Economic diversification and growth in CIS countries: Panel data analysis

By
Aruzhan Yertay

Submitted to Central European University - Private University
Department of Public Policy

In partial fulfilment of the requirements for the degree of Master of Public Policy

Supervisor: Anand Murugesan

Vienna, Austria
2025

COPYRIGHT NOTICE

[1]

[1] Icon by Font Awesome.

AUTHOR'S DECLARATION

I, the undersigned, **Aruzhan Yertay**, candidate for the MA degree in Public Policy declare herewith that the present thesis titled "Economic diversification and growth in CIS countries: Panel data analysis" is exclusively my own work, based on my research and only such external information as properly credited in notes and bibliography.
I declare that no unidentified and illegitimate use was made of the work of others, and no part of the thesis infringes on any person's or institution's copyright.

I also declare that no part of the thesis has been submitted in this form to any other institution of higher education for an academic degree.

Vienna, 03 06 2025

Aruzhan Yertay

ABSTRACT

This thesis investigates the relationship between economic diversification and economic growth in Commonwealth of Independent States (CIS). The study uses panel data for 9 CIS countries from 2002 to 2021, where Herfindahl-Hirschman Index (HHI) is used to measure the economic diversification and GDP per capita growth to measure economic growth. The analysis reveals a negative coefficient for HHI, showing that more diversified countries or lower HHI associated with higher GDP growth. Additionally, diversification benefits most with lower resource dependence and the effect diminishes as it increases after the threshold of 20.4% in CIS. The study shows that there is no one fit policy for all, and it should be tailored for each economy individually.

Keywords: Economic diversification, Oil dependence, Herfindahl-Hirschman Index, CIS.

TABLE OF CONTENTS

LIST OF FIGURES, TABLES, OR ILLUSTRATIONS

CHAPTER ONE: INTRODUCTION

For a long time, the primary goal of national development programs has been to promote economic growth and stability. However, the approaches to achieving these goals vary significantly and are often fraught with challenges, particularly for countries whose economies rely heavily on a single natural resource. Fluctuating commodity prices, growing environmental concerns, and rapid technological advancements underscore the urgent need for economic diversification. This thesis explores this imperative, using CIS region as a case study of resource-rich, upper-middle-income countries striving to transition toward a more stable and sustainable economic future.

It has been always interesting to know why some countries are poor and other are rich. There are different theories that suggested different reasons for this question. Yet, the division between rich and poor emerged only after 1500, where industrialization played a key role (Allen 2011). Industrialization and new technology afterwards had a huge impact to all countries (Allen 2011). According to Hidalgo et al. 2007, it is because of the differences in industrial development and the accumulation of knowledge and skills that are important for the production of these complex goods and services. This can be challenging for resource dependent countries, as even if it generates large revenues, they are in the risk of the "resource curse" or also "Dutch disease" (van der Ploeg 2011).

The discussion about economic diversification has been a topic for a long time, but it has been more talked in Commonwealth of Independent States (CIS) in recent years. While there is a huge amount of literature on economic diversification and the resource curse, there is not much quantitative research on the given region. First, most empirical studies focus on Sub-Saharan Africa, Latin America, or Gulf states, with limited systematic analysis of CIS countries

as a distinct group with shared institutional and historical legacies. Second, while individual country studies provide qualitative insights, few employ quantitative panel data approaches that can identify generalizable patterns across the region. Third, existing research rarely examines the conditional nature of diversification benefits, whether the growth effects of diversification vary with resource dependence levels. This thesis addresses these gaps by providing the first comprehensive quantitative analysis of the diversification and growth relationship across multiple CIS countries.

Kazakhstan's economy has long relied on extractive industries, making it vulnerable to commodity price fluctuations and limiting industry growth (Aubakirova, Zhuparova & Kozhakhmetova 2025; OECD 2024; Hausmann et al., 2023). Despite steady economic growth over the past two decades, Kazakhstan's export basket remains largely composed of primary commodities like petroleum oils, which made up 32.35% of total exports in 2023 (Harvard Growth Lab). In doing so, Kazakhstan has limited integration to global value chain and has underdeveloped private sector as small and medium sized enterprises (SMEs) (OECD 2024). When country depends only on limited exports and trading partners, it makes the country both less trade and economically stable (OECD 2024).

This thesis's main goal is to look into the relationship between economic diversification and economic growth in the CIS, and also looking how this relationship changes depending on how much a country depends on oil. By systematically analyzing macroeconomic indicators for nine CIS countries using panel data regression techniques, this study seeks to provide insights into how diversification affects economic trajectories under different resource endowment conditions. The methodology employs two-way fixed effects panel regression with Driscoll-Kraay standard errors, interaction terms to test conditional effects, descriptive trend analysis, and robustness checks to validate the findings. Kazakhstan receives focused policy attention given its position near the identified resource curse threshold.

This thesis is organized by chapters as follows: Chapter 2 presents a thorough literature review of concepts as economic diversification, the resource curse, theories of economic growth, and diversification efforts of CIS region and Kazakhstan. Chapter 3 presents the research design with the research question and hypotheses, data variables and their definitions, and the methodology, which includes panel regression and robustness checks. Chapter 4 provides empirical findings and their analysis, including descriptive statistics, panel regression outcomes, concluding with a discussion on policy implications and research limitations. In chapter 5, I will conclude the thesis and propose policy recommendations.

CHAPTER TWO: LITERATURE REVIEW

This chapter looks at the concepts of economic diversification, resource curse, and economic growth. A particular emphasis will be put on the definitions and nuances of economic diversification, as well as a detailed examination of the resource curse phenomenon and its mechanisms, as well as a brief review of the Solow growth model. Next, I will examine the more contemporary notions of the product space and economic complexity, while taking into consideration their implications for diversification strategies. Next, diversification trajectories will be explored, to then finally provide a contextual review of the CIS region, and more specifically, Kazakhstan's own economic development and diversification efforts and challenges.

2.1. Definitions

Economic diversification has frequently been associated with economic growth, particularly in developing economies. The debate is more prevalent in developing and resource rich countries. A more diversified economy is more resilient to external shocks such as price volatility and geopolitical crises (Delechat et al. 2024). It is also linked to more sustainable, inclusive, and innovative growth by also creating new opportunities and jobs (Delechat et al. 2024). A more detailed definition of economic diversification will allow us to better understand its relationship with economic growth.

Economic diversification is the process of expanding economic activities, products, and markets by diversifying them (Delechat et al. 2024). It involves moving economic resources like labor, capital, and technology from activities with lower productivity to those with higher productivity, both within and between economic sectors (World Bank 2017). The idea is

connected to changing the structure of economies and getting higher levels of productivity by moving resources within and between economic sectors (World Bank, 2017). This change in activities from less productive to more productive is needed for long term growth and welfare of the society as a whole (World Bank 2017).

Differentiation has two main dimensions: trade diversification and domestic production diversification (World Bank 2017). Trade diversification means exporting new or better products, going into new markets, and making the quality of the existing goods even better. (World Bank 2017). This includes both a lot of new products and new markets, as well as a lot of improved margins on existing products sold in existing markets. (World Bank 2017). On the other hand, domestic production diversification means balancing output across different sectors. This makes firms move resources between industries and within industries, which raises total factor productivity (World Bank, 2017). The success of East Asian countries is a prime example of interconnectedness of trade and domestic production diversification (World Bank 2017). That is, integrating into the global economy while diversifying into manufacturing have led to notable poverty reduction in the region (World Bank 2017). Beyond these, vertical diversification involves moving into higher value added products within existing sectors, such as processing raw materials rather than merely exporting them (Akhmedov 2017). This can be seen in the upgrading of resource based sectors, for instance, Latin America's expansion in global markets for worked metal products compared to raw ores (Gelb 2010). Economic diversification is important for long term growth because manufacturing and service sectors tend to grow more quickly than agriculture and mining and contribute significantly to GDP in developing economies (Esanov 2009).

However, the presence of natural resources because of the Dutch disease often makes the path to diversification harder. Rich countries in resources like minerals and fuels tend to have slower economic growth and less diversification than resource poor countries, according to a study by

Sachs and Warner in 1995 (van der Ploeg 2011). If a country explores new natural resources, like oil or gas, its currency tends to get stronger (Aubakirova et al 2025). As a result, this makes it more expensive for other countries to buy the country's products (Aubakirova et al 2025). In doing so, these non resource sectors become less competitive, and is a common problem known as Dutch disease (Aubakirova et al 2025). This can cause the "crowding out" of these crucial sectors, delaying their development. Furthermore, because of the the concentration of wealth from natural resources, it can cause corruption, and a lack of accountability, which undermines institutional quality and distorts economic incentives away from productive investments in diversified sectors (Gelb 2010). An illustration of crowding out can be seen in Nigeria and Angola, which have weaker long-term growth and persistent poverty despite substantial resource wealth (Akhmedov 2017). Even numerous diversification plans are not always a guarantee of success. Countries like Saudi Arabia and Kuwait have struggled to reduce their reliance on oil despite such efforts, as their non-oil industries often lack the competitive power to survive independently (Akhmedov 2017). Nigerian case shows how relying too much on oil can harm non oil sectors, leading to economic and political instability and a failure to meet people's basic needs (Chuku 2021). Abundant amount of natural resources can weaken democratic institutions, marginalize the manufacturing sector, cause instability in the economy as a whole, and increase the likelihood of civil unrest (Esanov 2009).

