades the value of trade openness has not increased or decreased uniformly in a significant way.

Figure 10 provides more information on the comparison between these two variables than was previously available: it can be seen that as the group of countries changes, so does the sign of the relationship between KOF and trade openness, and the relationship remains

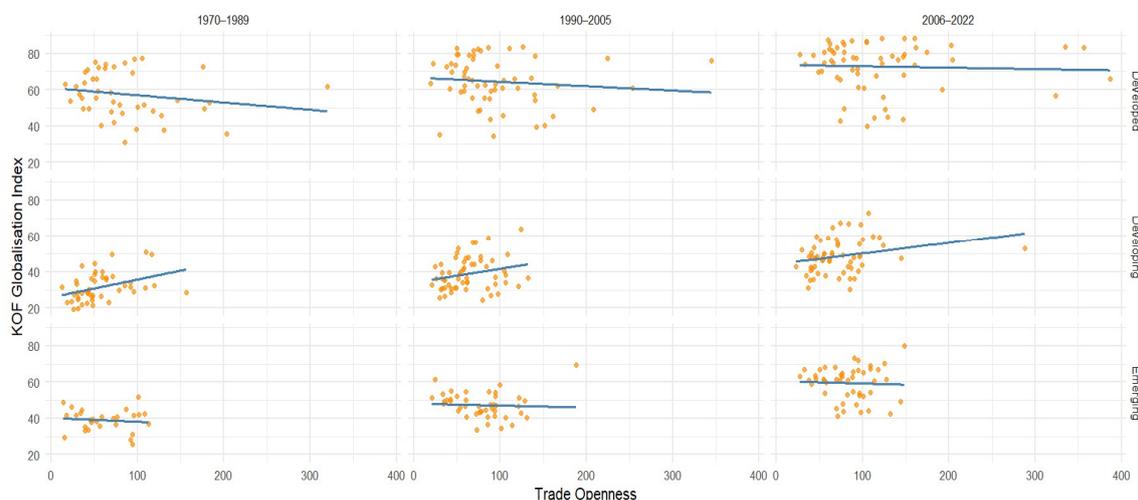
constant over the three periods. In fact, it is possible to see a slightly negative, certainly weak relationship between the two variables in all periods for developed countries, a much stronger and more positive relationship in all three periods for developing countries, and again an almost absent relationship, but slightly negative for so-called emerging countries. This may imply that high trade openness does not necessarily imply a high level of globalization; other factors, such as a country's income level, have a significant influence on whether or not trade openness is adopted, and in the case of developing countries, lower trade openness is more likely at the same level of globalization than in other income groups.

**Figure 9 - Trade openness and KOF Globalization index scatter plots**



*The scatter plots are divided into three periods: 1970–1989, 1990–2005, and 2006–2022. Each point represents a country in a given period. The line shows the estimated central trend for each time interval.*

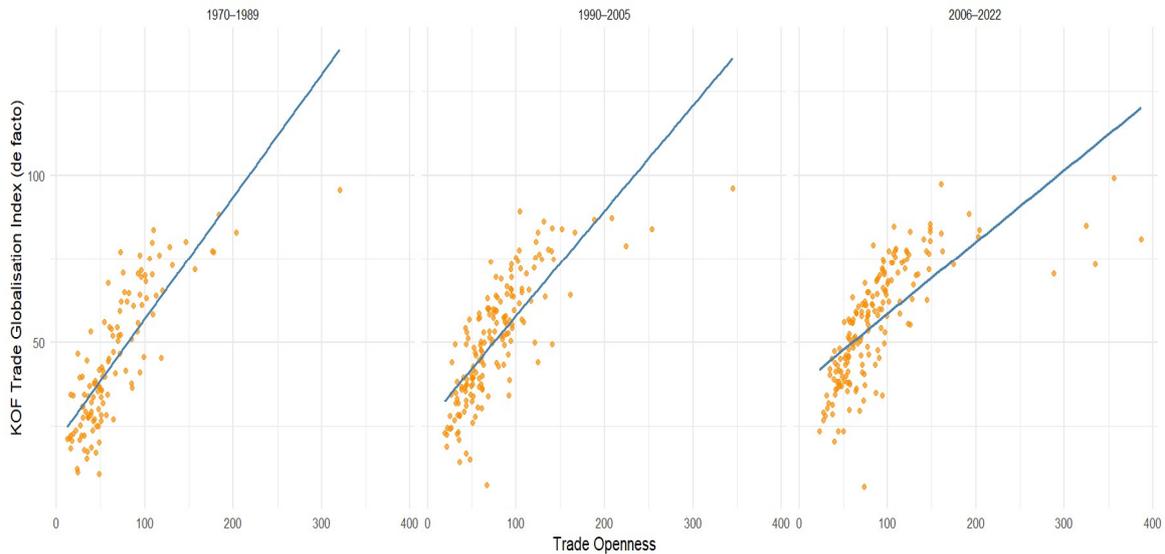
**Figure 10 - Trade openness and KOF Globalization index scatter plots by income group**



*The scatter plots are divided into three periods (1970–1989, 1990–2005, and 2006–2022) and three country groups based on economic development (developed, developing, emerging). Each point represents a country in a given period. The line shows the estimated central trend for each time interval.*

The next series of scatter plots - Figure 11 - compares trade openness with the de facto sub-dimension of economic globalization in the KOF index. We expect a positive relationship between the two variables, as trade openness is one of the components of the KOF's economic Index. The result confirms expectations: in all three periods, there is a strongly positive relationship between the variables, more positive than the relationship between the standard KOF and trade openness. The positive correlation appears to decrease in all three periods: it is therefore higher in the first period than in the third. This difference seems to be due to the progressive increase in the level of globalization measured by the KOF of the various countries: in fact, while it is not possible to note a significant variation in the level of trade openness of countries over the three periods, it is possible to see a “compaction” of the distribution of points along the y-axis (so on average the countries are closer to each other vertically), and this causes the line that directs the correlation to be more horizontal in the third period than in the previous two.

**Figure 11 - Trade openness and KOF Trade de facto Globalization index scatter plots**



*The scatter plots are divided into three periods: 1970–1989, 1990–2005, and 2006–2022. Each point represents a country in a given period. The line shows the estimated central trend for each time interval.*

## 4.2 Globalization, trade openness, and economic growth: GMM Regression

The previous section provided a correlation analysis of the KOF Index and one of its sub-components, trade openness, and economic growth. Now, the focus will shift on the effect of globalization and economic openness on economic growth, the most important analysis that will answer the research questions.

It is also important to mention that, in order to keep the actual analysis of the results separate from the more general discussion, in the first two sections, 4.2.1 and 4.2.2, the results will only be commented on with regard to the papers by Dreher (2002) and eventually Dreher (2008) and Gozgor (2017). In the discussion in section 4.3, however, the discussion will be broadened, and the results will be compared with what was said in the most relevant literature on the subject.

#### *4.2.1 GDP Growth and globalization*

The regressions analyzed in this sub-section, shown in Table 3, are divided into five different columns, three employing OLS/FE regressions and two GMM regressions. Column 1 includes the control variable, the logarithm of per capita GDP at the beginning of the period (at the end of the previous period for GMMs), three endogenous variables (investment, government consumption, logarithm of the fertility rate), and four exogenous variables (population growth, secondary school enrollment, logarithm of life expectancy, and inflation rate). Column 2 adds the KOF Globalization Index, considered an endogenous variable. Column 3 adds the political rights index and the civil liberty index as exogenous variables. Column 4 repeats the regression of Column 2, but this time using the GMM estimate. Column 5 still employs the GMM estimate, but instead of including the KOF Index as an endogenous variable, it includes its three sub-dimensions—economic, political, and social.

In this model we have considered the variables investment, government consumption, the KOF Index and its economic sub-dimension, together with the logarithm of the fertility rate (in line with what was stated by Barro and Lee, 1994) as endogenous. In addition, it was decided to also show the lagged coefficients for every endogenous variable (only for the GMM regressions), which will be indicated by the variable marked with the term “L1.” before its name, while the standard non-lagged variables have the term “L0.” before their name. This decision, in contrast to what Dreher and Gozgor did, in order to is taken to show the delayed effect of each endogenous variable and thus draw more in-depth conclusions about the effects these variables have on the dependent variable.

The overall results show substantial continuity with what was stated in the previous section. In fact, in no regression the main KOF Globalization Index has a significant effect, showing that there are no significant effect of globalization indices in economic growth. However, the influence of the KOF indices on the dependent variable is not zero. If we consider the regression in Column 5, there is a slightly positive and slightly significant effect of KOF’s economic sub-dimension, while the other two sub-dimensions – political and social – show no significant effect. Comparing our results with Dreher’s results, we see that Dreher finds

positive significant correlation with the KOF Index in every regression, both OLS/FEs and GMMs. This is contrary to what our analysis found. He also finds slightly positive significant effect of KOF's economic sub-dimension when testing all three globalization sub-dimensions (economic, political, social), therefore this result remains consistent with what we found.

The control variable "log\_GDPpc\_Initial" — the logarithm of GDP per capita at the beginning of the period — shows consistently significant coefficients, strongly negative in OLS regressions and positive in GMMs with values slightly below 1. These results are entirely consistent with Dreher's findings, suggesting that the higher the initial GDP of a country the lower the GDP will grow (OLS regressions) and the level of GDP at the beginning of the period influences the level of GDP at the end of the period. Gozgor too finds a positive and significant correlation with lagged per capita GDP used as a control variable when the dependent variable is real GDP per capita.

As for the other endogenous variables, both OLS/FEs and GMMs show that investment has a positive and highly significant effect. The OLS coefficients show an exaggeratedly large effect, but even the lower and more plausible GMM coefficients show a coefficient very close to 1. Moreover, government consumption is shown to have a negative effect, which is not significant in OLS/FEs but is otherwise significant in both GMMs, with a p-value < 0.001. Dreher (2002) findings for investment show a positive and significant influence in all regressions, but the coefficients are significantly lower, to the point that there is an even stronger effect of the KOF Index on the dependent variable. Similarly to investment, Dreher finds a much weaker significant negative effect for government consumption. With regard to lagged endogenous variables, an effect opposite to that of the standard variable can be observed for almost all variables: investment has a delayed negative and significant effect, as does the fertility rate; government consumption has a strong positive and significant effect. It is important to note that these results should be treated with caution, as they reflect the dynamic adjustment process and may be sensitive to specification. The lagged coefficients have a negative effect because it represents the adjustment phase. For example, with regard to investment, after a large boom in investment in a given year, the growth rate may slow down the following year, not because investment is harmful, but because the economy is

assimilating the new capital. The negative lagged effect indicates that growth is converging toward its long-term equilibrium rate.

With regard to exogenous variables, inflation is found to have a very weak negative and significant effect. The civil liberties index, on the other hand, has a significant positive relationship, visible in column 3 of Table 3, suggesting a positive relationship between economic growth and the level of civil liberties in a country. The latter results differ from those found by Dreher, who finds no significant effect with the civil liberties index. On the other hand, there does not appear to be a significant constant relationship, either positive or negative, with the other variables considered exogenous, with the exception of the logarithm of life expectancy, whose coefficients are positive and significant only in GMMs. Dreher's findings are consistent with these results.

