The Economic Complexity of Kazakhstan: A Roadmap for Sustainable and Inclusive Growth

Ricardo Hausmann, Douglas Barrios, Clement Brenot,
Nikita Taniparti, Eric Protzer, Sophia Henn

CID Faculty Working Paper No. 426
February 2023

Working Papers
Center for International Development
at Harvard University

Table of Contents

Executive Summary

Since the end of the 1990s, Kazakhstan has relied on oil and gas as the main drivers of economic growth. While this has led to rapid development of the country, especially during years of high oil prices, it has also subjected the economy to more severe downturns during oil shocks, bouts of currency overvaluation, and procyclicality in growth and public spending.

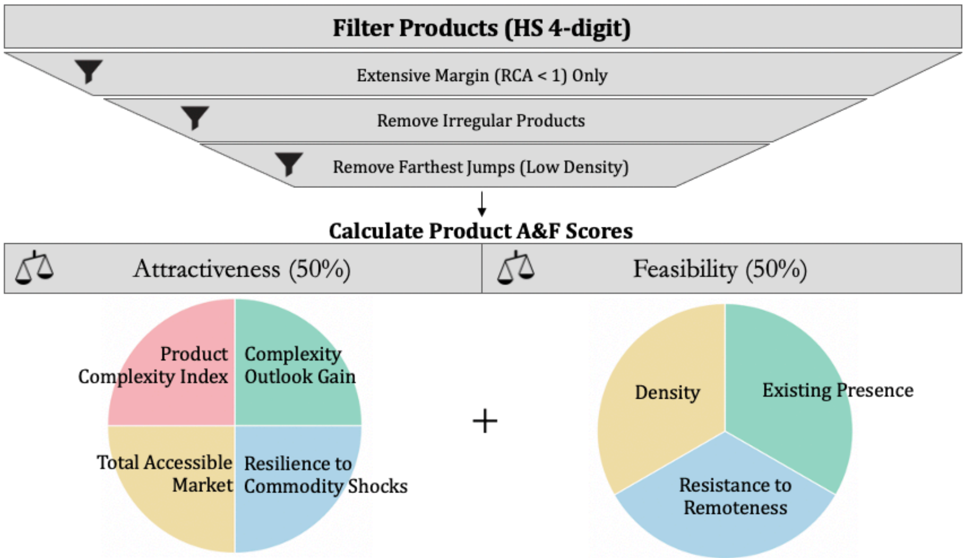
Stronger economic diversification has the potential to drive a new era of sustainable growth by supporting new sources of value added and export revenue, creating new and better jobs, and making the economy more resistant to fluctuations in oil dynamics. However, repeated efforts to stimulate alternative, non-oil engines of growth have so far been inconclusive.

This report introduces a new framework to identify opportunities for economic diversification in Kazakhstan. This framework attempts to improve upon previous methods, notably by building country and region-specific challenges to the development of the non-oil economy directly into the framework to identify feasible and attractive opportunities. These challenges are presented in detail in the *Growth Diagnostic of Kazakhstan* and are summarized along three high-level constraints: (i) an uneven economic playing field dominated by government-related public and private-entities; (ii) difficulties in acquiring productive capabilities, agglomerating them locally, and accessing export markets; and (iii) ongoing macroeconomic factors lowering external competitiveness lower and making the economy less stable.

Our approach applies the economic complexity paradigm to identify what specific products and industries are most feasible for diversification, based on the existing productive capabilities demonstrated in the economy. We examine Kazakhstan's economic complexity at the national but also subnational levels, highlighting the heterogeneity of export baskets across regions that makes an analysis of opportunities at the subnational level essential.

We formulate diversification strategies at the level of six regional groupings ("macro regions") based on their respective productive structures, chiefly using a local dataset on regional exports. Seven criteria are identified to evaluate the attractiveness and feasibility of 1,200 products as possible export opportunities for each macro region. These criteria are a product's existing presence in the macro region's export basket, its relatedness to the macro region's current productive capabilities, tolerance to geographic distance in global trade, resilience to oil price shocks, total global addressable market, as well two economic complexity measures (Product Complexity Index and Complexity Outlook Gain). Two sets of export opportunities are identified per macro region: new or nascent products to be introduced, and existing products to be scaled up. As an example, Fig. 1 illustrates the framework for identifying new or nascent opportunities.

Figure 1.Framework for Identifying New and Nascent Export Opportunities ("Extensive Margin")



Note: Framework for identifying existing products to be scaled up looks similar, but with a lower number of factors for consideration, as some do not apply to products that are already exported by a given region
Source: Own elaboration.

In order to incorporate tradable services into the analysis, an additional set of opportunities are identified in a similar manner on the basis of an employment dataset for the cities of Astana and Almaty, where opportunities in tradable services may be most likely to thrive.