The debate about economic growth has changed a lot over the years. Early theories, like the Solow growth model, focused on how capital accumulation, labor force growth, and technological progress improved economic growth (Solow 1956; Romer 1986). According to the Solow model, long term growth per capita is only caused by exogenous technological progress (Romer 1986). The Solow model was useful, but it was not very good at explaining why income levels stayed different between countries or why technological progress happened on its own (Romer 1986). This is because it assumed that technology was a public good that

everyone could use (Romer 1986). This led to the development of endogenous growth theories, pioneered by Paul Romer, which highlight the importance of human capital, innovation, and the generation of new ideas as internal drivers of sustained economic growth (Romer 1990). According to these ideas, long term growth depends on investing on things like education, research and development, and building institutions that encourage new ideas and innovation (Romer 1990). Small differences in annual growth rates can cause big differences in how much money different generations have. This shows how important growth is for reducing poverty and promoting prosperity (Johnson and Papageorgiou 2020).

2.2. The Product Space and Economic Complexity

The product space framework provides a structural explanation for why diversification matters for growth. Countries that export more complex products tend to have faster growth because complexity reflects productive capabilities that can be later used to enter new markets (Hidalgo & Hausmann 2009). Product Space and the related concept of Economic Complexity is influential framework that was introduced by Hidalgo and Hausmann (Hidalgo et al. 2007). According to this approach, the economy is not just a collection of sectors, but rather a network of products, where the ability to produce one product is related to the ability to produce others based on shared "know-how" and "capabilities" (Hidalgo et al. 2007).

In this framework, "know-how" is likened to letters in a Scrabble game, and products are like words (Hausmann and Hidalgo 2011). A country's economic capabilities are its "letters," and its productive structure is the set of "words" it can form. Producing more complex products requires a greater and more diverse set of capabilities. The "product space" is a visual representation of how products are related, showing that more complex products like machinery, chemicals are located in a densely connected "core," while less complex like primary products as agricultural goods or raw materials occupy a less connected "periphery"

(Hidalgo et al. 2007). Countries tend to move through product space by developing products that are nearby to those already being produced (Hidalgo et al. 2007). This approach builds on existing capabilities. The distance between products in this space can be understood as the ease or difficulty of transitioning from producing one good to another. For instance, a country exporting apples likely has capabilities easily transferable to other agricultural goods, but not necessarily to copper wires or home appliances (Hidalgo et al. 2007).. This is due to a constraint of new spheres requiring vastly different sets of knowledge and infrastructure (Hidalgo et al. 2007). The process of moving through the product space intentionally is critical for economic evolution, as it involves repurposing human, physical, and institutional capital towards different goods that are close in terms of shared capabilities (Hidalgo et al. 2007).

Economic complexity is measured by Economic complexity index (ECI) for countries and Product complexity index (PCI) for products (Hausmann and Hidalgo 2011). An ECI shows how large and varied a country's exports are and shows how much knowledge countries have about making these products. The higher ECI the faster economic growth, because more capabilities means more production and more variety of goods (Hausmann et al. 2020). That is, production of more complex goods is indicative of an inter-industry spillover, and an opportunity for the exploitation of increasing returns. This exact ability to exploit inter-industry spillovers is a key reason why what a country produces matters for its economic trajectory (Hausmann et al. 2020).

The framework of product space and economic complexity offers a powerful lens through which resource curse can be explored. Resource intensive sectors such as oil and mining entail a very specific and very limited set of capabilities. These usually cannot be utilized to expand and move through product space as there are not many nearby products that can be developed using similar capabilities (Hidalgo & Hausmann 2009). The resource rich countries are therefore limited in their "letters" or capabilities to arrange new "words" or products, thus

getting locked in their capabilities (Hidalgo & Hausmann 2009). This makes it difficult for resource dependent countries to diversify and strive for economic sophistication, which further contributes to the "Great Divergence" (Johnson and Papageorgiou 2020). As this divergence continues to exist, the "alphabet and words get longer," and countries with few "letters" find it more difficult to catch up (Hidalgo & Hausmann 2009). However, the concept also suggests pathways for "convergence," particularly through globalization and the fragmentation of global value chains. This allows countries to specialize in "syllables" or specific tasks within a complex production process, rather than needing to master all "letters" to produce a whole "word" or final product (Hidalgo & Hausmann 2009). This makes it easier for countries to enter global production networks and gradually acquire new capabilities, thus easing the path for diversification. The challenge for resource rich countries is that their resource intensive sectors are concentrated in a small portion of the product space, limiting the opportunities for inner industry spillovers. This explains why the resource curse is frequently associated with slow growth in the economy's non resource sectors (Lashitew, Ross, and Werker 2020).

2.3. Diversification efforts of the CIS region

The Commonwealth of Independent States (CIS) is a group of countries formed collectively after the fall of the Soviet Union. They are Armenia, Azerbaijan, Belarus, Kazakhstan, Kyrgyzstan, Moldova, Russia, Tajikistan, Turkmenistan, Ukraine, and Uzbekistan. The eleven post-Soviet countries of the CIS share a unique economic legacy rooted in centrally planned industrialization followed by abrupt transition toward market economies. In the early 1990s, many CIS countries' economies weakened dramatically, trade opened up quickly, and industries separated, which resulted in the fact that economies need to focus again on extracting resources or farming (Sarosh and Shiells 2004, Souza et al. 2006).

While some countries like Belarus maintained industrial continuity with Soviet-era production lines, others, such as Turkmenistan, Azerbaijan, and Kazakhstan, developed strong reliance on hydrocarbons. Armenia, Moldova, and Kyrgyzstan, with fewer extractive resources, pursued diversification through services as IT, tourism, and agriculture, but remained vulnerable to external shocks and migration (Petrovskaya et al. 2023, Song and Jai 2020, Sarosh and Shiells 2004, Souza et al. 2006).

In the last 20 years, CIS countries have used a variety of strategies to diversify their economies. These range from import substitution industrialization (e.g., Uzbekistan), integration into Eurasian economic union value chains (e.g., Belarus), to global market integration via Special economic zones (e.g., Azerbaijan) (Petrovskaya et al. 2023). Nevertheless, export concentration levels remain high across the region. Many countries exhibit relatively high Herfindahl-Hirschman Index (HHI) scores, indicating low levels of product diversification. Additionally, many of the activities in the non resource sector are focused on making low-value goods and do not have much technological progress or joining global value chains.

Institutional constraints such as weak governance, limited innovation ecosystems, and underinvestment in education remain common obstacles. Nonetheless, there are cases of emerging capability building in areas like IT in Armenia, agritech in Uzbekistan, and industrial development in Belarus (Petrovskaya et al. 2023). These variations across the CIS countries offer a compelling setting for testing the relationship between diversification and growth outcomes.

2.4. Diversification efforts of Kazakhstan

Kazakhstan is the largest landlocked country, ranking at 9th with a territory of 2, 724, 900 square kilometers (TRACECA 2025). Kazakhstan declared its sovereignty in 1990. Thus, the

available dataset on Kazakhstan comes only starting 1990. Kazakhstan has a population of 20.2 million (World Bank 2025).

Kazakhstan had a lot of natural resources, but the main oil and gas fields had not been explored before the country got its independence (Kalyuzhova et al. 2004:249). About a third of the world's chromium and manganese deposits, and large reserves of tungsten, lead, zinc, copper, bauxite, phosphorus, and coal and iron can be found here (Gleason 2003:48).

Since gaining independence in 1991, Kazakhstan has undergone significant economic and political transformations. Early on, the country prioritized macroeconomic stabilization, trade liberalization, and privatization. These first-generation reforms, coupled with a commodity boom from 2003 to 2008 and relatively sound fiscal management, supported robust economic growth averaging around 8.5% annually between 2000 and 2008 (World Bank 2025). However, much of this growth was driven by the extractive sector, particularly oil and gas, which continued to dominate the economy's structure and export earnings.

Aware of the long-term risks of excessive resource dependence, Kazakhstan's government introduced a series of strategic plans beginning with its long-term development strategy in 1997. Inspired by the East Asian development model, this strategy aimed to use resource windfalls to foster industrialization and support non-oil sectors. The government identified priority sectors and established development agencies to channel investment and coordinate diversification efforts (Esanov 2009). Despite these efforts, there haven't been many results in diversification. The hydrocarbon sector continues to have a big impact on both the structure of GDP and the export portfolio. (World Bank 2025). Multiple structural and institutional challenges have constrained Kazakhstan's diversification efforts. These include a lack of institutional reform to match economic liberalization, weak governance, limited financial sector regulation, and a geographic disadvantage due to its landlocked position (Vos and Koparanova 2011; Esanov

2009) . While various public agencies were created to support export-led growth, issues such as bureaucratic inefficiency, overlapping mandates, and weak accountability hindered effective implementation (Zhukov et al. 2023). The financial sector, in particular, tended to channel credit toward construction and consumption rather than productive manufacturing, reflecting a broader misalignment between policy intentions and market realities (Esanov 2009; Vos and Koparanova 2011).

Kazakhstan's diversification experience also suffered from the uncritical adoption of foreign models. Efforts to emulate the East Asian industrialization trajectory were not sufficiently adapted to Kazakhstan's institutional context or regional constraints. Consequently, structural transformation remained shallow. For instance, while the agricultural sector declined in share between 2000 and 2009, manufacturing failed to grow meaningfully, and the mining sector increased its share of total output (Esanov 2009). Though the service sector and construction expanded, these developments did not represent a shift toward technologically advanced or export-competitive industries.

Export diversification metrics confirm persistence of focus. Between 1998 and 2006, Kazakhstan's export concentration worsened significantly, as reflected in the increase of its export diversification index from less than 0.10 to over 0.30. The global financial crisis of 2008–2009 showed how weak a narrow economy can be (World Bank 2025). Falling oil prices triggered a fiscal shortfall, a liquidity crisis in the banking sector, and forced the government to withdraw $10 billion from the National Oil Fund to stabilize the economy (Esanov 2009).

In sum, Kazakhstan's experience illustrates the complexity of achieving diversification in a resource-rich, transition economy. While macroeconomic performance and initial reforms laid the foundation for development, insufficient institutional quality, policy misalignments, and external shocks undermined the effectiveness of diversification efforts. The lack of significant

structural change in both output and export composition suggests that future strategies must prioritize institutional development, transparent public policy, and more realistic assessments of domestic capabilities.