#### *4.2.2 GDP Growth and economic integration*

The regressions analyzed in this sub-section, shown in Table 4, are divided into nine different columns, five employing the OLS/FE estimate and four the GMM estimate. Column 1 always includes the logarithm of GDP per capita at the beginning of the period (the end of the previous period for GMMs) as a control variable, the economic sub-dimension of the KOF Index, investment, government consumption, and the logarithm of the fertility rate as endogenous variables, and population growth, secondary school enrollment, the logarithm of life expectancy, and the inflation rate as exogenous variables. Column 2 includes the same variables but includes the economic de facto sub-dimension of the KOF Index (instead of the standard economic sub-dimension) as an endogenous variable. Column 3 includes the economic de jure sub-dimension as the only KOF variable (endogenous). Column 4 does not include any KOF variable but includes trade openness as an endogenous variable. Column 5 repeats the regression of Column 1, but this time using the GMM estimate. Column 7 repeats the regression of Column 2, but this time using the GMM estimate. Column 8 includes the KOF economic sub-dimension as an endogenous variable and includes the three “restriction” variables as exogenous variables. Column 9 does not include any KOF variable nor any restriction variable but does include trade openness as an endogenous variable.

The results are consistent with those found in the previous model. Once again, the coefficients of KOF and its sub-dimensions, although positive and significant in Columns 1, 2, 3, and 7, are still too low to suggest a direct influence of globalization and trade liberalization on economic growth. Dreher, on the other hand, found significant results and a higher positive coefficient for the indices, with the exception of the sub-dimension of trade de jure (trade restrictions), whose coefficient is not significant.

With regard to the other endogenous variables, here too are found to have highly significant and positive coefficients for investment and highly significant and negative coefficients for government consumption, with results that are entirely consistent with those of Dreher (although once again he obtained lower coefficients, and they were not significant with regard to government consumption in GMMs). As for fertility rate, we find significant and negative results in OLS/FE regressions and significant and positive results in GMMs. The lagged endogenous variables show identical behavior to that shown in the previous model in Table 3: investment and the fertility rate always have a negative and significant coefficient (with the exception of the fertility rate in Column 9), and government consumption always has a positive and significant coefficient. For an explanation of why this happens, see section 4.2.1 above.

As for the exogenous variables, here too are found consistently weak negative (next to zero) and significant coefficients for inflation, and significant positive results (columns 5, 6, 7) and negative results (columns 8, 9) for the logarithm of life expectancy, and essentially no significant coefficients for the other variables. This also applies to the restriction variables, which are significant in OLS/FEs but not in GMMs, but in any case, their coefficients are close to zero.

**Table 3 - GDP growth and globalization**

	1	2	3	4	5
<b>KOFGI</b>		<b>-0.007</b>	<b>0.013</b>	<b>0.016</b>	
L1.KOFGI				-0.022*	
<b>KOFGI_Econ</b>					<b>0.019+</b>
L1.KOFGI_Econ					-0.018+
<b>KOFGI_Social</b>					<b>0.005</b>
L1.KOFGI_Social					-0.009
<b>KOFGI_Political</b>					<b>-0.017</b>
L1.KOFGI_Political					0.013
<b>log_GDPpc_Initial</b>	<b>-0.943+</b>	<b>-0.879</b>	<b>-0.658</b>	<b>0.912***</b>	<b>0.896***</b>
<b>Investment</b>	<b>7.282***</b>	<b>7.226***</b>	<b>7.837***</b>	<b>0.661+</b>	<b>0.865***</b>
L1.Investment				-0.533*	-0.892***
<b>Gov_Consumption</b>	<b>-0.259</b>	<b>-0.403</b>	<b>0.274</b>	<b>-1.307***</b>	<b>-0.839***</b>
L1.Gov_Consumption				1.136***	0.698**
<b>log_Fertility</b>	<b>-2.408**</b>	<b>-2.406**</b>	<b>-1.865**</b>	<b>0.513+</b>	<b>0.525***</b>
L1.log_Fertility				-0.609**	-0.716***
<b>Pop_Growth</b>	<b>0.172</b>	<b>0.160</b>	<b>-0.076</b>	<b>-0.019+</b>	<b>-0.001</b>
<b>Secondary_School_Enrollment</b>	<b>-0.012</b>	<b>-0.012</b>	<b>-0.004</b>	<b>0.001</b>	<b>0.001</b>
<b>log_Life_Expectancy</b>	<b>0.040</b>	<b>0.081</b>	<b>-0.506</b>	<b>0.286*</b>	<b>0.374*</b>
<b>Inflation</b>	<b>-0.008***</b>	<b>-0.008***</b>	<b>-0.008***</b>	<b>-0.002+</b>	<b>-0.002+</b>
<b>PR</b>			<b>-0.134</b>		
<b>CL</b>			<b>0.674*</b>		
Estimation Method	OLS	OLS	OLS	GMM	GMM
Number of observations	727	724	667	580	580
Number of countries	137	136	125	116	116
R <sup>2</sup>	0.147	0.146	0.159		
Hansen Test (p-level)				0.290	0.322
Arellano-Bond-Test (p-level)				0.127	0.110

*The table is adapted from Dreher (2002). In the OLS regressions, the dependent variable is the average GDP per capita growth rate. When estimated with GMM, the natural logarithm of per capita GDP at the end of each five-year period is employed. A dummy for each time period is included.*

*Significance levels: \*\*\* = p value < 0.001; \*\* = p value < 0.01; \* = p value < 0.05; + = p value < 0.1*

**Table 4 - GDP growth and economic integration**

	1	2	3	4	5	6	7	8	9
<b>KOFGI_Econ</b>	<b>0.042+</b>					<b>0.008</b>			
L1.KOFGI_Econ						-0.007			
<b>KOFGI_EconDeFacto</b>		<b>0.032+</b>					<b>0.011+</b>		
L1.KOFGI_EconDeFacto							-0.011+		
<b>KOFGI_TrRestr</b>			<b>0.023*</b>						
L1.KOFGI_TrRestr									
<b>Trade_Openness</b>					<b>0.006*</b>				<b>0.001</b>
L1.Trade_Openness						-			-0.001
<b>Restriction_Developed</b>				<b>0.002+</b>				<b>&lt; 0.001</b>	
L1.Restrictio_n_Developed								< 0.001	
<b>Restriction_Emerging</b>				<b>0.004**</b>				<b>&lt; 0.001</b>	
L1.Restrictio_n_Emerging								< 0.001	
<b>Restriction_Developing</b>				<b>0.003*</b>				<b>&lt; 0.001</b>	
L1.Restrictio_n_Developing								< 0.001	
<b>log_GDPpc_Initial</b>	<b>-1.261**</b>	<b>-1.161*</b>	<b>-1.696***</b>	<b>-1.586***</b>	<b>-1.707***</b>	<b>0.814***</b>	<b>0.761***</b>	<b>0.676***</b>	<b>0.842***</b>
<b>Investment</b>	<b>6.840**</b>	<b>6.514**</b>	<b>9.666***</b>	<b>9.789***</b>	<b>7.370***</b>	<b>0.962***</b>	<b>1.631***</b>	<b>1.404***</b>	<b>0.185</b>
L1.Investment						-0.496+	-1.732***	-1.341***	0.499**
<b>Gov_Consumption</b>	<b>-0.005</b>	<b>-0.868</b>	<b>-4.554*</b>	<b>-4.571*</b>	<b>-2.231</b>	<b>-1.85***</b>	<b>-1.144***</b>	<b>-1.27***</b>	<b>-1.983***</b>
L1.Gov_Consumption						2.068***	0.934***	1.427***	1.479***
<b>log_Fertility</b>	<b>-2.193**</b>	<b>-2.299**</b>	<b>-1.703**</b>	<b>-1.637*</b>	<b>-1.607*</b>	<b>0.571*</b>	<b>0.553*</b>	<b>0.259**</b>	<b>0.138</b>
L1.log_Fertility						-0.631***	-0.478*	-0.514***	-0.165
<b>Pop_Growth</b>	<b>0.164</b>	<b>0.138</b>	<b>-0.067</b>	<b>-0.057</b>	<b>-0.053</b>	<b>-0.005</b>	<b>-0.001</b>	<b>0.011</b>	<b>-0.012</b>
<b>Secondary_School_Enrollment</b>	<b>-0.016</b>	<b>-0.015</b>	<b>0.000</b>	<b>0.002</b>	<b>0.007</b>	<b>0.002</b>	<b>0.001</b>	<b>&lt; 0.001</b>	<b>&lt; 0.001</b>
<b>log_Life_Expectancy</b>	<b>0.269</b>	<b>0.636</b>	<b>2.645</b>	<b>2.133</b>	<b>5.135**</b>	<b>0.38**</b>	<b>0.225**</b>	<b>-0.387+</b>	<b>-0.368***</b>
<b>Inflation</b>	<b>-0.007***</b>	<b>-0.007***</b>	<b>-0.007***</b>	<b>-0.007***</b>	<b>-0.010***</b>	<b>-0.002</b>	<b>-0.002*</b>	<b>-0.002+</b>	<b>-0.002***</b>
Estimation Method	OLS	OLS	OLS	OLS	OLS	GMM	GMM	GMM	GMM
Number of observations	724	724	689	684	660	595	595	595	580
Number of countries	136	136	127	126	124	119	119	119	116
R <sup>2</sup>	0.158	0.158	0.286	0.294	0.284				
Hansen Test (p-level)						0.322	0.453	0.285	0.382
Arellano-Bond-Test (p-level)						0.100	0.179	0.186	0.201

*The table is adapted from Dreher (2002). In the OLS regressions, the dependent variable is the average GDP per capita growth rate. When estimated with GMM, the natural logarithm of per capita GDP at the end of each five-year period is employed. A dummy for each time period is included.*

*Significance levels: \*\*\* = p value < 0.001; \*\* = p value < 0.01; \* = p value < 0.05; + = p value < 0.1*

### 4.3 Discussion

The results do not show any significant effect of the KOF Indexes and trade openness on the dependent variable. They only show slightly positive weak significant coefficients for the economic dimension of the KOF Index and trade openness. The results also show an effect of fixed investment (strong and positive), government consumption (strong and negative), and inflation rate (very weak and negative) on GDP growth. Furthermore, we see that the initial GDP level is negatively correlated with GDP growth. These results are in line with what was found by Barro and Lee (1994) and Barro (1996), who found a negative correlation between GDP growth and government consumption, inflation, and initial GDP level. These findings have also been confirmed by the subsequent literature (Sturm & De Haan, 2005; Mo, 2007; Bergh and Karlsson, 2010; Adepaju & Ogundunmade, 2019). The literature also identifies fertility rates and secondary school enrollment as determinants of economic growth (e.g. Samini and Jenatabadi 2014; Hanif & Arshed, 2016) although we were unable to find any significant results in this regard, same as Dreher.

The fact that the models have effectively identified determinants of economic growth that are widely recognized in the literature supports the soundness of the methodology adopted. However, what deserves the most attention is the effect of globalization and trade openness on economic growth. With this regard, no particular effect of globalization on economic growth can be identified. However, there are some minimal significant results, such as the significant and positive (but low) coefficient of the economic sub-dimension of globalization, which is in line not only with Dreher but also with a significant part of the literature (Samini and Jenatabadi 2014; Gygli et al., 2018). Equally in line with the literature is the finding that, among the KOF sub-dimensions, the one that achieves significant results on GDP growth is the economic one (Dreher, 2002; Dreher, 2008; Gygli et al., 2018). But all these studies find a stronger and greater effect than the one found in the thesis' results. Therefore, our findings contradict part of the literature. However, similar results have also been found by Bergh and Karlsson (2010) and Villaverde and Maza (2011), while Osterloh (2010) even finds a significant negative influence of globalization on GDP growth. Moreover, Hasan (2019)

finds no significant effect of the three dimensions of globalization of the KOF Index on economic growth on the short run, and Radulović and Kostić (2024) find a significant negative effect of globalization on economic growth and a significant positive effect of the social dimension of the KOF Index, but do not find any significant effect within the economic dimension of KOF.