In total, 172 export opportunities are identified across the six macro regions, and an additional 60 industry-based opportunities are identified for Astana and Almaty City.[1] Overall, identified product opportunities are diverse but most numerous in the sectors of chemicals (fertilizers, plastics, organic chemicals), transportation (train parts and equipment, aircraft and defense equipment), and processed agricultural products (meat, dairy, others). Additional employment-based opportunities identified in Astana and Almaty include various information, financial, and engineering services.

Between macro regions, identified opportunities are distinct and often unique. For example, the macro region encompassing regions in the north, whose current export basket is characterized mostly by grains and iron ores, contains opportunities in various agricultural products as well as transportation equipment, inorganic chemicals, and agricultural machinery. The macro region encompassing the southern regions, whose export basket includes vegetables, animal products, uranium and some electronics, contains new opportunities in various chemicals, plastics and construction materials.[2]

In the conclusion, we discuss two policy avenues for actioning on identified opportunities, suggesting the formation of public-private partnerships in the form of productivity taskforces as well as targeted investment promotion strategies that have proven effective in other efforts towards diversification around the world.

[1] Note that Almaty City is distinct from the region of Almaty, which are separate administrative divisions in Kazakhstan.
[2] The full results by macro region are included in Section 6 and Section 7 of this report, in addition to the online tool published here: https://public.tableau.com/app/profile/growth.lab.kazakhstan/viz/IndustryTargetingDashboardKazakhstan/Dashboard.

1. Introduction

This Economic Complexity Report was drafted as part of a research engagement between the Growth Lab at Harvard University and the Astana International Financial Centre (AIFC) during 2021 and 2022. The purpose of the engagement was to formulate evidence-based policy options to address critical issues facing the economy of Kazakhstan through innovative frameworks such as growth diagnostics and economic complexity. This report speaks to findings from the *Growth Diagnostic of Kazakhstan* on economy-wide challenges to growth and diversification to identify attractive and feasible export opportunities for the country. A companion online tool to further explore the results in this report is also available at:

https://public.tableau.com/app/profile/growth.lab.kazakhstan/viz/IndustryTargetingDashboardKazakhstan/Dashboard.

Kazakhstan's recent growth trajectory—during which real GDP per capita multiplied by 2.5x—can be divided into two periods that underscore how development of the country has been correlated with oil and gas dynamics. The early and mid-2000s characterized by the global commodity supercycle led to an expansion of the economy upwards of 8% annually, with a mild slowdown during the global financial crisis. In 2014, Kazakhstan's growth slowed with the collapse of commodity prices, and alternative engines of growth have not been strong enough to fend against volatility since. These trends, along with growing uncertainty in the long-run demand of oil and gas, continue to highlight the limitations of relying on natural resources to drive development.

As in the case of other major oil producers, diversification of Kazakhstan's non-oil economy is a critical pathway to drive a new era of sustainable and inclusive growth and mitigate the impacts of commodity price shocks on the country's economy. Development of non-oil activities has been a policy objective of the government of Kazakhstan for some time, but previous efforts for target sectors have failed to generate sufficient exports and investments to produce alternative engines of growth. This report addresses the shortcomings of previous attempts at diversification and introduces a new framework to identify opportunities that incorporates the economic complexity paradigm and considers the country's economy-wide constraints to growth as detailed in *A Growth Diagnostic of Kazakhstan*. Using this framework, a specific set of attractive and feasible diversification opportunities are identified for groups of regions and main cities in Kazakhstan.

This report is organized into 7 sections following this brief introduction as Section 1. Section 2 describes the Economic Complexity methodology and related measures used in this report. Section 3 presents an overview of Kazakhstan's diversification trajectory and related policy efforts. Section 4 introduces the motivation for pursuing a sub-national complexity analysis of Kazakhstan and necessary adjustments to the data and methodology to fulfill the exercise. Section 5 presents an overview of Kazakhstan's economic complexity at the sub-national level. Section 6 describes the framework to identify export opportunities at the sub-national level and presents detailed results. Section 7 describes the framework to identify additional opportunities in services for the cities of Astana and Almaty and presents detailed results. Section 8 concludes with a brief discussion on pathways for actioning on identified diversification opportunities.

2. The Economic Complexity Paradigm

The theory of economic complexity introduced by Hausmann, Hidalgo et al. (2011) is based on the realization that the development of products and services not only requires raw materials, labor, and machinery, but also tacit knowledge ("knowhow" or "productive capabilities") of how inputs are combined to produce outputs and run business operations. This tacit knowledge tends to be the limiting factor for diversifying economic activities because it is the most difficult component of production to transfer. Whereas many other inputs to production—including materials, tools, and blueprints—are relatively easy to trade and transfer, tacit knowledge can only be acquired through experience. Moreover, modern production requires far more knowhow than any single individual can acquire. Therefore, tacit knowledge is necessarily spread across many individuals who coordinate across teams and organizations.