2.5. Challenges and opportunities specific to diversification of Kazakhstan

Despite significant financial resources and a series of strategic plans, Kazakhstan's economic diversification efforts have delivered limited results. Several key constraints continue to hinder progress for Kazakhstan. These include misaligned economic policies, an inadequately adapted diversification strategy, and underdeveloped political and economic institutions (Esanov 2009; Zhukov et al. 2023; Kalyuzheva et al. 2004; Souza et al. 2006).

As it was previously highlighted, Kazakhstan's initial diversification model drew heavily from East Asian development experiences but failed to adjust for local context. High-tech sectors were prioritized prematurely, bypassing intermediate stages of industrial learning (Esanov 2009). This led to missed opportunities in building entrepreneurial and organizational capabilities (Esanov 2009). Institutionally, the financial sector remained poorly regulated, leading to excessive borrowing by banks and credit expansion into construction and consumer goods, rather than into tradable manufacturing (Esanov 2009). The Samruk-Kazyna National Welfare Fund, intended to promote diversification and stabilization, has been criticized for weak transparency, overlapping mandates, and limited performance accountability (Esanov 2009). These issues undermined its ability to efficiently allocate capital and support productive transformation. Kazakhstan's investment patterns between 2000–2007 were also inconsistent with its stated diversification goals: the bulk of capital went to mining, construction, and real estate, while the share going to manufacturing remained stagnant at around 10% (Esanov 2009; Zhukov et al. 2023). This misallocation reflected a persistent gap between development strategies and implementation.

Geography adds a further challenge: as a landlocked country with limited access to global markets, Kazakhstan faces high transport costs and limited exposure to advanced technological partners. Its trade is oriented primarily toward Russia and China, which do not offer the same opportunities for technological upgrading as more industrialized economies (Vos and Koparanova 2011; Esanov 2009).

Nonetheless, opportunities do exist. Kazakhstan has relatively strong fiscal buffers, improving digital infrastructure, and a growing services sector. If reforms are prioritized, especially in institutions, investment policy, and regional integration, Kazakhstan could still unlock more diverse and sustainable growth pathways.

2.6. Research gap

There has been a lot of research on the link between economic diversification and growth in Sub-Saharan Africa, Latin America, and the Gulf states. However, post-Soviet economies, especially those in the CIS region, have not gotten as much attention. Much of the existing literature on CIS countries focuses either on macroeconomic transition, political economy, or resource dependence in isolation (Sarosh and Shiells 2004; Souza et al. 2006; Esanov 2009), often lacking a systematic analysis of how export concentration affects long term per capita growth.

Studies like Chuku (2021) provide relevant insights from other resource rich, middle income economies, yet their findings cannot be automatically generalized to the institutional and historical specificities of the CIS region. This is due to specific Soviet era industrial legacies, uneven institutional reform, and geographic constraints present unique challenges to diversification. Moreover, while individual country studies, particularly on Kazakhstan, offer qualitative assessments of diversification outcomes, few adopt a quantitative approach that compares multiple CIS economies using consistent indicators such as the Herfindahl-

Hirschman Index (HHI) for export concentration. Finally, few studies simultaneously control for key structural factors such as investment levels, labor force participation, human development, and exchange rate dynamics, which shape a country's growth capacity.

This study fills in these gaps by using panel data to do the first quantitative analysis of the relationship between diversification and growth in CIS region countries. In particular, it adds to the literature by first how economic concentration and oil dependence interact; resource curse thresholds that are specific to the CIS; marginal effect calculations that show how the benefits of diversification change with resource levels; giving countries tailored policy advice based on where they stand in relation to the identified thresholds. This method goes beyond just pushing for diversification, and instead to resource dependent transition economies more specific policy recommendation.

CHAPTER THREE: RESEARCH DESIGN

This chapter introduces the research design for the connection between economic diversification and growth in the Commonwealth of Independent States (CIS). First, I will state the primary research question and the hypothesis. Subsequently, I will describe the data, explaining its origin, the reason for selecting specific countries for the sample, and the definition of each variable. A crucial aspect of this chapter discusses the research methodology. This includes the rationale for using a panel data regression model with two-way fixed effects, the equations to be examined, and the measures implemented to address potential statistical issues such as endogeneity, cross-sectional dependence, and heteroskedasticity. The last part of the chapter goes over the diagnostic tests and robustness checks that were used to make sure the results were valid and accurate.

3.1. Research questions and hypotheses

This thesis looks at the relationship between economic diversification and economic growth in CIS countries. The main research question of the thesis: "Does economic diversification promote higher GDP per capita growth in CIS countries, and how does this relationship vary with oil dependence levels?"

This question addresses the policy challenge facing former Soviet countries as they build more stable and balanced economies while managing natural resource wealth.

Hypotheses:

H1: Economic diversification or lower Herfindahl-Hirschman Index (HHI) is positively associated with higher GDP per capita growth in CIS countries.

H2: The growth benefits of economic diversification diminish as oil dependence increases.

These hypotheses suggest that diversified economies are better positioned for sustained growth, and how much they benefit from this depends on the resources they have.

3.2. Data sources and variable definitions

The dataset was collected for all Commonwealth Independent states from 1990 until 2024 from the World Bank dataset. The countries include: Armenia, Azerbaijan, Belarus, Kazakhstan, Kyrgyzstan, Moldova, Russia, Tajikistan, Turkmenistan, Ukraine, and Uzbekistan. Table 1 shows all the variables collected from the World Bank dataset. The variables were adopted from the paper by Chuku, Chigozile Nnamdi 2021.

Table 1. Variables

Variable	**Definition**	**Source**
Independent variable		
HHI	Herfindahl-Hirschman Index (HHI). Constructed from sectoral shares below:	
1. *SER*	Services, value added (% of GDP)	World Bank
2. *MAN*	Manufacturing, value added (% of GDP)	World Bank

3. AGR	Agriculture, forestry, and fishing, value added (% of GDP)	World Bank
Dependent variable		
GDPCG	GDP per capita growth (annual %)	World Bank
Control variables		
OIL	Oil rents (% of GDP)	World Bank
GCF	Gross capital formation (% of GDP)	World Bank
HDI	Human development index	World Bank
LFPR	Labor force participation rate	World Bank
Log (EXCR)	Natural logarithm of official exchange rate	World Bank

3.3 Data preparation

The initial dataset was collected for all Commonwealth Independent states (CIS) from 1990 to 2024 from the World bank dataset. The countries include: Armenia, Azerbaijan, Belarus, Kazakhstan, Kyrgyzstan, Moldova, Russia, Tajikistan, Turkmenistan, Ukraine, and Uzbekistan. During the data preparation, there were missing values for certain variables and

countries. Thus, I chose the period from 2002-2021.. This time frame also includes the period after the transition when most CIS countries became more stable. Additionally, Turkmenistan and Armenia had missing data for several variables, thus, to achieve balanced panel data these 2 countries were removed from the dataset. Even though the sample size is reduced, the paper makes sure that the data is accurate and can be compared across the nine remaining countries. Even though Azerbaijan had times of very high GDP per capita growth, as shown in Figure 1, the country was still included in the final sample for robustness checks. However, because key variables like HHI, GCF, LFPR, and HDI are lagged by one period, the actual analysis period is from 2003 to 2021, with 171 observations from 9 countries spanning 19 years. To deal with endogeneity issues, this lagging approach is necessary because the effects of structural economic changes and investments usually take time to show up. The drop from 180 to 171 observations was intentional and not caused by missing data. The resulting panel is strongly balanced with with a total of 171 observations.