Similarly, we find no highly positive and significant correlation for trade openness. This contrasts with a whole series of findings in literature (Kim, 2011; Yanikkaya, 2003; Huchet-Bourdon et al., 2017; Yu and Meng, 2023). However, there are conflicting opinions on the real influence of trade openness on economic growth. As already mentioned in section 2.2.2, Dani Rodrik has devoted part of his economic career to this very issue. In his joint work with Francisco Rodriguez, it is stated that “we are in fact skeptical that there is a general, unambiguous relationship between trade openness and growth waiting to be discovered. We suspect that the relationship is a contingent one, dependent on a host of country and external characteristics” (Rodriguez & Rodrik, 1999, p.6). In line with what Rodrik and Rodriguez said, we also extend our assumption about the difficulty of finding a clear and definite correlation with economic growth to trade openness.

In light of what was found and discussed, it is now possible to answer the research questions posed in section 1.2. Globalization and trade openness were not found to have a significant and consistent impact on economic growth, with the exception of a weak effect of the economic dimension of globalization and trade openness. Of the two measures that showed any impact, it appears that the economic dimension of the KOF has a greater influence on the dependent variable than trade openness, but the results do not show a significant prevalence of one measure over the other. This does not mean that the hypothesis that globalization and trade openness influence economic growth is rejected, but rather that studies using panel data from countries over several decades do not always find a significant and sufficiently high effect of the main index of globalization and the main measure of trade openness on GDP growth. Furthermore, it can be seen in the correlation analysis in section 4.1 how the value of the globalization index for various countries has increased progressively over the years, and one consequence of this is that its correlation with trade openness has

apparently declined over the last 16 years. Similarly, it is possible that this average increase, accompanied by a gradual decrease in the variance of the KOF values of various countries, has the effect of showing a lower influence on GDP growth when using more recent data than in previous studies, such as that by Dreher (2002), which uses data from 1970 to 2000 to investigate a period with a undoubtedly higher variance in the distribution of global KOF values.

Essentially, we can summarize our findings with this statement: although economic globalization and trade openness appear to have a slight influence on economic growth, this correlation is too weak to be identified systematically, and there are other factors that have a much greater impact on globalization. We can also add that although our regression models have a solid and proven methodological basis and are sufficiently complex to take into account many facets, from strictly economic factors to civil rights and democratic freedoms, population growth, and education levels, they certainly cannot capture all the complexity and interdependence between the various factors that occur in the real world, and also that it should not be overlooked that a gradual global increase in the value of the globalization index and a decline in the level of globalization of various countries may diminish the impact of the KOF on GDP growth in panel data analyses such as in this thesis.

#### *4.3.1 Limitations and future research*

No study is without limitations, and this thesis represents certainly no exception. This section will outline the main limitations encountered and propose new directions for future research. The first limitation of the study is methodological. In section 1.3, it is mentioned that “two major ways [of structuring the methodology] are consistently used and have a solid background”. This thesis followed one of these two ways, but there is a trade-off in choosing one and not the other. If one chooses the first one, integrating in the methodology second generation panel unit root tests which consider cross-sectional dependence, permitting to verify whether there is cross-sectional dependence among the countries in the panel, i.e., whether shocks or changes in one country also affect the others. The methodology used by Dreher (2002) and Gozgor (2017) does not employ any tests to verify cross-sectional

dependence, but is more cautious with regard to endogeneity issues, which have also been pointed out by Potrafke (2014), among others. This shortcoming is, in any case, a limitation that should be noted. The other important limitation concerns the source of the variables: since it is not possible to obtain all the variables from a single source (apart from those variables that must necessarily be obtained from a specific source, such as the political rights index and the civil liberties index), some of the variables have to be taken from the World Bank and some from the Penn World Table. Since the variables are not obtained from a single source, their values differ from their versions in the different sources, and therefore the continuity and reliability of the analysis is affected. Furthermore, future studies could use this methodological approach and integrate second-generation panel unit root tests that consider cross-sectional dependence, that were omitted from this dissertation in line with the works of Dreher (2002 and 2008) and Gozgor (2017). These are problems that future research will have to overcome if it wants to continue, improve and expand the approach used in this thesis.

## 5. Conclusion

This thesis analyzes the effect that globalization and trade openness have on economic growth, as well as investigating which measure between the KOF Index and trade openness has the greatest effect on economic growth. After identifying the research question and the research gap, the literature is carefully analyzed through a bibliometric analysis, in which a dataset containing hundreds of papers on the subject is analyzed to establish the macro-trends on the subject, and then a qualitative review of the two measures identified—the KOF Index and trade openness—is carried out. Subsequently, the methodology is described, divided into two parts: the first structures a correlation analysis through the use of scatter plots, and the second structures a panel data regression analysis based of Dreher (2002) and updated according to new developments brought by Dreher (2008) and Gozgor (2017) to allow for a more solid and structured analysis to provide answers to the research questions. It is possible to draw a line of continuity between what was found in the comparative analysis and what was found in the regression analysis. What emerged is that it is not possible to identify a significant and constant effect of globalization on economic growth, nor of trade openness on economic growth. It was therefore not possible to establish which one between the KOF Index and trade openness has a greater effect on economic growth. Although no significant and consistent influence was found, this does not mean that no relationship was found at all. In fact, a weak, marginally significant positive relationship was found between economic globalization and economic growth (Table 3, Column 5 and Table 4, Column 7), as well as a slightly positive and marginally significant relationship between trade openness and economic growth (Table 4, Column 5). Significant relationships were also found with a whole series of variables, with results in line with those described in a significant portion of the literature (Sturm & De Haan, 2005; Mo, 2007; Bergh and Karlsson, 2010; Adepoju & Ogundunmade, 2019). Furthermore, correlation analysis using scatter plots should not be underestimated in finding a relationship between the various dimensions of the KOF Index and trade openness. In fact, although it is not part of the research question, this investigation has provided a clearer and more defined view of how the two variables have changed over time, and has allowed us to note a sort of “upward levelling” of the globalization index of

various countries over the last few years, which has also allowed us to propose a reason why the correlation between the KOF Index and economic growth is so weak, in contrast to the findings of other studies (Dreher, 2002; Dreher, 2008; Samini and Jenatabadi 2014; Gygli et al., 2018).

In conclusion, this thesis proposes research questions that have been asked by a series of well-established studies, attempting to answer them through a solid approach based on a mixed technique of literature review and a dual methodology, consisting of both comparative analysis and panel data analysis. It has provided results that have enriched the ongoing literary discussion on one of the most important and influential topics concerning economic globalization, demonstrating that we have not yet arrived at a clear and demonstrable strong positive effect of globalization on economic growth. Nevertheless, the methodology offered opens up new possibilities that can be explored by future research.

## References

- Adepoju, A. A., & Ogundunmade, T. P. (2019). Economic growth and its determinants: A cross-country evidence. *Statistics in Transition New Series*, 20(2), 69–84.  
<https://doi.org/10.21307/stattrans-2019-015>
- Al-Rodhan, N. R., & Stoudmann, G. (2006). Definitions of globalization: A comprehensive overview and a proposed definition. *Program on the geopolitical implications of globalization and transnational security*, 6(1-21).
- Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo Evidence and an application to employment equations. *The Review of Economic Studies*, 58(2), 277. <https://doi.org/10.2307/2297968>
- Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, 68(1), 29–51.  
[https://doi.org/10.1016/0304-4076\(94\)01642-d](https://doi.org/10.1016/0304-4076(94)01642-d)
- Baltagi, B. H. (1998). Panel data methods. In *Handbook of applied economic statistics* (pp. 311-323). CRC Press.
- Barro, R. (1996). *Determinants of Economic Growth: A Cross-Country Empirical Study*.  
<https://doi.org/10.3386/w5698>
- Barro, R. J., & Lee, J.-W. (1994). Sources of economic growth. *Carnegie-Rochester Conference Series on Public Policy*, 40, 1–46. [https://doi.org/10.1016/0167-2231\(94\)90002-7](https://doi.org/10.1016/0167-2231(94)90002-7)
- Baumer, B., Cetinkaya-Rundel, M., Bray, A., Loi, L., & Horton, N. J. (2014). R markdown: Integrating a reproducible analysis tool into introductory statistics. *Technology Innovations in Statistics Education*, 8(1). <https://doi.org/10.5070/t581020118>
- Bergh, A., & Karlsson, M. (2009). Government size and growth: Accounting for Economic Freedom and globalization. *Public Choice*, 142(1–2), 195–213.  
<https://doi.org/10.1007/s11127-009-9484-1>

- Brady, D., Beckfield, J., & Seeleib-Kaiser, M. (2005). Economic globalization and the welfare state in affluent democracies, 1975–2001. *American Sociological Review*, 70(6), 921–948. <https://doi.org/10.1177/000312240507000603>
- Busse, M., & Koeniger, J. (2012). Trade and economic growth: A re-examination of the empirical evidence. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2009939>
- Castillo-Vergara, M., Alvarez-Marin, A., & Placencio-Hidalgo, D. (2018). A bibliometric analysis of creativity in the field of Business Economics. *Journal of Business Research*, 85, 1–9. <https://doi.org/10.1016/j.jbusres.2017.12.011>
- Cipollina, M., De Benedictis, L., & Scibè, E. (2024). Environmental migration? A systematic review and meta-analysis of the literature. *Review of World Economics*, 160(4), 1393–1441. <https://doi.org/10.1007/s10290-024-00529-5>
- Coyne, C. J., & Williamson, C. R. (2009). Trade openness and culture (No. 09-05).
- Coyne, C. J., & Williamson, C. R. (2012). Trade openness and cultural creative destruction. *Journal of Entrepreneurship and Public Policy*, 1(1), 22-49.
- Destek, M. A., & Sinha, A. (2020). Renewable, non-renewable energy consumption, economic growth, trade openness and ecological footprint: Evidence from Organisation for Economic Co-operation and development countries. *Journal of Cleaner Production*, 242, 118537. <https://doi.org/10.1016/j.jclepro.2019.118537>
- Dollar, D. (1992). Outward-oriented developing economies really do grow more rapidly: Evidence from 95 LDCs, 1976-1985. *Economic Development and Cultural Change*, 40(3), 523–544. <https://doi.org/10.1086/451959>
- Dreher, A. (2002). Does globalization affect growth? *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.348860>
- Dreher, A., Gaston, N., & Martens, P. (2008). Measuring globalisation. Gauging its Consequences *Springer, New York*, 25-74.