Some products and services incorporate large amounts of knowhow and knowhow that are valuable for multiple uses. In contrast, other products and services incorporate much less knowhow or knowhow that is not transferable for other valuable uses. As an analogy, different products and services can be understood as "words" whose production requires "letters" (knowhow), like in a game of *Scrabble*. The production of long and sophisticated words requires many letters, including some high-value letters, while few are needed to generate short and simple words. The knowhow embedded in locations varies in terms of type and quantity. That is, some locations have many diverse letters, which they can use in many combinations to make many different and valuable words, while others have few letters and letters with limited uses, which limits the possibility of creating new words. The differences in productive capacities brought by uneven endowments of letters are further amplified by the fact that the number of words that can be constructed increases exponentially as new letters are added.[3]

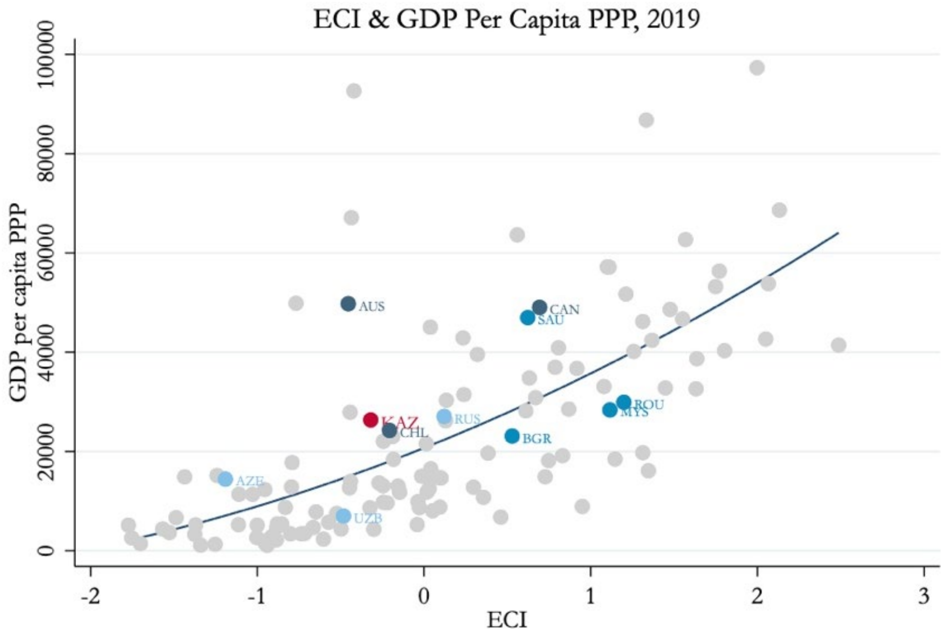
Applying this analogy to global trade, locations ultimately develop the products and services (words) that their knowhow (letters) can support. The Economic Complexity paradigm is focused on translating observations about locations and the products/services they trade into insights on the knowhow and productive capabilities present in their economy. The paradigm observes patterns about locations such as the number of products they export and the rarity of these products in global trade; the paradigm also realizes relationships between products such as the tendency for two products to be exported together by the same location. Using only data on the products and services traded by countries over time, the Economic Complexity paradigm builds measures that capture the complexity of economies (Economic Complexity Index, ECI) and products (Product Complexity Index). Locations with a large amount of diverse and specialized knowhow are the most complex economies and characterized by a high ECI. Locations that support a less diverse range of knowhow are the least complex economies and characterized by a low ECI.

Given that economic complexity reflects the amount of knowhow that is embedded in the productive structure of an economy, it is not surprising to find a strong correlation between country ECIs and their income per capita. Fig. 2.1 demonstrates this relationship.

---

[3] Thus, for example, in the English language, with 1 letter, "a", one word can be formed of up to 1 letter; with 3 letters, "a", "c" and "t", up to 4 words can be formed of up to 3 letters ("a", "at", "cat" and "act"); with 4 letters, "a", "c", "t" and "r", 9 words can be formed of up to 4 letters ("a", "at", "cat", "act", "rat", "car", "art", "tar" and "cart"); and with 10 letters, "a", "c", "t", "r", "o," l "," g "," s "," n "and" i ", 595 words can be formed of up to 10 letters.

Figure 2.1



Source: Atlas of Economic Complexity and World Bank WDI.

Hausmann, Hidalgo et al. (2014) also found that the prediction errors in Fig. 2.1—i.e., the difference between a country's actual income levels and those predicted by its ECI—are predictive of future growth dynamics. Countries with an economic complexity greater than expected given their level of income tend to grow faster than countries that display a level of income that is higher than expected for their current level of economic complexity. In other words, countries positioned below the regression line are often poised to enter periods of sustained growth, because if key constraints (such as infrastructure, access to financial capital, or institutional gaps) can be overcome, they can translate their existing stock of knowhow into higher output. Meanwhile, locations above the regression line may be in a more precarious position in terms of long-term growth as they may be benefitting from a temporary positive shock. If this boom is not leveraged to increase the complexity of the economy to a level consistent with the current level of income, they run the risk of having their level of income fall toward the regression line when the boom ends.