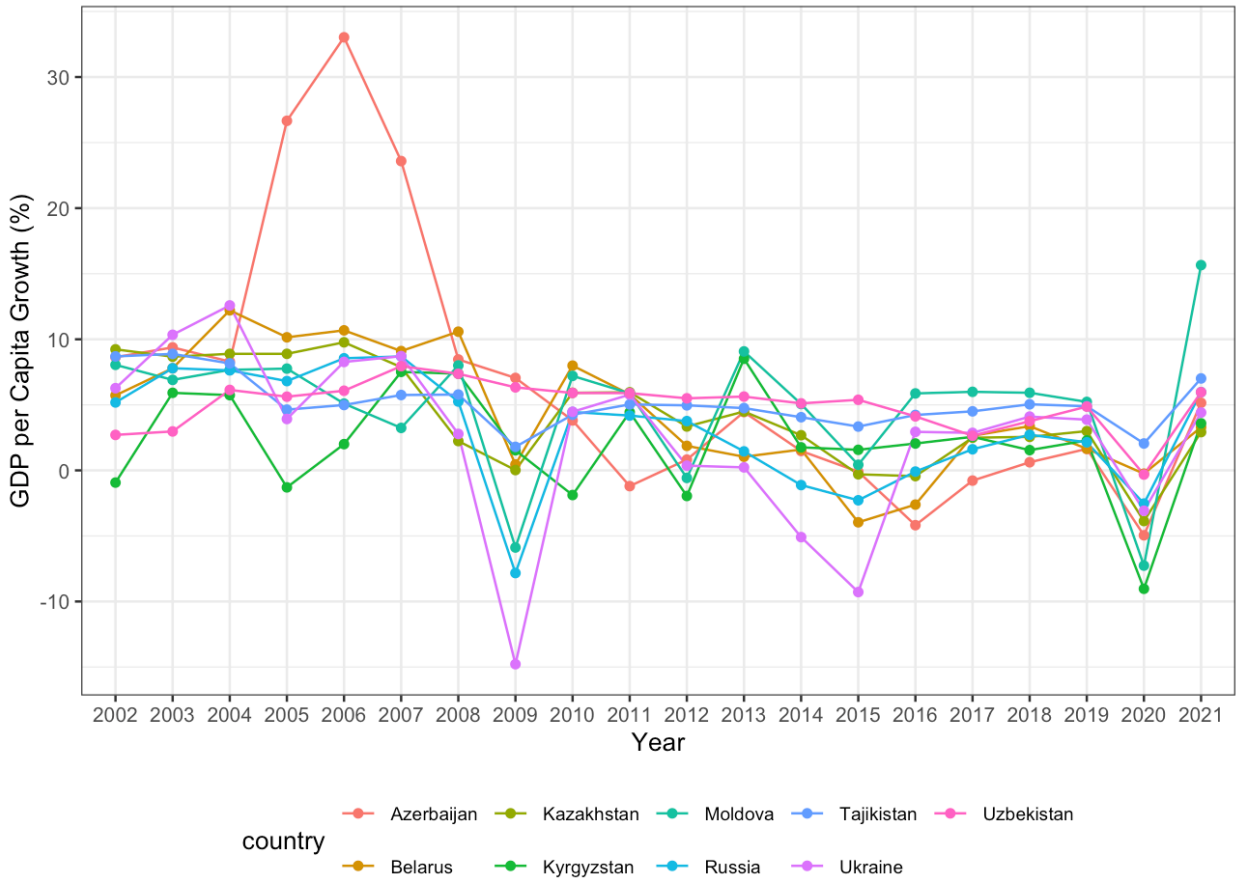


Figure 1. GDP per capita growth for CIS from 2002-2021

3.4 Variables and Herfindahl-Hirschman Index (HHI)

The main outcome this study looks at is the GDP per capita growth rate (GDPCG), which is a standard way to measure economic growth. The key independent variable used to measure economic diversification is the Herfindahl-Hirschman Index (HHI).

To find out how diverse the economy is as a whole and to compare economies over time or between countries, there is a single measure that can be used. For cross-country or longitudinal analyses, this is especially helpful because it gives a reliable measure. The Herfindahl-Hirschman Index (HHI) will be used for this particular study. According to Belbali et al. (2024), the HHI is a well-known and useful way to find out how concentrated and diverse an economy is.

The HHI is calculated as the sum of the squares of the share of each sector in total GDP (Belbali et al. 2024). The formula is:

$$HHI = \Sigma_{i=1}^{n}(S_i)^2$$

The HHI is calculated by adding up the shares (S_i) of different productive sectors in the country's total gross domestic product (GDP). Each $S_i$is given as a percentage, like 0.2 for a 20% share of total GDP. Most of the time, the HHI value is between 0 and 1. A lower HHI value close to 0 means that the economy is more diverse. This means that economic activity is spread out more evenly across different sectors and no single sector is in charge. If the HHI value is higher, closer to 1, it means that the economy is less diverse and more concentrated, meaning that it depends on a few sectors more than others (Belbali et al. 2024;3030). These economies are more likely to be affected by sector risks like changes in prices or demand (Belbali et al. 2024:3030). The study looks at the 9 CIS countries from 2002 to 2021. HHI will be used to get a full picture of these countries. It will include the most important economic sectors, such as Agriculture (AGR), Manufacturing (MAN), and Services (SER), along with a "Other Sectors" category that is calculated as

S_{other}=1−(S_{AGR}+S_{MAN}+S_{SER}) to make sure that all of the GDP components add up to one unit, or 100%. If a country's HHI goes down over time, it means that its economy is becoming more diverse. On the other hand, if the HHI stays the same or goes up, it could mean that the country isn't making much progress. The HHI will be used as an independent variable to assess the impact of diversification on economic growth.

So, it will be:

$$HHI = {S_{AGR}}^2 + {S_{MAN}}^2 + {S_{SER}}^2 + {S_{other}}^2$$

Furthermore, the study also includes a number of control variables that are known to be important in economic growth theories. These help to get a better picture of what causes growth. For example, oil rents as a percentage of GDP and oil² are used to look at the idea of the resource curse. We can use the squared term to see if the link between oil and growth is not linear. As is common in growth models, gross capital formation (GCF) is used to show how much money is being invested. Labor force participation rate (LFPR) are also used to look at the workforce. The official exchange rate's natural logarithm is used to track the value of the country's currency and the economy as a whole.

Some variables are delayed by one period to avoid the statistical problem of endogeneity. All of these indicators the HHI, GCF, LFPR, and HDI are "lagged," which means that the model uses the value from the previous year. This is because changes to the economy's structure, new investments, and better human capital all take time to have an effect on growth. However, oil rents and the exchange rate are used without the lagging because they can have an immediate impact on the economy.

3.5 Methods

First, the study will use descriptive analysis to look for patterns and trends in how the economies of CIS countries are becoming more diverse. Then, it will use panel data regression to quantify the relationship between diversification and economic growth

3.5.1 Panel Data

For this paper, panel data regression analysis is used to analyze the economic diversification and its relationship with economic growth in 9 CIS countries from 2002 to 2021.

Panel data, also known as longitudinal data, obtained by observing the same countries, firms, individuals across time (Wooldridge 1999). It is possible to look at differences between

variables and changes over time because it uses both cross-sectional and time series data. (Wooldridge 1999). This data structure is better than cross sectional or time series datasets because it gives more information, more variability, fewer variables that are similar, more degrees of freedom, and more efficiency (Wooldridge 1999). Specifically, panel data accounts for individual unobserved heterogeneity that stays the same over time and find and measure effects that cannot be found in cross-sectional or time-series data alone (Wooldridge 1999).

With panel data, it is possible to control for unobserved country specific factors that may be correlated to other variables (Wooldridge 1999). With more data points and degrees of freedom, it decreases collinearity among explanatory variables, which makes econometric estimates more accurate (Wooldridge 1999). Panel data is the best way to look at how things change and adapt over time, which is very important when looking at how economies diversify.

As this study looks at many CIS countries over 20 years, it's important to think about both factors that are unique to each country and shocks that happen at the same time every time. Thus, the two-way fixed effects (TWFE) model will be used. The TWFE model estimates the coefficients by including dummy variables for each country and for each year (Wooldridge 1999). Country fixed effects account for things such as geography, historical legacies, and stable institutional factors (Wooldridge 1999). Time fixed effects account for factors that affect all countries in a given year, such as global economic shocks, widespread technological changes, and major geopolitical events (Wooldridge 1999).

The TWFE model helps reduce the omitted variable bias caused by unobserved factors by taking into account both country and time fixed effects. This makes the estimates of the relationships between the variables more reliable. The model effectively analyses the relationship between deviations from a country's mean and a year's mean. To take into account

the possibility of heteroskedasticity and serial correlation between countries, robust standard errors will be used at the country level.

When economies are linked together, like in the CIS, cross-sectional dependence is a big problem. This is when economic shocks in one country can affect other countries. The Pesaran CD test and other diagnostic tests confirmed that this dataset has a lot of cross-sectional dependence. This is not taken into account by standard errors that are grouped by country, which can make statistical conclusions less reliable. Thus, Driscoll-Kraay (DK) standard errors are used to estimate the model so that valid and reliable results can be obtained. It is designed so that this method works well even when there are different types of cross-sectional dependence, as well as when there is heteroskedasticity and serial correlation.

The study uses three different versions of the model to test hypotheses in depth. The first one is a simple nonlinear model that tests the idea of a resource curse by adding the squared term for oil rents. The second is an interaction model, which is the most important one for this thesis. This model includes a term that multiplies both the HHI and the oil rents. This lets us look at how diversification affects oil wealth. The third is a simple linear model, which leaves out the non-linear and interaction terms and acts as a baseline for comparison.

(1) Model 1: Non-linear

$$GDPCG_{it} = \beta_0 + \beta_1 HHI_{it-1} + \beta_2 OIL_{it} + \beta_3 OIL^2{}_{it} + \beta_4 LFPR_{it-1} +$$

$$\beta_5 GCF_{it-1} + \beta_6 HDI_{it-1} + \beta_7 LOG(EXRATE)_{it} + \alpha_i + \lambda_t + \varepsilon_{itt}$$

(2) Model 2: Interaction

$$GDPCG_{it} = \beta_0 + \beta_1 HHI_{it-1} + \beta_2 OIL_{it} + \beta_3 OIL^2{}_{it} + \beta_4 (HHI_{it-1} \times OIL_{it}) +$$

$$+ \beta_5 LFPR_{it-1} + \beta_6 GCF_{it-1} + \beta_7 HDI_{it-1} + \beta_8 LOG(EXRATE)_{it} + \alpha_i + \lambda_t + \varepsilon_{itt}$$

(3) Model 3: Linear

$$GDPCG_{it} = \beta_0 + \beta_1 HHI_{it-1} + \beta_2 OIL_{it} + \beta_3 LFPR_{it-1} + \beta_4 GCF_{it-1} + \beta_5 HDI_{it-1} + \\ + \beta_6 LOG(EXRATE)_{it} + \alpha_i + \lambda_t + \varepsilon_{it}$$

Where:

$GDPCG_{it}$ is the GDP per capita growth for country i at time t. β_O is the overall intercept. β_k are the coefficients for the independent variables. α_i represents unobserved country specific fixed effects that capture time invariant features of each country. λ_t shows time fixed effects by capturing shocks or trends that happen in all countries in a given year, and ϵ_{it} is the error term.

3.6 Diagnostic tests

Several important diagnostic tests are done to make sure the econometric approach is appropriate in order to validate the main results from the TWFE model. Although the Generalized Method of Moments (GMM) would have been the best way to deal with possible endogeneity issues, the fact that there were only 171 observations from 9 countries makes GMM estimation impossible, since GMM estimation needs a larger sample size for accurate results (Wooldridge 2019). Therefore, the study uses different robust estimation methods. The Pesaran CD test is used to detect cross-sectional dependence. If there is dependence, Driscoll-Kraay standard errors will be used to address the problem in all model estimations. Additional tests include the Breusch-Godfrey test for serial correlation and the Breusch-Pagan test for heteroskedasticity. Evidence of serial correlation can be observed in the Breusch-Godfrey test,

and the Breusch-Pagan test confirms the presence of heteroskedasticity in the residuals. These findings justify the choice of Driscoll-Kraay standard errors as the appropriate estimation method. Additionally, Variance Inflation Factor (VIF) analysis is done to check for multicollinearity. Afterwards, a sensitivity analysis check if the results were not affected by outliers.