- Dreher A., Gaston N., Martens P., & Van Boxem L. (2010). Measuring globalization: opening the black box. A critical analysis of globalization indices. *Journal of Globalization Studies*, 1(1), 166-185.
- Edwards, M. S. (2006). Public opinion regarding economic and cultural globalization: Evidence from a cross-national survey. *Review of International Political Economy*, 13(4), 587–608. <https://doi.org/10.1080/09692290600839857>
- Federico, G., & Tena-Junguito, A. (2017). A tale of two globalizations: Gains from trade and openness 1800–2010. *Review of World Economics*, 153(3), 601–626. <https://doi.org/10.1007/s10290-017-0279-z>
- Fujii, E. (2017). What does trade openness measure? *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3043615>
- Gozgor, G. (2017). Robustness of the KOF index of economic globalisation. *The World Economy*, 41(2), 414–430. <https://doi.org/10.1111/twec.12546>
- Gräbner, C., Heimberger, P., Kapeller, J., & Springholz, F. (2020). Understanding economic openness: A review of existing measures. *Review of World Economics*, 157(1), 87–120. <https://doi.org/10.1007/s10290-020-00391-1>
- Grassman, S. (1980). Long-term trends in openness of National Economies. *Oxford Economic Papers*, 32(1), 123–130. <https://doi.org/10.1093/oxfordjournals.oep.a041465>
- Gygli, S., Haelg, F., Potrafke, N., & Sturm, J.-E. (2018). The KOF globalisation index - revisited. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3338784>
- Hanif, N., & Arshed, N. (2016). Relationship between school education and economic growth: SAARC countries. *International Journal of Economics and Financial Issues*, 6(1), 294-300.
- Harrison, A. (1995). Openness and Growth: A Time-Series, Cross-Country Analysis for Developing Countries. <https://doi.org/10.3386/w5221>

- Hasan, M. A. (2019). Does globalization accelerate economic growth? South Asian experience using panel data. *Journal of Economic Structures*, 8(1).  
<https://doi.org/10.1186/s40008-019-0159-x>
- He, Q., Wang, T., Chan, A. P. C., Li, H., & Chen, Y. (2019). Identifying the gaps in Project Success Research. *Engineering, Construction and Architectural Management*, 26(8), 1553–1573. <https://doi.org/10.1108/ecam-04-2018-0181>
- Hout, M. C., Papesh, M. H., & Goldinger, S. D. (2013). Multidimensional scaling. *Wiley Interdisciplinary Reviews: Cognitive Science*, 4(1), 93-103.
- Huchet-Bourdon, M., Le Mouël, C., & Vijil, M. (2017). The relationship between trade openness and economic growth: Some new insights on the openness measurement issue. *The World Economy*, 41(1), 59–76. <https://doi.org/10.1111/twec.12586>
- Kilic, C. (2015). Effects of globalization on economic growth: Panel data analysis for developing countries. *Petroleum-Gas University of Ploiesti Bulletin, Technical Series*, 67(1).
- Kim, D.-H. (2011). Trade, growth and income. *The Journal of International Trade & Economic Development*, 20(5), 677–709. <https://doi.org/10.1080/09638199.2011.538966>
- Lloyd, P. J., & MacLaren, D. (2002). Measures of trade openness using CGE analysis. *Journal of Policy Modeling*, 24(1), 67–81. [https://doi.org/10.1016/s0161-8938\(01\)00096-5](https://doi.org/10.1016/s0161-8938(01)00096-5)
- Lockwood, Ben. (2004). How robust is the kearney/foreign policy globalisation index? *World Economy*, 27(4), 507–523.
- Lutz, M., & Singer, H. W. (1994). The link between increased trade openness and the terms of trade: An empirical investigation. *World Development*, 22(11), 1697–1709.  
[https://doi.org/10.1016/0305-750x\(94\)90001-9](https://doi.org/10.1016/0305-750x(94)90001-9)
- Martens, P., & Zywiets, D. (2006). Rethinking globalization: A modified globalization index. *Journal of International Development*, 18(3), 331–350.  
<https://doi.org/10.1002/jid.1223>

Merigó, J. M., Rocafort, A., & Aznar-Alarcón, J. P. (2016). Bibliometric overview of Business & Economics Research. *Journal of Business Economics and Management*, 17(3), 397–413. <https://doi.org/10.3846/16111699.2013.807868>

Mo, P. H. (2007). Government expenditures and economic growth: The supply and demand sides\*. *Fiscal Studies*, 28(4), 497–522. <https://doi.org/10.1111/j.1475-5890.2007.00065.x>

Norris, P. (2000). Global governance and cosmopolitan citizens. *Governance in a globalizing world*, 155, 173-175.

Osterloh, S. (2010). Words speak louder than actions: The impact of politics on Economic Performance. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.1739743>

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews.

Palan, N., Simoes, N., & Crespo, N. (2020). Measuring Fifty Years of Trade globalisation. *The World Economy*, 44(6), 1859–1884. <https://doi.org/10.1111/twec.13016>

Pesaran, M. H., Shin, Y., & Smith, R. P. (1999). Pooled mean group estimation of dynamic heterogeneous panels. *Journal of the American statistical Association*, 94(446), 621-634.

Pesaran, M. H. (2007). A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics*, 22(2), 265–312. <https://doi.org/10.1002/jae.951>

Potrafke, N. (2014). The evidence on globalization. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2425513>

Radulović, M., & Kostić, M. (2024). Globalization and economic growth in Western Balkan countries. *Regional Science Policy & Practice*, 16(8), 100031. <https://doi.org/10.1016/j.rspp.2024.100031>

Rodriguez, F., & Rodrik, D. (1999). Trade Policy and Economic Growth: A Skeptic's Guide to Cross-National Evidence. <https://doi.org/10.3386/w7081>

Rodrik, D. (1998). *Why Do More Open Economies Have Bigger Governments?*

<https://doi.org/10.3386/w5537>

Roodman, D. (2009). How to do xtabond2: An introduction to difference and system GMM in stata. *The Stata Journal: Promoting Communications on Statistics and Stata*, 9(1), 86–136. <https://doi.org/10.1177/1536867x0900900106>

Samimi, P., & Jenatabadi, H. S. (2014). Globalization and economic growth: Empirical evidence on the role of complementarities. *PLoS ONE*, 9(4).

<https://doi.org/10.1371/journal.pone.0087824>

Scholte, J.A., (2002) *What is globalization? The definitional issue - again*. Working Paper. Coventry: University of Warwick. Centre for the Study of Globalisation and Regionalisation. (No.109).

Sturm, J.-E., & de Haan, J. (2005). Determinants of long-term growth: New results applying robust estimation and extreme bounds analysis. *Empirical Economics*, 30(3), 597–617. <https://doi.org/10.1007/s00181-005-0252-x>

Squalli, J., & Wilson, K. (2011). A new measure of trade openness. *The World Economy*, 34(10), 1745–1770. <https://doi.org/10.1111/j.1467-9701.2011.01404.x>

Villaverde, J., & Maza, A. (2011). Globalisation, growth and Convergence. *The World Economy*, 34(6), 952–971. <https://doi.org/10.1111/j.1467-9701.2011.01359.x>

Yanikkaya, H. (2003). Trade openness and economic growth: A cross-country empirical investigation. *Journal of Development Economics*, 72(1), 57–89.

[https://doi.org/10.1016/s0304-3878\(03\)00068-3](https://doi.org/10.1016/s0304-3878(03)00068-3)

## Appendix I: Detailed process of Literature review

This appendix will explain, in greater detail and depth than in Chapter 2, the entire process that led to the production of the outputs analyzed in the quantitative review of the literature. The backbone of this thesis is an R Markdown document. R Markdown is an “easy-to-use system that enables students to combine statistical computing in an environment of their choosing and written analysis in one document. At a high-level, it renders a well-annotated R script into a self-contained HTML file, replete with graphics, commands, and stylized text” (Baumer et al, 2014, p. 5).

The first step consists of loading the relevant packages, which will be done on the markdown in the following strings. The next steps will be unveiled starting from the next chapter.

```
library(pwt10)
library(tidyverse)
library(stringr)
library(dplyr)
library(tidyr)
library(purrr)
library(textstem)
library(visNetwork)
library(ggplot2)
library(ggrepel)
library(stringdist)
library(textstem)
library(fuzzyjoin)
library(readxl)
library(tinytable)
library(WDI)
library(plm)
library(modelsummary)
library(knitr)
library(wbstats)
library(broom)
library(countrycode)
library(car)
library(webshot2)
```

The first step of the literature review consists of the identification of the key papers that must be included in the analysis, and from there the creation of the data-set of papers following a quantitative approach. The key papers identified are four:

- Robustness of the KOF index of economic globalisation, Gozgor, 2017

- Understanding economic openness: A review of existing measures, Gräbner et al., 2020
- The KOF globalisation index - revisited, Gygli et al., 2018
- Measuring Fifty Years of Trade globalisation, Palan et al., 2020

```
core_papers <- read.csv("core_papers.csv", stringsAsFactors = FALSE,
colClasses = "character")
```

All the references of each paper have been then exported from Scopus in four different csv files – named respectively scopus1, scopus2, scopus3 and scopus4 – comprehensive of all data that Scopus allows to export. The files have been then unified in a single object in R through the code that will be running after this sentence.

```
# Loading the 4 files containing the papers used as references
scopus1 <- read.csv("scopus1.csv", stringsAsFactors = FALSE, colClasses =
"character")
scopus2 <- read.csv("scopus2.csv", stringsAsFactors = FALSE, colClasses =
"character")
scopus3 <- read.csv("scopus3.csv", stringsAsFactors = FALSE, colClasses =
"character")
scopus4 <- read.csv("scopus4.csv", stringsAsFactors = FALSE, colClasses =
"character")

# Unifying all files in one object
scopus_data <- bind_rows(scopus1, scopus2, scopus3, scopus4, core_papers)
```

The total amount of papers (observations) of the object “scopus\_data” is 189. The sample was then enriched with the papers that cited the four core papers, in such a way as to include all subsequent literature that has used and applied the index and composite indicators presented in the core papers. In order to do that, we exported from Scopus the citations of each paper in four different csv files, loaded in our working environment.

```
### Loading the csv files exported from Scopus
scopus_openness <- read.csv("scopus51.csv", stringsAsFactors = FALSE,
colClasses = "character")
Gozgor <- read.csv("scopusA.csv", stringsAsFactors = FALSE, colClasses =
"character")
Grabner <- read.csv("scopusB.csv", stringsAsFactors = FALSE, colClasses =
"character")
Gygli <- read.csv("scopusC.csv", stringsAsFactors = FALSE, colClasses =
"character")
Palan <- read.csv("scopusD.csv", stringsAsFactors = FALSE, colClasses =
"character")
```

We also decided to include papers from one last research on Scopus, it consisted of several stages, where the papers were progressively skimmed, until the final 51 papers were identified as relevant. We tried to identify the papers that contain the following keywords:

- Globalization
- Economic openness
- Trade openness
- Financial openness
- Composite indicators

During the research, we tried to include at least one keyword among “trade openness,” “financial openness,” or “economic openness”, at least one among “economic growth,” “GDP growth,” or “macroeconomic indicators”, and at least one among “composite indicators,” “index-based measurement,” or “economic performance”:

```
( TITLE-ABS-KEY ( "trade openness" OR "financial openness" OR "economic openness" ) AND TITLE-ABS-KEY ( "economic growth" OR "GDP growth" OR "macroeconomic indicators" ) AND TITLE-ABS-KEY ( "composite indicators" OR "index-based measurement" OR "economic performance" ) )
```

The search returned 68 results, but some of them were related to sustainability issues.