The implication for developing countries is that long-term growth and corresponding improvements in wellbeing tend to require a process of structural transformation where economic stakeholders gradually gain productive capabilities. This allows the revealed comparative advantages of the economy to evolve and diversify over time. Countries that have transitioned from low-income to high-income economic systems have tended to diversify from primarily agricultural production into particular types of labor-intensive manufacturing (like garments) and onward to more sophisticated manufacturing and tradable services. As they grow, they do not abandon most of the economic activities of the past but rather become more productive in those activities as they add new industries. This diversification process leads to rising wages across both old and new industries and makes countries more resilient to a variety of shocks—whether natural, macroeconomic, financial, and technological—as economic activity and jobs are less concentrated and therefore less vulnerable to a single shock.

The implications for countries like Kazakhstan that find themselves above the regression line are noteworthy. Many of these countries benefit from substantial resource wealth, while some—but not all—benefit also from very strong institutions that diminish the negative impacts of resource wealth, including a tendency toward inequality and boom-and-bust cycles. However, these countries stand to benefit from recognizing the risk inherent with resource-driven wealth: long-term economic performance will be driven by the exogenously determined value of the resource and it will likely be hard to reallocate productive resources to alternative engines of growth. Additionally, some of these countries struggle with the subnational implications associated with these types of economies. For instance, stabilization mechanisms that work well at a national level may not work as well at a subnational level. They also often struggle sharing the benefits of that wealth, as in many instances rents from natural resources do not always translate into long-term material benefit for the locations they were extracted from. Finally, countries endowed with natural resource wealth face challenges for diversification of job opportunities that derive from the distorting macroeconomic influences that natural resources can have. Most commonly, this wealth can appreciate national currencies, which can crowd out the emergence of other economic opportunities that would provide more and better jobs. Such countries may want to actively pursue diversification for the benefits of resilience and inclusiveness, but the nature of their diversification paths will likely be idiosyncratic.

Another critical theoretical foundation of economic complexity was introduced by Hausmann and Klinger (2006). They showed that the probability that a location develops the ability to produce a new product varies based on the set of products that it already produces. This allowed for the measurement of the similarity between products based on their shared capabilities. Based on this pattern, they introduce a measure of proximity between products as the minimum conditional probability that they are co-exported by the same country. The collection of all product proximity pairs can be visualized in a network, known as the Product Space, and used to study the productive structure of locations.

The structure of the Product Space and a location's position within it is crucial as it affects the ability of locations to move into new products. A highly connected position in the Product Space reflects relatively easier paths to diversification than a sparse position. Hausmann and Klinger (2006) find that the Product Space is highly heterogeneous: some sections are composed of densely connected groups of products whereas others are more loosely connected. This heterogeneity has significant implications for the speed and patterns with which structural transformation takes place; thus, the ability of countries to diversify into products that are more complex is crucially dependent on their starting location in the Product Space. The complete Product Space and Kazakhstan's position in the space in 2019 are shown in Fig. 2.2 and 2.3.

The position of a country in the Product Space captures information regarding both the productive knowledge that it possesses and the capacity to expand that knowledge by moving into other nearby products. The strategic positioning of a place in the Product Space can be leveraged as an insightful tool for formulating economic diversification strategies.