CHAPTER FOUR: EMPIRICS AND DISCUSSION OF RESULTS

In this chapter, I will present the empirical findings of the research and provide a detailed interpretation and discussion of these results in the context of CIS's economic diversification efforts and the hypothesis outlined in Chapter 3. The chapter will first present descriptive statistics and data visualizations. This is followed by the panel data regression results and its interpretation. Finally, I will discuss the robustness of these estimates and explore the policy implications derived from the findings, and the research limitations.

4.1 Descriptive Analysis

The descriptive statistics provide valuable insights of the variables used in this analysis for the sample of nine CIS countries from 2002 to 2021.

4.1.1 Summary Statistics

The summary statistics for the final panel dataset, covering nine CIS countries from 2003 to 2021, are presented in Table №2.

Table 2. Summary of variables

Variable	N	Mean	St. Dev.	Min	Max
gdpcg	171	4.14	5.40	-14.78	33.03
hhi	171	0.36	0.05	0.27	0.50
oil	171	5.97	9.09	0.01	39.58
lfpr	171	62.63	9.68	30.07	82.10
gcf	171	26.55	7.33	8.93	57.99
hdi	171	0.73	0.06	0.58	0.84
log_exrate	171	3.01	2.48	-1.56	9.27

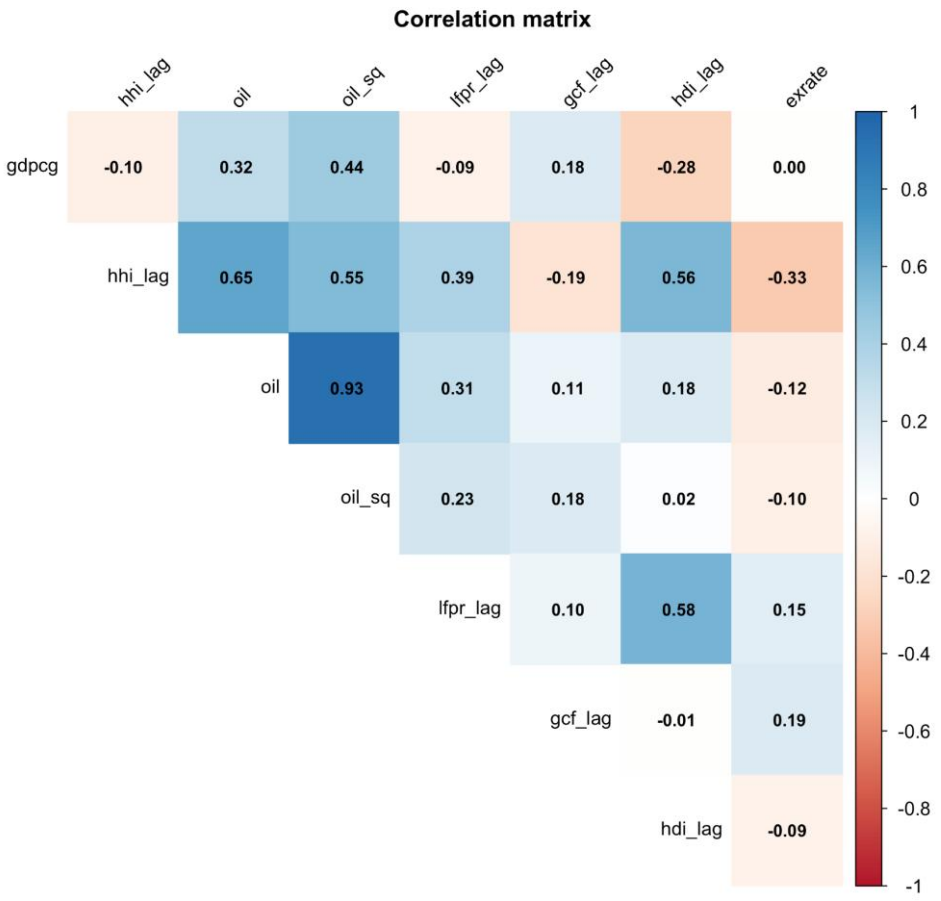
Source: Author's own calculations from World Bank data

The summary statistics reveal several key patterns in the data. GDP per capita growth rates vary greatly, ranging from -14.78% to 33.03%, with an average of 4.14% and a standard deviation of 6.83%. This wide range reflects the volatile nature of resource-dependent economies, particularly during commodity price cycles and economic crises. The extreme maximum value of 33.03% corresponds to Azerbaijan's oil boom period from 2005 to 2007, while the minimum represents crisis periods such as the 2008-2009 global financial crisis. The Herfindahl-Hirschman Index (HHI) values range between 0.268 and 0.499, with a mean of 0.355 and a standard deviation of 0.048. This relatively narrow range suggests that most CIS economies in the sample have similar levels of economic concentration, though there are significant differences. Oil rents as a percentage of GDP go from almost nothing (0.001%) to 39.58%, with a mean of 5.97% and a standard deviation of 9.09%. This wide range of differences is very important for figuring out how resource dependence affects the relationship between diversification and growth in different ways. The distribution is very skewed to the right, with most countries not being too dependent on oil and others like Azerbaijan, Kazakhstan, and Russia being highly dependent.

Labor force participation rates range from 30.07% to 82.10%, with a mean of 62.63%, reflecting differences in demographic structures and labor market conditions between countries. Gross capital formation ranges from 8.93% to 57.99% of GDP, with mean of 26.55%, indicating varying investment intensities across the region. The Human development index ranges from 0.581 to 0.839, indicating that the sample includes countries at different stages of development.

4.1.2 Correlation

Table 3. Correlation matrix



Source: Author's own calculations

The correlation matrix reveals several important relationships that help us understand how the variables interact. GDP per capita growth has a moderate positive correlation with oil rents (0.324) and a negative correlation with the HHI (-0.10), supporting both the resource curse at moderate levels and the diversification growth relationship.

The strong positive correlation (0.55) between HHI and oil rents confirms that greater oil dependence leads to less economic diversification, which is consistent with theoretical predictions about resource curse mechanisms. This relationship suggests that resource intensive economies have more concentrated economic structures.

The (-0.28) correlation between GDP growth and HDI is likely due to convergence effects, in which countries with lower initial development levels experience faster catch up growth. This pattern is consistent with conditional convergence theory, emphasising the importance of controlling for development levels in regression analysis.

4.1.3 Trends

The HHI trends show mixed patterns of diversification across the region. Some countries, particularly Tajikistan and Uzbekistan, had relatively stable and lower HHI values throughout the period, indicating more diverse economic structures. In contrast, Azerbaijan consistently had higher HHI values, indicating its heavy reliance on oil extraction, particularly during the 2005-2007 oil boom. Kazakhstan had moderate HHI levels with some volatility, reflecting its diverse resource base of oil, mining, and agriculture. Despite having a large and potentially diverse economy, Russia's HHI values ranged from moderate to high.

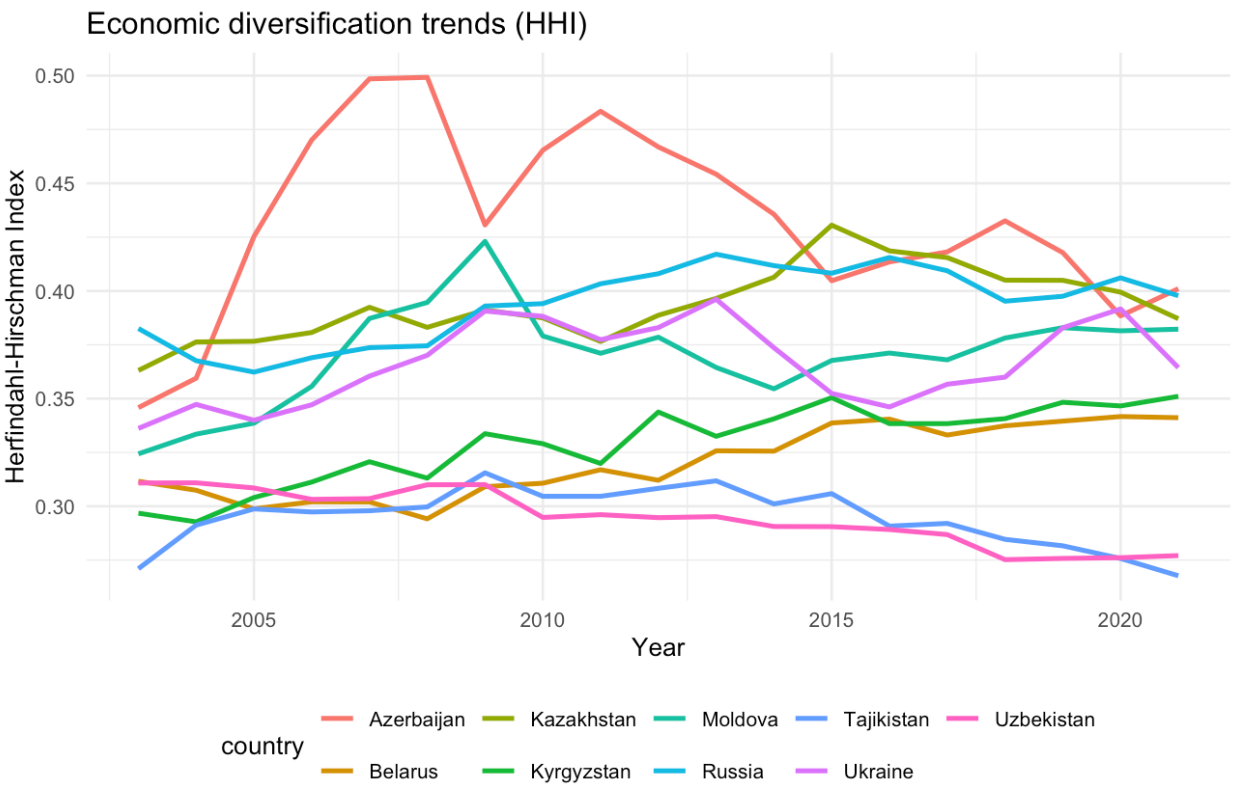


Figure 2. Economic diversification trends

4.2. Key empirical findings

The empirical analysis uses fixed effects panel data regression with two-way fixed effects (country and time) to account for unobserved heterogeneity. Three model specifications are estimated to assess the reliability of the findings and investigate various aspects of the diversification and growth relationship.