Therefore, the query was modified once again to filter these sustainability-related papers:

```
(TITLE-ABS-KEY("trade openness" OR "financial openness" OR "economic openness") AND TITLE-ABS-KEY("economic growth" OR "GDP growth" OR "macroeconomic indicators") AND TITLE-ABS-KEY("composite indicators" OR "index-based measurement" OR "economic performance")) AND NOT TITLE-ABS-KEY("climate" OR "carbon" OR "sustainability" OR "environment" OR "china" OR "renewable")
```

We then loaded the exported csv file in our working environment.

```
### Loading the csv file exported from Scopus
scopus_openness <- read.csv("scopus52.csv", stringsAsFactors = FALSE)
```

The last step consists of integrating the top 50 papers in terms of the number of citations in Scopus when searching the following prompt in the title research in Scopus:

```
globalization OR globalisation
```

The reason for this is to further include the most influential articles on the subject of globalization, in order to ensure that we do not leave out the most influential literature on the subject. We then unify the different data-sets in a single object which was called "scopus\_data2".

```
### Loading the csv file exported from Scopus
scopus_globalization <- read.csv("scopusGlobalization.csv",
stringsAsFactors = FALSE)

# Ensuring that the Year column in all tables is of the same type
scopus_data <- scopus_data %>% mutate(Year = as.character(Year))
scopus_openness <- scopus_openness %>% mutate(Year = as.character(Year))
Gozgor <- Gozgor %>% mutate(Year = as.character(Year))
Grabner <- Grabner %>% mutate(Year = as.character(Year))
Gygli <- Gygli %>% mutate(Year = as.character(Year))
Palan <- Palan %>% mutate(Year = as.character(Year))
scopus_globalization <- scopus_globalization %>% mutate(Year =
as.character(Year))

# Making sure the Volume column in all tables is of the same type
scopus_data <- scopus_data %>% mutate(Volume = as.character(Volume))
scopus_openness <- scopus_openness %>% mutate(Volume =
as.character(Volume))
Gozgor <- Gozgor %>% mutate(Volume = as.character(Volume))
Grabner <- Grabner %>% mutate(Volume = as.character(Volume))
Gygli <- Gygli %>% mutate(Volume = as.character(Volume))
Palan <- Palan %>% mutate(Volume = as.character(Volume))
scopus_globalization <- scopus_globalization %>% mutate(Volume =
as.character(Volume))

# Making sure the Art..No. column in all tables is of the same type
scopus_data <- scopus_data %>% mutate(Art..No. = as.character(Art..No.))
scopus_openness <- scopus_openness %>% mutate(Art..No. =
as.character(Art..No.))
Gozgor <- Gozgor %>% mutate(Art..No. = as.character(Art..No.))
Grabner <- Grabner %>% mutate(Art..No. = as.character(Art..No.))
Gygli <- Gygli %>% mutate(Art..No. = as.character(Art..No.))
Palan <- Palan %>% mutate(Art..No. = as.character(Art..No.))
scopus_globalization <- scopus_globalization %>% mutate(Art..No. =
as.character(Art..No.))

# Making sure the Page.start column in all tables is of the same type
scopus_data <- scopus_data %>% mutate(Page.start =
as.character(Page.start))
scopus_openness <- scopus_openness %>% mutate(Page.start =
as.character(Page.start))
Gozgor <- Gozgor %>% mutate(Page.start = as.character(Page.start))
Grabner <- Grabner %>% mutate(Page.start = as.character(Page.start))
Gygli <- Gygli %>% mutate(Page.start = as.character(Page.start))
Palan <- Palan %>% mutate(Page.start = as.character(Page.start))
```

```

scopus_globalization <- scopus_globalization %>% mutate(Page.start =
as.character(Page.start))

# Making sure the Page.end column in all tables is of the same type
scopus_data <- scopus_data %>% mutate(Page.end = as.character(Page.end))
scopus_openness <- scopus_openness %>% mutate(Page.end =
as.character(Page.end))
Gozgor <- Gozgor %>% mutate(Page.end = as.character(Page.end))
Grabner <- Grabner %>% mutate(Page.end = as.character(Page.end))
Gygli <- Gygli %>% mutate(Page.end = as.character(Page.end))
Palan <- Palan %>% mutate(Page.end = as.character(Page.end))
scopus_globalization <- scopus_globalization %>% mutate(Page.end =
as.character(Page.end))

# Making sure the Page.count column in all tables is of the same type
scopus_data <- scopus_data %>% mutate(Page.count =
as.character(Page.count))
scopus_openness <- scopus_openness %>% mutate(Page.count =
as.character(Page.count))
Gozgor <- Gozgor %>% mutate(Page.count = as.character(Page.count))
Grabner <- Grabner %>% mutate(Page.count = as.character(Page.count))
Gygli <- Gygli %>% mutate(Page.count = as.character(Page.count))
Palan <- Palan %>% mutate(Page.count = as.character(Page.count))
scopus_globalization <- scopus_globalization %>% mutate(Page.count =
as.character(Page.count))

# Making sure the Cited.by column in all tables is of the same type
scopus_data <- scopus_data %>% mutate(Cited.by = as.character(Cited.by))
scopus_openness <- scopus_openness %>% mutate(Cited.by =
as.character(Cited.by))
Gozgor <- Gozgor %>% mutate(Cited.by = as.character(Cited.by))
Grabner <- Grabner %>% mutate(Cited.by = as.character(Cited.by))
Gygli <- Gygli %>% mutate(Cited.by = as.character(Cited.by))
Palan <- Palan %>% mutate(Cited.by = as.character(Cited.by))
scopus_globalization <- scopus_globalization %>% mutate(Cited.by =
as.character(Cited.by))

# Making sure the Conference.code column in all tables is of the same
type
scopus_data <- scopus_data %>% mutate(Conference.code =
as.character(Conference.code))
scopus_openness <- scopus_openness %>% mutate(Conference.code =
as.character(Conference.code))
Gozgor <- Gozgor %>% mutate(Conference.code =
as.character(Conference.code))
Grabner <- Grabner %>% mutate(Conference.code =
as.character(Conference.code))
Gygli <- Gygli %>% mutate(Conference.code =
as.character(Conference.code))
Palan <- Palan %>% mutate(Conference.code =
as.character(Conference.code))
scopus_globalization <- scopus_globalization %>% mutate(Conference.code =
as.character(Conference.code))

```

```

# Making sure the Issue column in all tables is of the same type
scopus_data <- scopus_data %>% mutate(Issue = as.character(Issue))
scopus_openness <- scopus_openness %>% mutate(Issue =
as.character(Issue))
Gozgor <- Gozgor %>% mutate(Issue = as.character(Issue))
Grabner <- Grabner %>% mutate(Issue = as.character(Issue))
Gygli <- Gygli %>% mutate(Issue = as.character(Issue))
Palan <- Palan %>% mutate(Issue = as.character(Issue))
scopus_globalization <- scopus_globalization %>% mutate(Issue =
as.character(Issue))

# Making sure the ISSN column in all tables is of the same type
scopus_data <- scopus_data %>% mutate(ISSN = as.character(ISSN))
scopus_openness <- scopus_openness %>% mutate(ISSN = as.character(ISSN))
Gozgor <- Gozgor %>% mutate(ISSN = as.character(ISSN))
Grabner <- Grabner %>% mutate(ISSN = as.character(ISSN))
Gygli <- Gygli %>% mutate(ISSN = as.character(ISSN))
Palan <- Palan %>% mutate(ISSN = as.character(ISSN))
scopus_globalization <- scopus_globalization %>% mutate(ISSN =
as.character(ISSN))

# Making sure the ISSN column in all tables is of the same type
scopus_data <- scopus_data %>% mutate(PubMed.ID =
as.character(PubMed.ID))
scopus_openness <- scopus_openness %>% mutate(PubMed.ID =
as.character(PubMed.ID))
Gozgor <- Gozgor %>% mutate(PubMed.ID = as.character(PubMed.ID))
Grabner <- Grabner %>% mutate(PubMed.ID = as.character(PubMed.ID))
Gygli <- Gygli %>% mutate(PubMed.ID = as.character(PubMed.ID))
Palan <- Palan %>% mutate(PubMed.ID = as.character(PubMed.ID))
scopus_globalization <- scopus_globalization %>% mutate(PubMed.ID =
as.character(PubMed.ID))

### Unifing the two files in a single object
scopus_data2 <- bind_rows(scopus_data, scopus_openness, Gozgor, Grabner,
Gygli, Palan)
scopus_data2 <- scopus_data2 %>%
  distinct(Title, .keep_all = TRUE)
scopus_data2 <- scopus_data2 %>%
  mutate(Title = if_else(Title == "[No title available]", Source.title,
Title))

### Removing duplicates
scopus_data2 <- scopus_data2 %>%
  mutate(Title_clean = Title %>%
    str_to_lower() %>%
    str_replace_all("[[:punct:]]", "") %>%
    str_squish())
scopus_data2 <- scopus_data2 %>%
  distinct(Title_clean, .keep_all = TRUE)
scopus_data2 <- scopus_data2 %>%
  distinct(Title_clean, Year, Authors, .keep_all = TRUE)
scopus_data2 <- scopus_data2 %>%
  select(-Title_clean)

```

Then, the data-set is filtered in order to avoid repetitions and avoid papers that focus on different topics from our research scope.

```
# List of keywords to be excluded
exclude_keywords <- c("carbon dioxide", "renewable energy",
"sustainability",
"climate change", "greenhouse gases", "gender
relation", "wind", "biofuel",
"natural resource", "co2 emission", "environment",
"urbanization",
"alternative energy", "environmental", "energy",
"covid-19", "income inequality",
"human", "male", "female")

# Filtering data-set removing these keywords
scopus_data2 <- scopus_data2 %>%
  filter(!str_detect(tolower(Title), paste(exclude_keywords, collapse =
"|")) &
!str_detect(tolower(Abstract), paste(exclude_keywords,
collapse = "|")) &
!str_detect(tolower(Author.Keywords),
paste(exclude_keywords, collapse = "|")) &
!str_detect(tolower(Index.Keywords),
paste(exclude_keywords, collapse = "|")))

```

Then, the next step is to include the object “scopus-globalization” in our data-set.

```
### Unifing the two files in a single object
scopus_data2 <- bind_rows(scopus_data2, scopus_globalization)

```

The data-set now consists of 655 papers. The last step is to manually remove those papers whose scope of research is not suitable for our analysis, but that were nonetheless included in the data-set until now. To do that, since manually check the full text of 655 papers would be too time-consuming and inefficient for the purpose of this dissertation, a csv file exported directly from R was prepared, containing all the columns of the “scopus\_data2” object, corresponding to our dataset of papers. The csv was then attached to a new chat on ChatGPT, to which the following prompts were provided:

```
I am a researcher in international economics. I am writing a thesis on
trade openness and globalisation. I will describe it step by step. The
first objective of my thesis is to do a literature review on the topic. I
identified four core papers, namely: * Robustness of the KOF index of
economic globalisation, Gozgor, 2017
* Understanding economic openness: A review of existing measures, Gräbner
et al., 2020

```

\* The KOF globalisation index - revisited, Gygli et al., 2018  
\* Measuring Fifty Years of Trade globalisation, Palan et al., 2020  
were the first to be included in the dataset. After that, I imported from scopus both the reference and the papers that cited these four core papers. Then, I integrated with other papers found on scopus obtained by typing the following search(...). Finally, I supplemented the dataset by entering the top 50 papers on scopus by number of citations by typing the following search(...). Screening was done by removing duplicates and papers related to sustainability topics that do not interest me using the following code(...)

The dataset I have given you is the one filtered by screening, all the papers have passed the step and are at the eligibility stage. I have no criteria regarding year of publication, country of publication and author. you know the prism graph? If I send you a csv with the data of the papers in my dataset, would you be able to help me at the eligibility stage?

ChatGPT then evaluated which papers should be included in the analysis, which papers should be excluded, and which papers are in doubt. It proceeded to create an excel file with the same columns and rows of the previously mentioned csv file, with two new columns being added: one column specifies ChatGPT's evaluation of eligibility of each paper, using three different words: Included, Excluded and Maybe. The second column provides the reason why certain papers should be excluded or are in doubt. I mostly considered the first column in the analysis. First of all, I highlighted in green all the "Included", in red all the "Excluded" and in yellow all the "Maybe". All the green "Included" person were not to be analyzed. I manually assessed the eligibility of all papers that were flagged as "Maybe" and "Excluded" considering the title, the abstract and the keywords. The titles of the papers that did not pass the eligibility phase were then included in a new excel file, with only one column "Title". The creation of the excel file was aimed at removing the above-mentioned paper from the data set, which will be run on the markdown after this sentence.

```
# 1. Loading non eligible papers from the excel file
non_eligible <- read_excel("non_eligible_titles.xlsx")

# 2. Cleaning of papers' titles
non_eligible <- non_eligible %>%
  mutate(Title_clean = str_to_lower(Title),
         Title_clean = str_replace_all(Title_clean, "[[:punct:]]", ""),
         Title_clean = str_squish(Title_clean))

# 3. Applying same cleaning to our data-set
scopus_data2_tmp <- scopus_data2 %>%
  mutate(Title_clean = str_to_lower(Title),
         Title_clean = str_replace_all(Title_clean, "[[:punct:]]", ""))
```

```
Title_clean = str_squish(Title_clean)

# 4. Removing non-eligible papers
scopus_data2_filtered <- scopus_data2_tmp %>%
  filter(!Title_clean %in% non_eligible$Title_clean) %>%
  select(-Title_clean) # rimuove colonna temporanea

# 5. Overwrite original dataset
scopus_data2 <- scopus_data2_filtered
```

The final data-set consists of 543 papers.