Figure 2.2

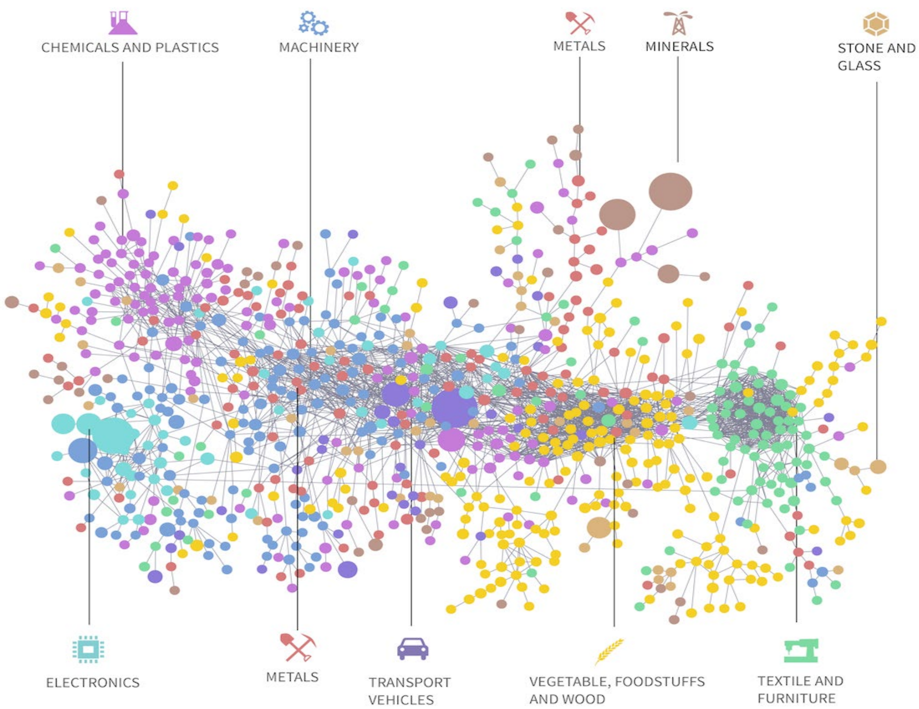
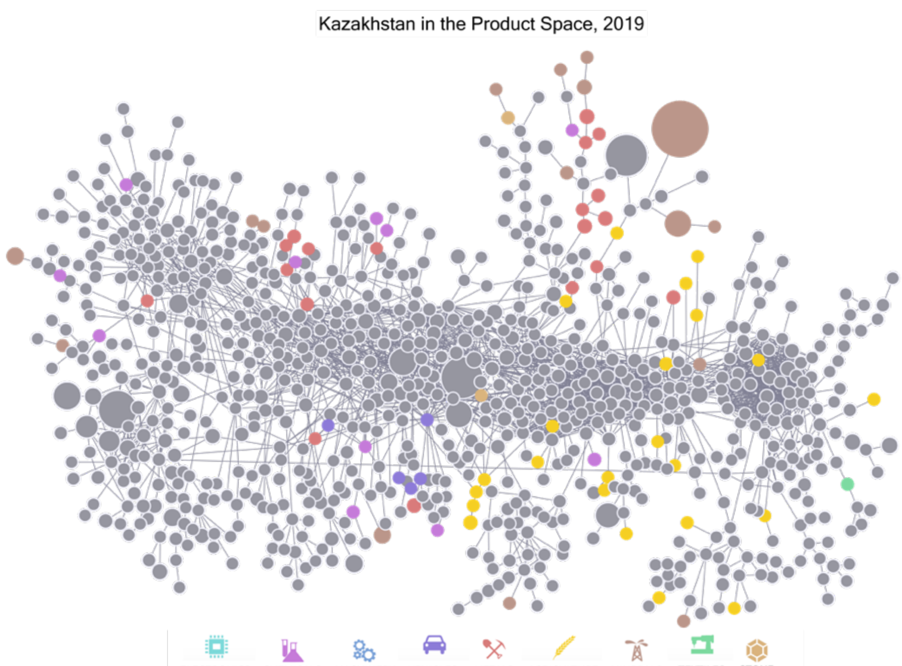


Figure 2.3



The quantitative methodology to construct the Product Space will be iterated upon in this report in the context of a complexity analysis at the subnational level in Kazakhstan. This report will address the tendency of possible diversification opportunities in the Product Space to be extremely sparse for resource-intensive countries like Kazakhstan and develop an alternative measure of density that improves upon its predictive power in anticipating the products for which a country has a comparative advantage.

Box: Relevant Concepts in Economic Complexity

The following glossary is meant to provide an intuitive explanation for several measures within the framework of economic complexity employed in this report. Additional mathematical detail can be found online at the Growth Lab website: www.atlas.cid.harvard.edu/glossary.

* ***Revealed Comparative Advantage (RCA)*****:** This is a location-product measure that captures the relative prevalence of a product in a location. Following the methodology of Balassa (1964), it is usually calculated as the ratio between the proportion of the product in the export basket of a location and the proportion of the product in world trade. If this relationship is greater than one, the location has a "revealed comparative advantage" in that product, which is equivalent to saying that the location produces the product more intensively than the rest of the world.

* ***Matrix Country Product (M_{cp}):*** A binarized matrix of the location-product RCA values, where M_{cp} = 1 when RCA is greater than 1 and 0 otherwise. A location that has M_{cp} = 1 in a product may also be simply referred to as "having an M_{cp}" in the product.

* ***Diversity:*** A location-specific measure that indicates the number of products for which that location has a comparative advantage (RCA > 1 or equivalently M_{cp}=1).

* ***Ubiquity*****:** A product-specific measure that indicates the number of locations that have a comparative advantage in that product (RCA > 1 or equivalently M_{cp}=1).

* ***Product Complexity Index (PCI)*****:** This is a product-specific measure that ranks the diversity and ubiquity of the productive knowledge required for production of the product. It is calculated through a recursion of the measures of diversity and ubiquity, examining the average diversity of countries that make the product, as well as the average ubiquity of the other products that these countries make.

* ***Economic Complexity Index (ECI)*****:** A location-specific measure that captures how diversified and complex a location's export basket is. Locations that are home to a great diversity of productive knowhow, particularly complex specialized knowhow, are able to export a great diversity of products, including highly unique products. ECI is found to be highly predictive of current income levels and future growth dynamics. It is calculated through a recursion of the measures of diversity and ubiquity and can also be expressed as the average PCI of the products exported by the location.