4.2.1 Diagnostic tests

Before showing the main results, a number of diagnostic tests are done to make sure that the econometric approach is correct. The Pesaran CD test is the first way to show cross-sectional dependence. The graph shows that $z = -2.93$, $p = 0.003$, which means that economic events in one CIS country tend to have an impact on other countries in the region. To solve this problem,

Driscoll-Kraay standard errors are used to estimate all models. These errors are strong against cross-sectional dependence, heteroskedasticity, and serial correlation.

The next tests are the Breusch-Godfrey and Breusch-Pagan tests. Some evidence of serial correlation can be seen in the Breusch-Godfrey test ($\chi^2 = 34.33$, $p = 0.017$), and the Breusch-Pagan test confirms heteroskedasticity ($BP = 65.88$, $p < 0.001$). This makes the choice of Driscoll-Kraay standard errors justifiable.

4.2.2 Multicollinearity

Table 4. Variance inflation factors

```
Variance Inflation Factors:
           hhi_lag                oil             oil_sq hhi_oil_interaction            lfpr_lag
              4.09             203.91              20.38             133.49                1.92
           gcf_lag            hdi_lag        log_exrate
              1.21               3.61               1.38
```

Variance inflation factor analysis shows high levels of multicollinearity among several important variables, which should be carefully thought through when interpreting the results. High VIF values (203.91) for oil rents show severe multicollinearity, and a VIF value of 133.49 for the HHI and oil interaction term also shows problematic collinearity. The VIF for the oil squared term is 20.38, which is higher than average. Even though these high VIF values are to be expected since they include interaction and quadratic terms with the same base variables, they suggest that the estimates of the individual coefficients should be reviewed carefully. While multicollinearity is most noticeable when it comes to variables that involve oil rents, this is because of the mathematical connections between oil, oil squared, and the interaction terms, not because of a problem with the data itself.

4.2.3 Regression results

Table 5. Regression results

	Dependent variable:		
	GDP per capita growth (%)		
	Non-linear	With Interaction	Linear
	(1)	(2)	(3)
hhi_lag	-15.955	-44.537***	3.161
	(19.915)	(8.912)	(24.192)
oil	-0.912***	-1.144***	0.463***
	(0.241)	(0.250)	(0.122)
oil_sq	0.031***	0.028***	
	(0.004)	(0.004)	
hhi_oil_lag		1.077***	
		(0.225)	
lfpr_lag	0.090**	0.083*	0.138**
	(0.042)	(0.046)	(0.056)
gcf_lag	0.108***	0.147***	0.155***
	(0.029)	(0.038)	(0.053)
hdi_lag	-64.123***	-86.072***	-56.657***
	(22.712)	(28.425)	(15.273)
log_exrate	-3.555***	-3.756***	-2.828***
	(0.460)	(0.487)	(0.834)
Country FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
SE Type	Driscoll-Kraay	Driscoll-Kraay	Driscoll-Kraay
Observations	171	171	171
R2	0.426	0.485	0.298
Adjusted R2	0.288	0.357	0.135
F Statistic	14.525*** (df = 7; 137)	16.035*** (df = 8; 136)	9.765*** (df = 6; 138)
Note:			*p<0.1; **p<0.05; ***p<0.01

Model 1. Basic non linear

There is some evidence for the resource curse hypothesis in the basic non linear model, but there is mixed evidence for the diversification and growth relationship. The HHI coefficient is negative (-15.955) but not statistically significant ($p > 0.1$). This means that the effect is not

statistically different from zero in this simple setting, even though the direction supports the diversification hypothesis.

There is a clear inverted U shaped relationship between the oil rents variables and the resource curse. It's very important to note that the coefficient for oil rents is negative (-0.912, $p < 0.01$), but the coefficient for oil rents squared is positive (0.031, $p < 0.01$). It was calculated that the turning point will happen when oil rents reach about 20.4% of GDP. In other words, this means that countries benefit from developing oil up to a certain point, but after that, relying too much on oil hurts growth.

Participation in the labor force has a positive and significant effect (0.090, $p < 0.05$), which shows that being active in the workforce helps the economy grow. Gross capital formation has a strong and positive relationship (0.108, $p < 0.01$), which shows how important investment is for growth. The HDI has a negative and highly significant coefficient (-64.123, $p < 0.01$), which shows that countries with lower levels of development at the start grow faster as they catch up. The exchange rate has a big, negative, and highly significant effect (-3.555, $p < 0.01$), which suggests that when currencies lose value, they slow down growth in economies that depend on imports mostly.

Model 2. Oil interaction

The HHI coefficient is negative and very significant (-44.537, $p < 0.01$), which strongly supports Hypothesis 1 that a more diverse economy leads to growth.

It's important to note that the interaction term (HHI × Oil) has a positive and highly significant coefficient (1.077, $p < 0.01$). This positive interaction shows that as oil dependence rises, the negative effect of economic lack of diversification becomes less harmful. This means that diversification helps countries that are not too dependent on oil the most and less as they become

more dependent on oil. In the oil rents, the linear term is always negative and significant (-1.144, $p < 0.01$), while the squared term is always positive and significant (0.028, $p < 0.01$). The relationships between control variables stay the same. The labor force participation rate stays positive and slightly significant (0.083, $p < 0.1$), the gross capital formation rate stays strong and positive (0.147, $p < 0.01$), the HDI rate stays negative and highly significant (-86.072, $p < 0.01$), and the exchange rate rate stays very negative (-3.756, $p < 0.01$).

Model 3. Linear

The HHI coefficient is positive (3.16), but it is not statistically significant ($p > 0.1$), so it does not show any benefits of diversification. Additionally, the oil rents coefficient is positive and very significant (0.463, $p < 0.01$), which goes against the resource curse theory. This finding shows why the non-linear specifications in Models 1 and 2 are so important for figuring out how things really connect.

Control variables show patterns that are mostly the same as other specifications, but the sizes of the patterns are sometimes different. The relationship between labor force participation and GDP is still positive and significant (0.138, $p < 0.05$), and the relationship between HDI and GDP is still highly significant (0.155, $p < 0.01$). The expected negative effect for HDI is seen (-56.657, $p < 0.01$), and the exchange rate has its usual negative effect (-2.828, $p < 0.01$), though it is smaller than in the non-linear models.

Control variables

The R-squared values for all three models are moderate, ranging from 0.298 to 0.485. This means that the specifications can explain between 29- 48% of the differences in GDP per capita growth between countries and time periods. This level of explanatory power is reasonable and

in line with how cross country growth regressions usually work, which have a lot of variation that cannot be explained because of how complicated economic development is.

Overall model significance F-statistics are large and highly significant across all specifications. This is strong evidence that the models explain meaningful variation in growth outcomes rather than just capturing noise. Using a two way fixed effects structure, which includes both country and time dummy variables, effectively addresses concerns about omitted variable bias caused by unobserved country specific traits and common temporal shocks that affect the whole region.

Before estimating, diagnostic tests were done to make sure that using Driscoll-Kraay standard errors was the right thing to do. These errors are strong even when cross-sectional dependence, heteroskedasticity, and serial correlation are found in the data. This choice in methodology makes sure that the reported standard errors and significance levels show the true statistical uncertainty in the estimates.

4.2.4 Robustness check

Table 6. Regression without outliers

	Dependent variable:		
	GDP per capita growth (%)		
	non linear	with interaction	linear
	(1)	(2)	(3)
hhi_lag	-5.911	-26.926	-2.198
	(21.646)	(16.603)	(21.251)
oil	-0.097	-0.249	0.177***
	(0.276)	(0.272)	(0.066)
oil_sq	0.007	0.004	
	(0.006)	(0.005)	
hhi_oil_lag		0.831***	
		(0.210)	
lfpr_lag	0.084***	0.079***	0.092***
	(0.022)	(0.027)	(0.024)
gcf_lag	0.089***	0.103***	0.088***
	(0.014)	(0.018)	(0.015)
hdi_lag	-39.912**	-55.752**	-35.783*
	(19.669)	(22.000)	(19.713)
log_exrate	-2.890***	-2.990***	-2.764***
	(0.731)	(0.601)	(0.856)
Country FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
SE Type	Driscoll-Kraay	Driscoll-Kraay	Driscoll-Kraay
Observations	161	161	161
R2	0.199	0.287	0.187
Adjusted R2	-0.010	0.094	-0.016
F Statistic	4.496*** (df = 7; 127)	6.325*** (df = 8; 126)	4.905*** (df = 6; 128)
Note:			*p<0.1; **p<0.05; ***p<0.01

The robustness check using the outlier adjusted sample mostly backs up the main results and gives us more confidence in how stable the results are. When Azerbaijan is taken out of the equation, mostly to get rid of the very high GDP growth rates during the oil booms, the main relationships still hold true, but with some important changes.