## Appendix II: Code of Methodology and Analysis

```
# Loading World Bank file (Trade Openness)
wb <- read_excel("API_NE.TRD.GNFS.ZS_DS2_en_excel_v2_327564.xls", skip =
3)

# Removing columns "2023" and "2024"
wb <- wb %>%
  select(-`2023`, -`2024`)

# Transforming in long format
wb_long <- wb %>%
  pivot_longer(cols = `1970`:`2022`, names_to = "year", values_to =
"Trade_Openness") %>%
  select(country = `Country Name`, year, Trade_Openness) %>%
  mutate(year = as.integer(year))

# Loading KOF file
kof <- read_excel("KOFGI_2024_public.xlsx")

# Selecting the relevant columns (Country, Year, KOFGI)
kof_clean <- kof %>%
  select(country, year, KOFGI) %>%
  mutate(year = as.integer(year))

merged <- left_join(kof_clean, wb_long, by = c("country", "year"))

# Creating historical periods
merged <- merged %>%
  mutate(Period = case_when(
    year >= 1970 & year <= 1989 ~ "1970-1989",
    year >= 1990 & year <= 2005 ~ "1990-2005",
    year >= 2006 & year <= 2022 ~ "2006-2022",
    TRUE ~ NA_character_
  )) %>%
  filter(!is.na(Period)) # Elimina eventuali righe fuori range

merged_clean <- merged %>%
  filter(!is.na(KOFGI), !is.na(Trade_Openness))

# Reassigning periods
merged_clean <- merged_clean %>%
  mutate(Period = case_when(
    year >= 1970 & year <= 1989 ~ "1970-1989",
    year >= 1990 & year <= 2005 ~ "1990-2005",
    year >= 2006 & year <= 2022 ~ "2006-2022",
    TRUE ~ NA_character_
  )) %>%
  filter(!is.na(Period))

# Calculating averages by country and period
period_means <- merged_clean %>%
```

```

group_by(country, Period) %>%
summarise(
  Mean_KOF = mean(KOFGI, na.rm = TRUE),
  Mean_Trade = mean(Trade_Openness, na.rm = TRUE),
  .groups = "drop"
)

# Scatter plots with linear relationship
ggplot(period_means, aes(x = Mean_Trade, y = Mean_KOF)) +
  geom_point(alpha = 0.7, color = "darkorange") +
  geom_smooth(method = "lm", color = "steelblue", se = FALSE) +
  facet_wrap(~Period) +
  labs(
    title = "Average Trade Openness vs KOF Index by Country and Period",
    x = "Average Trade Openness",
    y = "Average KOF Globalization Index"
  ) +
  theme_minimal(base_size = 13)

# Obtaining country classifications from the World Bank
wb_meta <- wb_countries()
income_groups <- wb_meta %>%
  select(Country = country, iso2c, Income_Group = income_level) %>%
  mutate(Group = case_when(
    Income_Group == "High income" ~ "Developed",
    Income_Group == "Upper middle income" ~ "Emerging",
    Income_Group %in% c("Lower middle income", "Low income") ~
"Developing",
    TRUE ~ NA_character_
  )) %>%
  filter(!is.na(Group))

# Creating a temporary version of merged_clean without the Group column
# Performing the select(-Group) operation only on a temporary object
# Leaving "merged_clean" intact within the pipe
if (!"Country" %in% colnames(merged_clean) && "country" %in%
colnames(merged_clean)) {
  merged_clean <- merged_clean %>% rename(Country = country)
}

# Now joining
plot_data <- merged_clean %>%
  select(-any_of("Group")) %>%
  left_join(income_groups %>% select(Country, Group), by = "Country") %>%
  filter(!is.na(Group))

# Calculating averages by country, period and group
period_means_grouped <- plot_data %>%
  group_by(Country, Period, Group) %>%
  summarise(
    Mean_KOF = mean(KOFGI, na.rm = TRUE),
    Mean_Trade = mean(Trade_Openness, na.rm = TRUE),
    .groups = "drop"
  )

```

```

# Creating the 9 scatter plots
ggplot(period_means_grouped, aes(x = Mean_Trade, y = Mean_KOF)) +
  geom_point(alpha = 0.7, color = "darkorange") +
  geom_smooth(method = "lm", se = FALSE, color = "steelblue") +
  facet_grid(Group ~ Period) +
  labs(
    title = "Trade Openness vs KOF Globalisation Index\nby Income Group
and Period",
    x = "Average Trade Openness (World Bank)",
    y = "Average KOF Globalisation Index"
  ) +
  theme_minimal(base_size = 13)

# =====
# KOF Globalization index vs GDP per capita growth
# =====
wdi_growth <- WDI::WDI(
  indicator = c(GDPpc_Growth = "NY.GDP.PCAP.KD.ZG"),
  start = 1970,
  end = 2022
) %>%
  rename(country = country, year = year) %>%
  mutate(year = as.integer(year)) %>%
  select(country, year, GDPpc_Growth)

# Joining with KOF
merged_growth <- left_join(kof_clean, wdi_growth, by = c("country",
"year"))

# Creating historical periods
merged_growth <- merged_growth %>%
  mutate(Period = case_when(
    year >= 1970 & year <= 1989 ~ "1970-1989",
    year >= 1990 & year <= 2005 ~ "1990-2005",
    year >= 2006 & year <= 2022 ~ "2006-2022",
    TRUE ~ NA_character_
  )) %>%
  filter(!is.na(Period), !is.na(GDPpc_Growth), !is.na(KOFGI))

# Calculating averages by country and period
period_means_growth <- merged_growth %>%
  group_by(country, Period) %>%
  summarise(
    Mean_KOF = mean(KOFGI, na.rm = TRUE),
    Mean_Growth = mean(GDPpc_Growth, na.rm = TRUE),
    .groups = "drop"
  )

# Scatter plots with linear regression
ggplot(period_means_growth, aes(x = Mean_KOF, y = Mean_Growth)) +
  geom_point(alpha = 0.7, color = "darkorange") +
  geom_smooth(method = "lm", color = "steelblue", se = FALSE) +

```

```

facet_wrap(~Period) +
labs(
  title = "KOF Globalization Index vs GDP per capita Growth",
  x = "Average KOF Globalization Index",
  y = "Average GDP per capita Growth (%)"
) +
theme_minimal(base_size = 13)

# =====
# Trade openness vs per capita GDP growth
# =====

# Let's use the wb_long dataset already created (1970–2022, World Bank
Trade Openness)

# Join with per capita GDP growth (WDI)
merged_trade_growth <- left_join(wb_long, wdi_growth, by = c("country",
"year"))

# Creating historical periods
merged_trade_growth <- merged_trade_growth %>%
  mutate(Period = case_when(
    year >= 1970 & year <= 1989 ~ "1970–1989",
    year >= 1990 & year <= 2005 ~ "1990–2005",
    year >= 2006 & year <= 2022 ~ "2006–2022",
    TRUE ~ NA_character_
  )) %>%
  filter(!is.na(Period), !is.na(Trade_Openness), !is.na(GDPpc_Growth))

# Calculation of averages by country and period
period_means_trade_growth <- merged_trade_growth %>%
  group_by(country, Period) %>%
  summarise(
    Mean_Trade = mean(Trade_Openness, na.rm = TRUE),
    Mean_Growth = mean(GDPpc_Growth, na.rm = TRUE),
    .groups = "drop"
  )

# Scatter plots with linear regression
ggplot(period_means_trade_growth, aes(x = Mean_Trade, y = Mean_Growth)) +
  geom_point(alpha = 0.7, color = "darkorange") +
  geom_smooth(method = "lm", color = "steelblue", se = FALSE) +
  facet_wrap(~Period) +
  labs(
    title = "Trade Openness vs GDP per capita Growth",
    x = "Average Trade Openness (World Bank)",
    y = "Average GDP per capita Growth (%)"
  ) +
  theme_minimal(base_size = 13)

# KOF vs GDPpc Growth (3×3 matrix)
# Join KOF with GDP growth (WDI)
merged_kof_growth <- left_join(kof_clean, wdi_growth, by = c("country",
"year"))

```

```

# Add historical periods
merged_kof_growth <- merged_kof_growth %>%
  mutate(Period = case_when(
    year >= 1970 & year <= 1989 ~ "1970-1989",
    year >= 1990 & year <= 2005 ~ "1990-2005",
    year >= 2006 & year <= 2022 ~ "2006-2022",
    TRUE ~ NA_character_
  )) %>%
  filter(!is.na(Period), !is.na(KOFGI), !is.na(GDPpc_Growth))

# Add income groups
plot_data_kof_growth <- merged_kof_growth %>%
  rename(Country = country) %>%
  left_join(income_groups %>% select(Country, Group), by = "Country") %>%
  filter(!is.na(Group))

# Calculate averages by Country, Period and Group
period_means_kof_growth <- plot_data_kof_growth %>%
  group_by(Country, Period, Group) %>%
  summarise(
    Mean_KOF = mean(KOFGI, na.rm = TRUE),
    Mean_Growth = mean(GDPpc_Growth, na.rm = TRUE),
    .groups = "drop"
  )

# 3x3 scatter plots
ggplot(period_means_kof_growth, aes(x = Mean_KOF, y = Mean_Growth)) +
  geom_point(alpha = 0.7, color = "darkorange") +
  geom_smooth(method = "lm", se = FALSE, color = "steelblue") +
  facet_grid(Group ~ Period) +
  labs(
    title = "KOF Globalisation vs GDP per capita Growth\nby Income Group
and Period",
    x = "Average KOF Globalisation Index",
    y = "Average GDP per capita Growth (%)"
  ) +
  theme_minimal(base_size = 13)

# Trade Openness vs GDPpc Growth (3x3 matrix)
# Joining Trade Openness with GDP pc growth
merged_trade_growth <- left_join(wb_long, wdi_growth, by = c("country",
"year"))

# Adding historical periods
merged_trade_growth <- merged_trade_growth %>%
  mutate(Period = case_when(
    year >= 1970 & year <= 1989 ~ "1970-1989",
    year >= 1990 & year <= 2005 ~ "1990-2005",
    year >= 2006 & year <= 2022 ~ "2006-2022",
    TRUE ~ NA_character_
  )) %>%
  filter(!is.na(Period), !is.na(Trade_Openness), !is.na(GDPpc_Growth))