* ***Product Proximity*****:** A product-product measure indicating how similar is the knowhow required to produce each product. It is calculated as the minimum conditional probability that the two products are co-exported with a comparative advantage by the same location, observed using global data. These product proximities are the basis on which the network of product relationships is built (Product Space).

* ***Density*****:** A location-product measure indicating how proximate (in knowhow) any given product is to the knowhow currently exhibited in the economy by the set of products already exported with a comparative advantage. It is calculated as the sum of the proximities between the given product and every product exported with a comparative advantage by the location, divided by the sum of the proximities between the given product and all products. Density is sometimes conveyed as a *Distance* measure, which is equivalent to 1 – Density.

Box: Relevant Concepts in Economic Complexity

* ***Complexity Outlook Gain (COG)***: A location-product measure that quantifies the extent to which adding a new product to the current export basket can open up opportunities to diversify into more and more complex products. A high product COG implies that the product is in the vicinity of many potential new products and/or of products that are more complex, while a low product COG implies that it is near many existing products and/or new products that are less complex. It is calculated as the weighted sum of PCIs for products not currently exported with a comparative advantage from a location, weighted by each product's proximity to the new product in question.

* ***Complexity Outlook Index (COI)***: A location-specific measure evaluating how overall well-positioned a location is to diversify into higher complexity products. A high location COI implies that the location has a smaller gap in productive knowhow to high complexity products, while a low location COI means that the location has a larger gap in productive knowhow to high complexity products, implying that achieving them may be more difficult. It is calculated as the weighted sum of PCIs for products not currently exported with a comparative advantage from the location, weighted by each product's distance (density) from the location's current knowhow.

3. Kazakhstan's Diversification Challenge

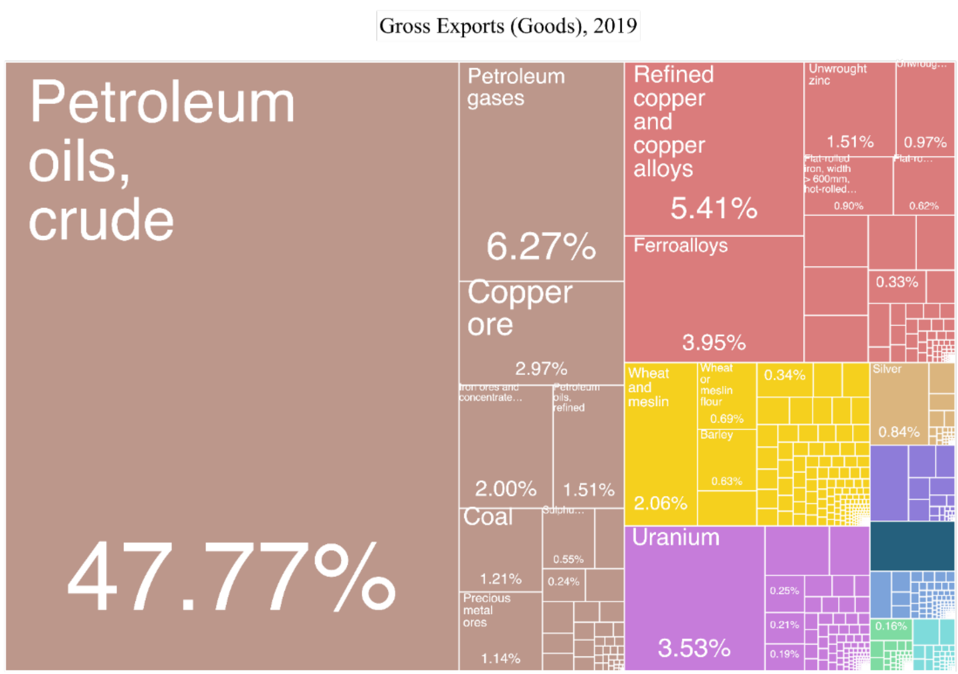
This section overviews Kazakhstan's historical trajectory of exports, national-level performance on measures of economic complexity, and limitations of several previous diversification plans in stimulating non-oil engines of economic growth.

Export Profile

Kazakhstan witnessed a growth boom in the early 2000s that was primarily driven by mineral exports. The country underwent a major economic expansion for most of the previous two decades, with real per capita GDP growth in constant USD averaging 6.4% annually between 2000-2014.[4] This growth was spurred by an investment and export boom of oil and mineral exports due to favorable terms of trade during the global commodity supercycle. Overall export growth of goods[5] between 2000-2014 was 462%,[6] but most of this growth was concentrated in the oil and mineral sector. Since 2000, mineral products — particularly crude petroleum oils and ores — have comprised between 51.9% to 76.1% of Kazakhstan's annual gross exports of goods.