In Model 1, the HHI coefficient stays negative (-5.911) but loses statistical significance. This suggests that the basic effect of diversification may change when extreme cases are included.

The oil rents coefficients also lose statistical significance in the robustness check, but they still have the expected signs. This suggests that the resource curse relationship may be driven in part by Azerbaijan's periods of extreme oil dependence.

In Model 2, the interaction term (HHI × Oil) stays positive and highly significant (0.831, $p < 0.01$). This shows that the conditional diversification effects are strong across the larger CIS sample and not just a reflection of Azerbaijan's unique situation. This means that diversification benefits CIS economies that depend on natural resources in general.

The significance and signs of the control variables stay the same across the robustness specifications. The effects of labor force participation, gross capital formation, HDI, and exchange rates are still the same as they were in the main results. Overall model fit goes down a bit, where R squared goes from 0.485 to 0.287 in Model 2, but the fact that key relationships stay the same suggests that the results are not due to outliers and show real patterns across the CIS region.

4.3 Marginal effect

I calculated the marginal effect of HHI at different levels of oil dependence using the interaction model coefficients:

Table 7. Marginal effect

```
Marginal effect of HHI at different oil rent levels:
  At Oil Rents =  5%, Marginal Effect = -39.151
  At Oil Rents = 10%, Marginal Effect = -33.766
  At Oil Rents = 15%, Marginal Effect = -28.381
  At Oil Rents = 20%, Marginal Effect = -22.996
  At Oil Rents = 25%, Marginal Effect = -17.611
```

This study shows an relationship between diversification and growth using marginal effects analysis. Interestingly, the marginal effect of HHI is strongly negative (-39.151) at low levels of oil dependence (5% of GDP), showing that diversification is very good for growth. Nevertheless, this effect gets less negative as oil dependence rises, peaking at -17.611 when oil dependence is 25%.

Evidence strongly supports Hypothesis 2: as resource dependence rises, the marginal benefit of diversification falls. The traditional need to diversify becomes less important for countries that depend a lot on oil, but diversification is still a good idea in general.

4.4 Policy implications

The findings indicate that since diversification benefits vary greatly depending on the resource levels, CIS nations should tailor their diversification based on their oil dependency levels. Economic diversification and expanding product space should be the main policy focus for low oil dependent nations such as Moldova, Tajikistan, and Ukraine. While countries like Kazakhstan, Russia, and Azerbaijan, who are dependent on oil should keep their efforts at diversification to control commodity price volatility and stop Dutch disease effects. The nations approaching 20.4% oil dependence have a resource curse threshold beyond which further oil dependence will constrain over generating the growth for the countries. Thus, diversification is much more important. All CIS nations should keep strong investment rates regardless of their degree of oil dependence since gross investment is a policy priority for long term economic development and that promotes growth.

4.4 Policy implications: Kazakhstan

Kazakhstan is moderately dependent on oil, with oil rents making up about 15 to 20 percent of its GDP on average. This means that the marginal benefits of diversification range from -28.381

to -22.996. This leads to a number of specific policy implications that support the country's current development plans.

The marginal effects analysis shows that Kazakhstan is still getting a lot out of its efforts to diversify, even though the effects are still strongly negative at the level of resource dependence it is at now. This supports the government's effort for economic diversification through programs like the Strategic Plan Kazakhstan-2025 and the State Program of Industrial-Innovative Development. The country should continue and even further. They should focus on areas that can make the most of what they already have while also growing into new areas of production.

Kazakhstan is getting close to the critical resource curse threshold of 20.4%, which is a major concern for resource policy. At 15% oil dependence, the marginal benefit of diversification is still big -28.381, but as oil dependence rises to 20%, this benefit drops to -22.996, which is a 19% drop in the effectiveness of diversification. This finding strongly supports policies that aim to limit the oil sector's growth in relation to GDP, manage resource revenues through the National Oil Fund, and take steps to avoid Dutch disease effects such as real exchange rate appreciation, which harms the competitiveness of other sectors than oil.

Kazakhstan should keep focusing on building up its infrastructure and industrial capacity because there is a positive relationship between gross capital formation and growth. Its investment strategy should focus on projects that support the current economy and develop the skills necessary for future diversification, as the country is only moderately dependent on oil. This means further investment in transportation infrastructure to address landlocked challenges, industrial facilities that can process raw materials domestically, and technology and skills development that can support more sophisticated manufacturing and service sectors.

The study found that when the exchange rate goes down, it has large negative effects. This shows how important it is to keep the economy stable so that Kazakhstan's import dependent economy can keep growing. Because the country needs to import capital goods and intermediate inputs for industrial growth, monetary policy should continue to focus on keeping the currency stable. According to the marginal effects analysis, this stability is even more important as countries become more dependent on oil. At this point, diversification stops being as helpful and the economy is more likely to be affected by shocks from outside the country.

Kazakhstan should see its current situation as a chance to make things better in the future. As the country becomes more dependent on oil, the benefits of diversification are still big, but they are going down. This gives the country strong reasons to speed up its efforts to diversify before they reach a point where they are less effective. Setting clear goals to keep oil rents below 20% of GDP is something the government should think about. At the same time, they should build up institutions that can handle resource revenues and help the growth of sectors other than oil.

4.5 Limitations and future research

There are several limitations in the paper. First, there is a relatively small sample size of only nine CIS countries over 19 years and overall 171 observations. While CIS countries have shared institutional legacies and similar transition experiences, the results may not apply to resource rich economies in other areas with different historical, institutional, and developmental contexts, such as Middle East, Latin America, or Africa.

Because the cross-sectional dimension is limited, there are also fewer econometric techniques that can be used to deal with possible endogeneity issues and look into how treatment effects vary across countries. Using lagged variables and two-way fixed effects can help with some endogeneity problems, but larger samples and more advanced dynamic panel methods or instrumental variable approaches might give us more information. Additionally, due to low

numbers, GMM method could not be applied. Although the study period from 2002 to 2021 includes some important post-transition dynamics, it may not fully capture the long-term structural transformation processes that happen over many decades.

There are some problems with using the Herfindahl-Hirschman Index as the main way to measure economic diversification, even though it is a good method and is widely accepted. While the HHI measures concentration across broad sectoral categories, it might not fully show how complex and advanced economic activities are within sectors. Other ways of measuring, like the Economic Complexity Index, export product space analysis, or measures of productive capabilities, might help us understand more about the quality of diversification, not just the quantity of it. It's also possible that the HHI calculated from value-added shares doesn't take into account patterns of export diversification or integration into global value chains, both of which are becoming more and more important aspects of structural transformation.

The analysis also has to deal with data problems that are common to studies of economies in transition. For example, there are not that much of data points for some variables. For example, Turkmenistan and Armenia had to be omitted from the dataset to keep the balance, but it makes the results less representative of the whole CIS region. The study also doesn't directly look at how the quality of institutions changes over time within countries, which could play a big role in how diversification affects growth.

These problems could be fixed in future research in a number of ways. Including more countries and longer time series in the sample would increase statistical power and make it possible to use more advanced econometric methods. Comparative studies between developing regions that include both CIS countries and resource rich economies from other regions could test the hypothesis's external validity and find factors that are unique to each region that affect the process of diversification.

More complex information about the quality of diversification, not just the amount of it, could be gained by studying the role of export destinations, participation in the global value chain, and product sophistication. It might be easier to understand how diversification policies work or don't work in different situations if future studies include measures of institutional quality, governance, and political economy.

More research could be done in the future on how the link between diversification and growth changes as countries grow and their institutions get stronger. Researchers who look into how diversification plans affect environmental concerns could help resource-dependent economies make better policy decisions at a time when the world economy is moving towards lower carbon intensity.

Even though it has some flaws, this study strongly supports the idea that economic diversification can be beneficial in transition economies that depend on natural resources. These results show that while diversification generally helps growth in the CIS region, its small benefits depend a lot on how many resources are available. This goes against simple policy suggestions that say everyone should diversify, and it shows how important it is for resource-rich countries to have individual ways to economic growth. Countries that depend on oil have a tough time with growth because they need to balance the traditional need to diversify with the need to manage resource wealth effectively. The best strategies depend on how dependent the country is on resources and how well its institutions can handle resource wealth.

CHAPTER FIVE: CONCLUSION

This thesis looked into one of the most important problems that resource-based economies in the Commonwealth of Independent States (CIS) face when they try to grow: how to achieve long-term economic growth through diversification. At the start, the main question was whether economic diversification, which means moving away from centralized economic structures, leads to higher GDP per capita growth in this particular region. A panel dataset of nine CIS countries from 2002 to 2021 was used in this quantitative study to answer this question. This last chapter gives a clear answer to this question by summarising the main points of the arguments, restating the main findings, and thinking about what these results mean for the bigger picture. It also thinks about what this research has added to our knowledge of economic development in transition economies.

This thesis shows that the relationship between economic structure and growth is complex, especially in countries shaped by Soviet central planning and heavy reliance on natural resources. This was put to the test by looking at how economic concentration, as measured by the Herfindahl-Hirschman Index (HHI), affected GDP per capita growth. The study took into account important factors like investment, human capital, and, most importantly, oil dependence. The study also looked at how this relationship between diversification and economic growth might change depending on the amount of resources available. This was done to see how the benefits of diversification might change at different levels of resource wealth.

The study's empirical results give strong and clear answers to the research questions. Firstly , the results support the idea that economic diversification leads to higher GDP per capita growth in the CIS region. It was found that a lower HHI, which means a more diversified economy, has a statistically significant and positive effect on growth across a number of model

specifications. This finding backs up with the general idea that relying less on a small set of economic activities is a key part of long term prosperity.