```

```

# Adding income groups
plot_data_trade_growth <- merged_trade_growth %>%
  rename(Country = country) %>%
  left_join(income_groups %>% select(Country, Group), by = "Country") %>%
  filter(!is.na(Group))

# Calculate averages by Country, Period and Group
period_means_trade_growth <- plot_data_trade_growth %>%
  group_by(Country, Period, Group) %>%
  summarise(
    Mean_Trade = mean(Trade_Openness, na.rm = TRUE),
    Mean_Growth = mean(GDPpc_Growth, na.rm = TRUE),
    .groups = "drop"
  )

# 3x3 scatter plots
ggplot(period_means_trade_growth, aes(x = Mean_Trade, y = Mean_Growth)) +
  geom_point(alpha = 0.7, color = "darkgreen") +
  geom_smooth(method = "lm", se = FALSE, color = "steelblue") +
  facet_grid(Group ~ Period) +
  labs(
    title = "Trade Openness vs GDP per capita Growth\nby Income Group and
Period",
    x = "Average Trade Openness (World Bank)",
    y = "Average GDP per capita Growth (%)"
  ) +
  theme_minimal(base_size = 13)

# Scatter plots using the de facto trade component of the KOF
# Selecting and renaming useful columns from the KOF dataset
kof_trade <- kof %>%
  select(country, year, KOFTrGIdf) %>%
  rename(Country = country, Year = year, KOF_Trade_DeFacto = KOFTrGIdf)
%>%
  mutate(Year = as.integer(Year))

# Cleaning the World Bank dataset and renaming the columns
wb_trade <- wb_long %>%
  rename(Country = country, Year = year, Trade_Openness = Trade_Openness)

# Joining the datasets
merged <- left_join(kof_trade, wb_trade, by = c("Country", "Year"))

# Creating historical periods
merged <- merged %>%
  mutate(Period = case_when(
    Year >= 1970 & Year <= 1989 ~ "1970-1989",
    Year >= 1990 & Year <= 2005 ~ "1990-2005",
    Year >= 2006 & Year <= 2022 ~ "2006-2022",
    TRUE ~ NA_character_
  )) %>%
  filter(!is.na(Period))

```

```

# Removing rows with missing values
merged_clean <- merged %>%
  filter(!is.na(KOF_Trade_DeFacto), !is.na(Trade_Openness))

# Calculating averages by country and period
period_means <- merged_clean %>%
  group_by(Country, Period) %>%
  summarise(
    Mean_KOF_Trade = mean(KOF_Trade_DeFacto, na.rm = TRUE),
    Mean_Trade_Openness = mean(Trade_Openness, na.rm = TRUE),
    .groups = "drop"
  )

# Creating the scatter plots
ggplot(period_means, aes(x = Mean_Trade_Openness, y = Mean_KOF_Trade)) +
  geom_point(alpha = 0.7, color = "darkorange") +
  geom_smooth(method = "lm", color = "steelblue", se = FALSE) +
  facet_wrap(~Period) +
  labs(
    title = "Average Trade Openness vs KOF Trade Globalisation (de
facto)",
    x = "Average Trade Openness (World Bank)",
    y = "KOF Trade Globalisation Index (de facto)"
  ) +
  theme_minimal(base_size = 13)

# CODE FOR REGRESSIONS
# Data from Penn World Table (PWT)
# --- 1. Penn World Table (PWT) ---

data("pwt10.0")
pwt <- pwt10.0

pwt_data_raw <- pwt %>%
  select(Country = country, Year = year, rgdpna, pop,
         Investment_Share = csh_i, Gov_Consumption_Share = csh_g) %>%
  filter(Year >= 1992, Year <= 2022) %>%
  mutate(GDPpc_Level = rgdpna / pop,
         log_GDPpc_Level = log(GDPpc_Level)) %>%
  rename(Investment = Investment_Share,
         Gov_Consumption = Gov_Consumption_Share) %>%
  select(Country, Year, log_GDPpc_Level, Investment, Gov_Consumption)

# --- 2. World Bank (WDI) ---

wdi_data <- WDI::WDI(
  indicator = c(
    GDPpc_Growth = "NY.GDP.PCAP.KD.ZG",
    Inflation = "FP.CPI.TOTL.ZG",
    Life_Expectancy = "SP.DYN.LE00.IN",
    Fertility = "SP.DYN.TFRT.IN",
    Pop_Growth = "SP.POP.GROW",
    Secondary_School_Enrollment = "SE.SEC.ENRR"
  )
)

```

```

    ),
    start = 1992,
    end = 2022
  ) %>%
  rename(Country = country, Year = year) %>%
  mutate(
    Year = as.integer(Year),
    log_Life_Expectancy = log(Life_Expectancy),
    log_Fertility = log(Fertility)
  ) %>%
  select(-Life_Expectancy, -Fertility, -iso2c)

# --- 3. Loading KOF Index ---

kof_data <- read_excel("KOFGI_2024_public.xlsx") %>%
  select(Country = country, Year = year,
         KOFGI = KOFGI,
         KOFGI_EconDeFacto = KOFecGIdf,
         KOFGI_Econ = KOFecGI,
         KOFGI_TrRestr = KOFTrGIdj,
         KOFGI_Social = KOFSoGI,
         KOFGI_Political = KOFPoGI) %>%
  mutate(Year = as.integer(Year)) %>%
  filter(Year >= 1992, Year <= 2022)

# --- 3.1 Trade Openness Data ---

wb_trade_openness_raw <-
read_excel("API_NE.TRD.GNFS.ZS_DS2_en_excel_v2_327564.xls", skip = 3)

wb_trade_openness <- wb_trade_openness_raw %>%
  select(-`2023`, -`2024`, -`Indicator Name`, -`Indicator Code`, -
`Country Code`)

trade_openness_data <- wb_trade_openness %>%
  pivot_longer(cols = starts_with("19") | starts_with("20"), names_to =
"Year", values_to = "Trade_Openness") %>%
  select(Country = `Country Name`, Year, Trade_Openness) %>%
  mutate(Year = as.integer(Year)) %>%
  filter(Year >= 1992, Year <= 2022)

# --- 3.2 Freedom House Data ---

fh_raw <- read_excel("Country Status and Ratings, 1973-2013
(FINAL)_0.xls", skip = 6, col_names = FALSE)

years <- 1972:2012
col_names <- c("Country")
for (yr in years) {
  col_names <- c(col_names, paste0("PR_", yr), paste0("CL_", yr),
paste0("Status_", yr))
}

colnames(fh_raw) <- col_names[1:ncol(fh_raw)]

```

```

fh_raw <- fh_raw[-1, ]

fh_raw <- fh_raw %>%
  filter(!is.na(`Country`)) %>%
  rename(Country = `Country`)

fh_long_1 <- fh_raw %>%
  pivot_longer(
    cols = -Country,
    names_to = "Year_Var",
    values_to = "Value"
  ) %>%
  mutate(
    Year = as.integer(str_extract(Year_Var, "\\d{4}")),
    VarType = case_when(
      str_detect(Year_Var, "PR") ~ "PR",
      str_detect(Year_Var, "CL") ~ "CL",
      TRUE ~ NA_character_
    )
  ) %>%
  filter(!is.na(VarType)) %>%
  pivot_wider(names_from = VarType, values_from = Value)

fh_2013_2024 <- read_excel("All_data_FIW_2013-2024.xlsx", sheet = 2, skip
= 1, col_names = TRUE)

fh_long_2 <- fh_2013_2024 %>%
  select(Country = `Country/Territory`, Year = Edition, PR = `PR rating`,
CL = `CL rating`) %>%
  mutate(Year = as.integer(Year))

fh_long_1 <- fh_long_1 %>%
  mutate(
    PR = as.numeric(PR),
    CL = as.numeric(CL)
  )

fh_long_2 <- fh_long_2 %>%
  mutate(
    PR = as.numeric(PR),
    CL = as.numeric(CL)
  )

fh_all <- bind_rows(fh_long_1 %>% select(Country, Year, PR, CL),
                  fh_long_2 %>% select(Country, Year, PR, CL)) %>%
  arrange(Country, Year) %>%
  mutate(
    PR = as.numeric(PR),
    CL = as.numeric(CL)
  )

# --- 4. Merge dataset and initial cleaning ---

dreher_data <- pwt_data_raw %>%

```

```

left_join(wdi_data, by = c("Country", "Year")) %>%
left_join(kof_data, by = c("Country", "Year")) %>%
left_join(trade_openness_data, by = c("Country", "Year")) %>%
left_join(fh_all, by = c("Country", "Year")) %>%
mutate(Country = as.factor(Country), Year = as.integer(Year)) %>%
arrange(Country, Year)

# --- 5. Calculating five-year averages + initial/final levels ---

calculate_period_start <- function(year) {
  floor((year - 1992) / 5) * 5 + 1992
}

dreher_data_5y <- dreher_data %>%
  mutate(
    Period = calculate_period_start(Year),
    init_val = if_else(Year == calculate_period_start(Year),
log_GDPpc_Level, NA_real_),
    end_val = if_else(Year == calculate_period_start(Year) + 4L,
log_GDPpc_Level, NA_real_)
  ) %>%
  group_by(Country, Period) %>%
  summarise(
    # livelli "alla Dreher"
    log_GDPpc_Initial =
dplyr::coalesce(dplyr::first(stats::na.omit(init_val)),

dplyr::first(stats::na.omit(log_GDPpc_Level))),
    log_GDPpc_End =
dplyr::coalesce(dplyr::first(stats::na.omit(end_val)),
dplyr::last
(stats::na.omit(log_GDPpc_Level))),
    # medie quinquennali (come prima)
    across(
      c(log_GDPpc_Level, GDPpc_Growth, Investment, Gov_Consumption,
log_Life_Expectancy, log_Fertility, Pop_Growth, Inflation,
Secondary_School_Enrollment,
KOFGI, KOFGI_EconDeFacto, KOFGI_Econ, KOFGI_TrRestr,
KOFGI_Social, KOFGI_Political,
Trade_Openness, PR, CL),
      ~ mean(.x, na.rm = TRUE),
      .names = "Avg_{.col}"
    ),
    .groups = "drop"
  ) %>%
  rename_with(~ gsub("^Avg_", "", .x)) %>%
  # per OLS/FE servono crescita e livello iniziale; per GMM serve il
livello finale
  filter(!is.na(GDPpc_Growth), !is.na(log_GDPpc_Initial),
!is.na(log_GDPpc_End)) %>%
  arrange(Country, Period)

dreher_data_5y <- dreher_data_5y %>%
  left_join(income_groups, by = "Country") %>%

```

```

mutate(
  Restriction_Developed = if_else(Group == "Developed", KOFGI_TrRestr
* log_GDPpc_Initial, 0),
  Restriction_Emerging = if_else(Group == "Emerging", KOFGI_TrRestr
* log_GDPpc_Initial, 0),
  Restriction_Developing = if_else(Group == "Developing", KOFGI_TrRestr
* log_GDPpc_Initial, 0)
)

# --- Defining regression vectors ---

# Group col4
xvars_col_4.1 <- c("Investment", "Gov_Consumption", "log_Fertility",
"Pop_Growth", "Secondary_School_Enrollment", "log_Life_Expectancy",
"Inflation")
xvars_col_4.2 <- c("KOFGI", "Investment", "Gov_Consumption",
"log_Fertility", "Pop_Growth", "Secondary_School_Enrollment",
"log_Life_Expectancy", "Inflation")
xvars_col_4.3 <- c("KOFGI", "Investment", "Gov_Consumption",
"log_Fertility", "Pop_Growth", "Secondary_School_Enrollment",
"log_Life_Expectancy", "Inflation", "PR", "CL")

# Group col5
xvars_col_5.1 <-
c("Investment","Gov_Consumption","log_Fertility","KOFGI_Econ",
"Pop_Growth","Secondary_School_Enrollment","log_Life_Expectancy","Inflati
on")
xvars_col_5.2 <-
c("Investment","Gov_Consumption","log_Fertility","KOFGI_EconDeFacto",
"Pop_Growth","Secondary_School_Enrollment","log_Life_Expectancy","Inflati
on")
xvars_col_5.3 <-
c("Investment","Gov_Consumption","log_Fertility","KOFGI_TrRestr",
"Pop_Growth","Secondary_School_Enrollment","log_Life_Expectancy","Inflati
on")
xvars_col_5.4 <-
c("Investment","Gov_Consumption","log_Fertility","Restriction_Developed",
"Restriction_Emerging","Restriction_Developing","Pop_Growth","Secondary_S
chool_Enrollment","log_Life_Expectancy","Inflation")
xvars_col_5.5 <- c("Trade_Openness", "Investment", "Gov_Consumption",
"log_Fertility", "Pop_Growth", "Secondary_School_Enrollment",
"log_Life_Expectancy", "Inflation")

# --- Function to filter datasets for each model ---

prepare_data_for_ols <- function(df, dep_var, init_var, xvars, group_var
= "Country") {
  df %>%
    filter(if_all(all_of(c(dep_var, init_var, xvars)), ~ !is.na(.))) %>%
    group_by(across(all_of(group_var))) %>%
    filter(n() >= 3) %>%
    ungroup()
}