The high-growth years were volatile and the nonmineral economy failed to expand at this time into strong complementary engines of growth. At the outset of the 2000s, nonmineral products made up 47.98% of total goods exports. Between 2000-2014, nonmineral exports of goods expanded by approximately 20% each year and peaked at US$ 21.7 billion in 2012, followed by a period of stunted export growth. By 2019, gross nonmineral exports of goods stood at US$ 16.2 billion, representing 34.76% of Kazakhstan's total goods exports (Fig. 3.1).

Figure 3.1



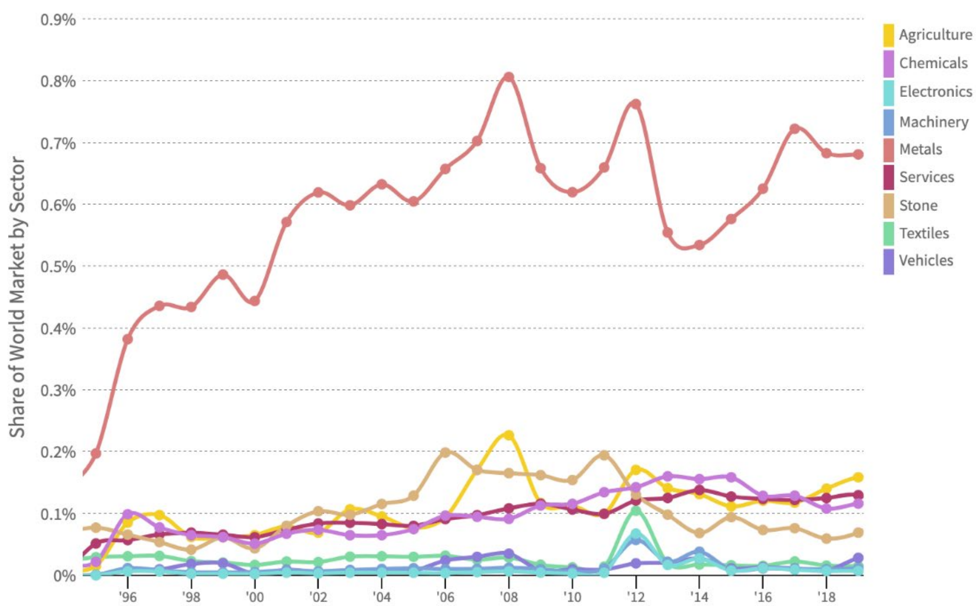
Source: Atlas of Economic Complexity.

[4] World Bank, "World Development Indicators (WDI)."
[5] Overall export growth inclusive of services was 479% between 2000-2014.
[6] The Growth Lab at Harvard University, "The Atlas of Economic Complexity."

This growth of Kazakhstan's nonmineral economy can be largely traced to the expansion of metal and chemical exports through the early 2010s and the continuous growth of services through the latter half of the most recent decade. Increased exports in refined copper, copper alloys, and ferroalloys were responsible for much of the growth in metals. Exports of these products grew by 489% from 2000-2014 and presently occupy over half of the metal industry's overall exports. The sizable growth in metals was accompanied by growth in chemical exports—the large majority of which has been in uranium, whose exports grew twenty-fold in the same period—in addition to a variety of raw and unprocessed agricultural products. Overall, Kazakhstan's export growth between 2000-2014 outperformed many of its regional and global peers, but most of this was driven by mineral exports.

Thus, export growth in the nonmineral economy was concentrated in a few products in the metal, agriculture, and chemical industries, and nonmineral exports are still a relatively small share of the total export basket. While the country gained global market share in its trade of minerals, the share of its global exports of metals, chemicals, and vegetable products was less impressive. Since 2000, Kazakhstan has seen substantive growth in the global market share of its nonmineral products between 2000-2010, with their global market shares almost doubling. However, since the 2010s, these shares have remained low and volatile in their trajectories (Fig. 3.2). Kazakhstan's metal and chemical exports comprised 0.682% and 0.065% of trade in their respective global markets in 2019. Moreover, after the fall in commodity prices in the early 2010s, exports of mineral products also fell and lost global market share, never fully regaining the previous export growth momentum. This trend highlights potential challenges around sustaining long periods of rapid growth with a concentrated export basket.