Second, and this may be even more important, this thesis gives strong evidence that the benefits of diversification can depend on how much a country depends on oil. The study found a significant interaction between diversification and oil rents. It showed that countries that depend less on oil get the most growth from diversification. Diversification has less of a positive effect as reliance on oil grows. Diversification is always a good idea, but it's especially important for economies that are not yet deeply locked in the resource curse. This is because diversification has a smaller but still positive effect on growth. The study also confirmed the existence of an inverted U-shaped resource curse relationship. It found a turning point at about 20.4% of GDP in oil rents, above which relying too much on resources is bad for growth.

These results add to literature on economic development as they provide a quantitative and comparative analysis of diversification to the CIS countries, an area that has been missed in many cross-country studies of the resource curse. CIS countries have been grouped with other emerging markets in many studies, but this thesis provides specific evidence for the region. Overall, the results show that the general ideas of diversification and the resource curse do apply to the CIS. However, they also show the specific limits and conditions that affect how these transition economies develop.

This research has important policy implications, especially for a country like Kazakhstan, which was used as a main case study for this thesis. The results suggest that there is no such thing as a one size fits all approach to diversification. Kazakhstan's oil rents have put it close to the identified turning point in the past. The results are both proof of its current strategy to diversify and a clear warning. The evidence strongly supports the government continuing to invest in non extractive sectors, as doing so has big benefits for growth. At the same time, it shows how

important it is for the country to carefully manage its existing resources so that its oil sector does not become a a curse on long term growth. The insight is that strategies need to be tailored to each country's level of resource dependence. This is so that the long-term goal of building a diversified economy can coexist with the immediate challenge of managing resource revenues effectively.

Finally, it's important to note that this study has some limitations. The analysis was limited by the small sample size of nine countries, which means that the results cannot be generalized for bigger picture. Furthermore, the Herfindahl-Hirschman Index is a common way to measure diversification, but it is also a pretty simple way to do it. It might not show the full complexity of structural economic change, like when new skills are developed or products become more complex.Thus, these limitations could be fixed in future research. If there is enough data, adding a longer time series or more countries to the dataset would make the results more reliable.

This research adds several contributions to the diversification literature. First, it gives quantitative panel data analysis for the CIS region using interaction terms and marginal effect calculations. Second, it sets clear limits and small benefits that allow tailored policy advice instead of broad support for diversification. Third, it shows that the effects of the resource curse are very different in different countries based on where they are in relation to the 20.4% threshold. These additions help us learn more about how resource endowments affect how well diversification strategies work.

In the end, this thesis shows that the CIS countries' ability to diversify their economies is largely linked to their ability to achieve long term economic growth. It shows that this path is still necessary, even though it is made harder by the strong influence of resource wealth. This study helps policymakers and scholars understand and work towards a more stable and prosperous

future for this important region by showing not only that diversification is important, but also how and when it is most important.

BIBLIOGRAPHY

Abderrahim, B., Belbali, A., Gasmi, M. A., Ouldbahammou, S., Yousfat, A., & Oukacha Azizi, A. (2024). Measuring and analyzing economic diversification using the Herfindahl-Hirschman Index. European Economic Letters, 14(3), 3025–3037. https://eelet.org.uk/index.php/journal/article/view/2078

Akhmedov, E. (2017). Economic diversification issues in oil-exporting countries: Case of the Republic of Kazakhstan.

Allen, Robert C. 2011. "1. The Great Divergence." In *Oxford University Press eBooks*, xvi–13. https://doi.org/10.1093/actrade/9780199596652.003.0001.

Aubakirova, G. M., Zhuparova, A. S., & Kozhakhmetova, A. K. (2025). Diversification of the Kazakhstan economy through global value chains: A quantitative assessment. Economy: Strategy and Practice, 25(1), 103–119.

Chuku, C. N. (2021). The impact of economic diversification on economic development in oil-producing middle-income countries.

Delechat, Corinne C., Giovanni Melina, Monique Newiak, Chris Papageorgiou, Ke Wang, and Nikola Spatafora (2024). *Economic Diversification in Developing Countries – Lessons from Country Experiences with Broad-Based and Industrial Policies.* Washington, DC: International Monetary Fund.Esquenazi,

Gelb, A. (2010). Economic diversification in resource-rich countries. International Monetary Fund.

Gleason, G. (2003). Markets and politics in Central Asia: Structural reform and political change. Routledge.

Hartmann, D., & Pyka, A. (2013). Innovation, economic diversification and human development. FZID Discussion Paper, No. 65-2013.

Hidalgo, C. A., & Hausmann, R. (2009). The building blocks of economic complexity. Proceedings of the National Academy of Sciences, 106(26), 10570–10575.

Hausmann, R., & Hidalgo, C. A. (2011). The network structure of economic output. Journal of Economic Growth, 16(4), 309–342.

Hausmann, R., Hidalgo, C. A., Stock, D. P., & Yıldırım, M. A. (2020). Implied comparative advantage (CID Faculty Working Paper No. 276, Revised July 2020). Center for International Development, Harvard University. https://growthlab.cid.harvard.edu/publications/implied-comparative-advantage

Hausmann, R., Taniparti, N., Brenot, C., Barrios, D., Soylu, C., El Houda, R., Vashkinskaya, E., Belostecinic, F., & Henn, S. (2023). A growth diagnostic of Kazakhstan (CID Faculty Working Paper No. 427). https://growthlab.cid.harvard.edu/files/growthlab/files/wp427.pdf

Harvard Growth Lab. (2023). Kazakhstan | The Growth Lab. Harvard University. https://growthlab.hks.harvard.edu/applied-research-projects/kazakhstan

Hidalgo, C. A., Klinger, B., Barabási, A.-L., & Hausmann, R. (2007). The product space conditions the development of nations. Science, 317(5837), 482–487.

Johnson, P., & Papageorgiou, C. (2020). What remains of cross-country convergence? Journal of Economic Literature, 58(1), 129–175. https://doi.org/10.1257/jel.20181207

Kalyuzhova, Y., Pemberton, J., & Mukhamediyev, B. (2004). Natural resources and economic growth in Kazakhstan. In G. Ofer & R. Pomfret (Eds.), The economic prospects of the CIS – Sources of long-term growth (pp. xx–xx). Edward Elgar.

Lashitew, A. A., Ross, M. L., & Werker, E. (2020). What drives successful economic diversification in resource-rich countries? The World Bank Research Observer, 35(2), 1–33. https://doi.org/10.1093/wbro/lkaa001

OECD. (2024). Diversifying Kazakhstan's exports: Institutions, policies, infrastructures. OECD Publishing.

Petrovskaya, M. V., Tsypin, A. P., & Sharykina, E. A. (2023). Comparative statistical analysis of industrial development in post-Soviet countries for the period 1990–2021. Izvestiâ Saratovskogo Universiteta. Seriâ Èkonomika. Upravlenie. Pravo, 23(4), 411–419.

Romer, P. M. (1986). Increasing returns and long-run growth. Journal of Political Economy, 94(5), 1002–1037.

Romer, P. M. (1990). Endogenous technological change. Journal of Political Economy, 98(5, Part 2), S71–S102. https://www.jstor.org/stable/2937632

Sachs, J. D., & Warner, A. M. (1995). Natural resource abundance and economic growth (NBER Working Paper No. 5398). National Bureau of Economic Research. https://doi.org/10.3386/w5398

Sattar, S., & Shiells, C. R. (2004). The low-income countries of the Commonwealth of Independent States: Progress and challenges in transition. International Monetary Fund.

Schumpeter, J. A. (1939). Business cycles: A theoretical, historical, and statistical analysis of the capitalist process. McGraw-Hill.

Solow, R. M. (1956). A contribution to the theory of economic growth. The Quarterly Journal of Economics, 70(1), 65–94.

Song, M. Y., & Mah, J. S. (2020). The Belarusian model of economic transition and industrialization. Perspectives on Global Development and Technology, 19(3), 359–382.

Souza, L. V. de, Havrylyshyn, O., Pastor, G., & Klingen, C. (2006). Return to growth in CIS countries: Monetary policy and macroeconomic framework. Springer.

TRACECA. (2025) “About the Republic of Kazakhstan.” https://tracecaorg.org/kg/countries/kazakhstan/about-the-republic-of-kazakhstan/.

United Nations, Vos, R., & Koparanova, M. (2011). Globalization and economic diversification: Policy challenges for economies in transition (1st ed.). Bloomsbury Publishing.

Ullah, S., Akhtar, P., & Zaefarian, G. (2018). Dealing with endogeneity bias: The generalized method of moments (GMM) for panel data. Industrial Marketing Management, 71, 69–78. https://doi.org/10.1016/j.indmarman.2017.11.010

Van der Ploeg, F. (2011). Natural resources: Curse or blessing? Journal of Economic Literature, 49(2), 366–420.

World Bank. (2017). Economic diversification: Guidance note. Trade and Competitiveness Global Practice.

World Bank. Kazakhstan overview. https://www.worldbank.org/en/country/kazakhstan/overview

Wooldridge, J. M. (2019). Introductory econometrics: A modern approach (7th ed.).

Zhukov, S. V., Kopytin, I. A., Maslennikov, A. O., & Sinitsyn, M. V. (2023). Oil exporting countries: Analysis of economic growth opportunities through export diversification in a low-carbon world. Studies on Russian Economic Development, 34(5), 573–582.

APPENDIX

Table 8. Pesaran CD test

```
data:  gdpcg ~ hhi_lag + oil + oil_sq + lfpr_lag + gcf_lag + hdi_lag +     log_exrate
z = -2.9274, p-value = 0.003418
alternative hypothesis: cross-sectional dependence
```

Table 9. Breusch- Godfey test

```
data:  test_formula
chisq = 34.334, df = 19, p-value = 0.01678
alternative hypothesis: serial correlation in idiosyncratic errors
```

Table 10. Breusch-Pagan test

```
        Breusch-Pagan test

data:  test_formula
BP = 65.884, df = 7, p-value = 9.987e-12
```