```

```

# --- Function for OLS regression with time-dependent FE and cluster
errors ---

run_single_regression_ols <- function(df, xvars, id = "Country", time =
"Period",
                                dep = "GDPpc_Growth", init =
"log_GDPpc_Initial") {
  rhs <- paste(c(init, xvars), collapse = " + ")
  fml <- as.formula(paste(dep, "~", rhs, "+", paste0("factor(", time,
)"))))
  reg_model <- plm(fml, data = df, model = "pooling")
  vcov_rob <- vcovHC(reg_model, type = "HC1", cluster = "group")
  list(model = reg_model, vcov = vcov_rob)
}

# --- Preparing data and estimate models for col4 ---

data_ols_4.1 <- prepare_data_for_ols(dreher_data_5y, "GDPpc_Growth",
"log_GDPpc_Initial", xvars_col_4.1)
data_ols_4.2 <- prepare_data_for_ols(dreher_data_5y, "GDPpc_Growth",
"log_GDPpc_Initial", xvars_col_4.2)
data_ols_4.3 <- prepare_data_for_ols(dreher_data_5y, "GDPpc_Growth",
"log_GDPpc_Initial", xvars_col_4.3)

results_col4.1 <- run_single_regression_ols(data_ols_4.1, xvars_col_4.1)
results_col4.2 <- run_single_regression_ols(data_ols_4.2, xvars_col_4.2)
results_col4.3 <- run_single_regression_ols(data_ols_4.3, xvars_col_4.3)

# --- Preparing data and estimate models for col5 ---

data_ols_5.1 <- prepare_data_for_ols(dreher_data_5y, "GDPpc_Growth",
"log_GDPpc_Initial", xvars_col_5.1)
data_ols_5.2 <- prepare_data_for_ols(dreher_data_5y, "GDPpc_Growth",
"log_GDPpc_Initial", xvars_col_5.2)
data_ols_5.3 <- prepare_data_for_ols(dreher_data_5y, "GDPpc_Growth",
"log_GDPpc_Initial", xvars_col_5.3)
data_ols_5.4 <- prepare_data_for_ols(dreher_data_5y, "GDPpc_Growth",
"log_GDPpc_Initial", xvars_col_5.4)
data_ols_5.5 <- prepare_data_for_ols(dreher_data_5y, "GDPpc_Growth",
"log_GDPpc_Initial", xvars_col_5.5)

results_col5.1 <- run_single_regression_ols(data_ols_5.1, xvars_col_5.1)
results_col5.2 <- run_single_regression_ols(data_ols_5.2, xvars_col_5.2)
results_col5.3 <- run_single_regression_ols(data_ols_5.3, xvars_col_5.3)
results_col5.4 <- run_single_regression_ols(data_ols_5.4, xvars_col_5.4)
results_col5.5 <- run_single_regression_ols(data_ols_5.5, xvars_col_5.5)

### --- GMM regressions ---
# Defining two helper functions to allow modelsummary to read pdynmc
models.

```

```

# This function extracts the coefficients and standard errors.
tidy.pdynmc <- function(x, ...) {
  s <- summary(x)$coefficients
  tibble(
    term = rownames(s),
    estimate = s[, "Estimate"],
    std.error = s[, "Std.Err.rob"],
    p.value = s[, "Pr(>|z.rob|)"]
  )
}

glance.pdynmc <- function(x, ...) {
  jtest_p <- jtest.fct(x)$pvalue
  m2_p <- mttest.fct(x, m = 2, type = "all")$pvalue
  nobs <- nobs(x)

  tibble(
    N = nobs,
    r.squared = NA_real_,
    adj.r.squared = NA_real_,
    jtest.pvalue = jtest_p,
    m2.pvalue = m2_p
  )
}

dpd_model_dreher_col4.1 <- pdynmc(
  dat = dreher_data_5y_gmm_col4,
  varname.i = "Country",
  varname.t = "Period",

  use.mc.diff = TRUE,
  use.mc.lev = TRUE,
  use.mc.nonlin = FALSE,

  include.y = TRUE,
  varname.y = "log_GDPpc_End",
  lagTerms.y = 1,
  maxLags.y = 3,
  include.x = TRUE,

  inst.collapse = TRUE,
  w.mat = "iid.err",
  std.err = "corrected",
  estimation = "twostep",

  varname.reg.end = c("KOFGI", "Investment", "Gov_Consumption",
"log_Fertility"),
  lagTerms.reg.end = rep(1,4),
  maxLags.reg.end = rep(3,4),

  varname.reg.ex = c("Pop_Growth",
"Secondary_School_Enrollment",
"log_Life_Expectancy",
"Inflation"),

```

```

lagTerms.reg.ex = rep(0,4),
maxLags.reg.ex = rep(3,4),

include.dum = TRUE,
dum.diff = TRUE,
dum.lev = FALSE,
varname.dum = "Period"
)

dpd_model_dreher_col4.2 <- pdynmc(
  dat = dreher_data_5y_gmm_col4,
  varname.i = "Country",
  varname.t = "Period",

  use.mc.diff = TRUE,
  use.mc.lev = TRUE,
  use.mc.nonlin = FALSE,

  include.y = TRUE,
  varname.y = "log_GDPpc_End",
  lagTerms.y = 1,
  maxLags.y = 3,
  include.x = TRUE,

  inst.collapse = TRUE,
  w.mat = "iid.err",
  std.err = "corrected",
  estimation = "twostep",

  varname.reg.end = c("Investment", "Gov_Consumption", "log_Fertility",
                     "KOFGI_Econ", "KOFGI_Social", "KOFGI_Political"),
  lagTerms.reg.end = rep(1,6),
  maxLags.reg.end = rep(3,6),

  varname.reg.ex = c("Pop_Growth",
                    "Secondary_School_Enrollment",
                    "log_Life_Expectancy",
                    "Inflation"),
  lagTerms.reg.ex = rep(0,4),
  maxLags.reg.ex = rep(2,4),

  include.dum = TRUE,
  dum.diff = TRUE,
  dum.lev = FALSE,
  varname.dum = "Period"
)

dpd_model_dreher_col5.1 <- pdynmc(
  dat = dreher_data_5y_gmm_col5,
  varname.i = "Country",
  varname.t = "Period",

  use.mc.diff = TRUE,
  use.mc.lev = TRUE,

```

```

use.mc.nonlin = FALSE,

include.y = TRUE,
varname.y = "log_GDPpc_End",
lagTerms.y = 1,
maxLags.y = 3,
include.x = TRUE,

inst.collapse = TRUE,
w.mat = "iid.err",
std.err = "corrected",
estimation = "twostep",

varname.reg.end = c("KOFGI_Econ", "Investment", "Gov_Consumption",
"log_Fertility"),
lagTerms.reg.end = rep(1, 4),
maxLags.reg.end = rep(3, 4),

varname.reg.ex = c("Pop_Growth",
                  "Secondary_School_Enrollment",
                  "log_Life_Expectancy",
                  "Inflation"),
lagTerms.reg.ex = rep(0, 4),
maxLags.reg.ex = rep(2, 4),

include.dum = TRUE,
dum.diff = TRUE,
dum.lev = FALSE,
varname.dum = "Period"
)

dpd_model_dreher_col5.2 <- pdynmc(
  dat = dreher_data_5y_gmm_col5,
  varname.i = "Country",
  varname.t = "Period",

  use.mc.diff = TRUE,
  use.mc.lev = TRUE,
  use.mc.nonlin = FALSE,

  include.y = TRUE,
  varname.y = "log_GDPpc_End",
  lagTerms.y = 1,
  maxLags.y = 3,
  include.x = TRUE,

  inst.collapse = TRUE,
  w.mat = "iid.err",
  std.err = "corrected",
  estimation = "twostep",

  varname.reg.end = c("Investment", "Gov_Consumption",
                    "KOFGI_EconDeFacto", "log_Fertility"),
  lagTerms.reg.end = rep(1, 4),

```

```

maxLags.reg.end = c(2,3,3,3),

varname.reg.ex = c("Pop_Growth",
                  "Secondary_School_Enrollment",
                  "log_Life_Expectancy",
                  "Inflation"),
lagTerms.reg.ex = rep(1, 4),
maxLags.reg.ex = c(3,2,2,2),

include.dum = TRUE,
dum.diff = FALSE,
dum.lev = FALSE,
varname.dum = "Period"
)

dpd_model_dreher_col5.3 <- pdynmc(
  dat = dreher_data_5y_gmm_col5,
  varname.i = "Country",
  varname.t = "Period",

  use.mc.diff = TRUE,
  use.mc.lev = TRUE,
  use.mc.nonlin = FALSE,

  include.y = TRUE,
  varname.y = "log_GDPpc_End",
  lagTerms.y = 1,
  maxLags.y = 3,
  include.x = TRUE,

  inst.collapse = TRUE,
  w.mat = "iid.err",
  std.err = "corrected",
  estimation = "twostep",

  varname.reg.end = c("Investment",
                    "Gov_Consumption", "log_Fertility"),
  lagTerms.reg.end = rep(1, 3),
  maxLags.reg.end = c(2,3,3),

  varname.reg.ex = c("Pop_Growth",
                    "Inflation", "Restriction_Developed",
"Restriction_Emerging",
                    "Restriction_Developing", "log_Life_Expectancy"),
  lagTerms.reg.ex = rep(1, 6),
  maxLags.reg.ex = c(3,2,2,2,2,2),

  include.dum = TRUE,
  dum.diff = TRUE,
  dum.lev = FALSE,
  varname.dum = "Period"
)

dpd_model_dreher_col5.4 <- pdynmc(

```

```

dat = dreher_data_5y_gmm_col4,
varname.i = "Country",
varname.t = "Period",

use.mc.diff = TRUE,
use.mc.lev = TRUE,
use.mc.nonlin = FALSE,

include.y = TRUE,
varname.y = "log_GDPpc_End",
lagTerms.y = 1,
maxLags.y = 2,
include.x = TRUE,

inst.collapse = TRUE,
w.mat = "iid.err",
std.err = "corrected",
estimation = "twostep",

varname.reg.end = c("Investment", "Gov_Consumption",
                   "Trade_Openness", "log_Fertility"),
lagTerms.reg.end = rep(1,4),
maxLags.reg.end = c(2,2,3,2),

varname.reg.ex = c("Pop_Growth",
                  "Secondary_School_Enrollment",
                  "log_Life_Expectancy",
                  "Inflation"),
lagTerms.reg.ex = rep(1,4),
maxLags.reg.ex = c(2,2,2,2),

include.dum = TRUE,
dum.diff = TRUE,
dum.lev = FALSE,
varname.dum = "Period"
)

```