Figure 3.2



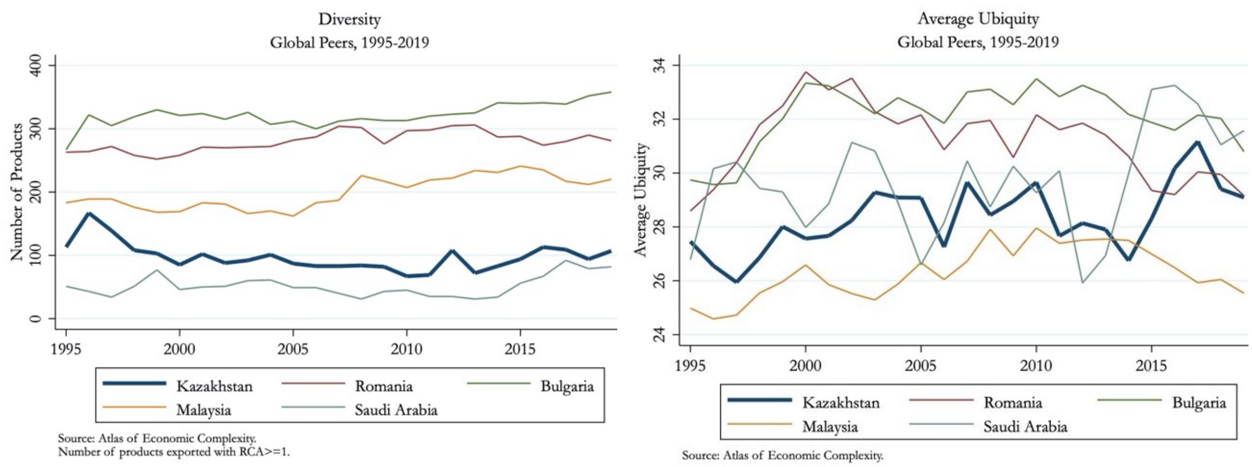
Source: Atlas of Economic Complexity.

Economic Complexity Measures

The Economic Complexity paradigm develops several measures that translate information on the export composition of a country into insights describing the quantity and sophistication of productive knowhow in the economy. To compare Kazakhstan's performance on measures of economic complexity with relevant benchmarks, we undergo a peer selection process that considers countries with similar historical contexts, endowments, and economic challenges. The final list of peers include regional peers (Russia, Uzbekistan, Azerbaijan), global peers (Bulgaria, Malaysia, Romania, Saudi Arabia) and more aspirational OECD peers (Australia, Canada, Chile). See Appendix A for more details on the peer selection process.

A country's export diversity denotes the number of distinct products that a country exports intensively compared to global trade. A country's average ubiquity denotes how commonly its exports are also intensively exported from other countries.[7] Kazakhstan has consistently demonstrated a lower export diversity and higher average ubiquity compared to global peers (Fig. 3.3).

Figure 3.3



The Economic Complexity Index (ECI) combines the measures of diversity and ubiquity to convey the complexity of a country's economy. In 2019, Kazakhstan ranked #79 out of 133 countries on the ECI, having declined over time on the index between 1996 and 2010.[8] Kazakhstan's low ECI positions itself below regional and international peers and among other oil exporters including Kuwait, Oman, and Qatar (Fig. 3.4 and 3.5).

Figure 3.4

[7] A product's ubiquity is moderately correlated with its complexity, though not always (the most complex products are only exported by a handful of countries, but the same is also true about certain rare minerals). In 2019, frozen fish was among the products with the highest ubiquity, being intensively exported by 85 countries. Audio recording equipment was among the products with the lowest ubiquity, being intensively exported by just 8 countries.

[8] A country's ECI rank over time will be influenced by the evolution of their own export basket in addition to the changes in the export basket of other countries. Therefore, a change in a country's ECI reflects not only its own trajectory of exports but also that of other countries. Also important to note is that the negative or positive quality of ECI does not have significance in meaning; ECI is relative and a normalized measure around 0, so negative values simply indicate below average.

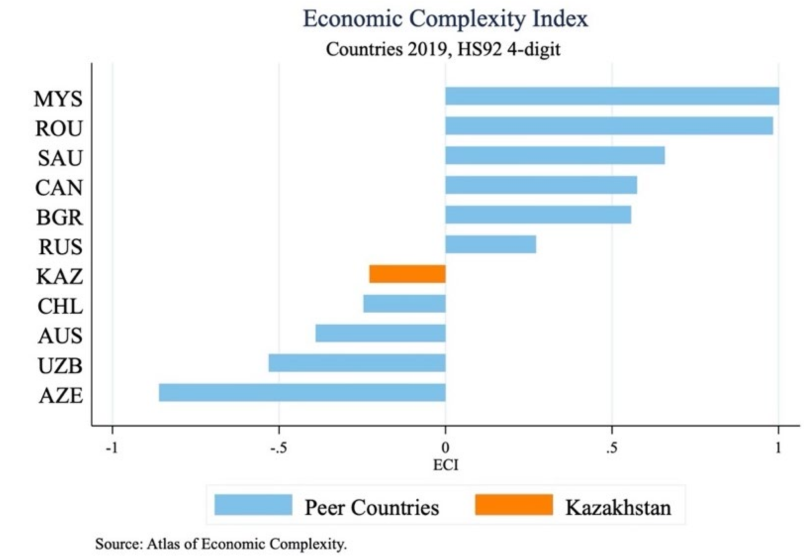
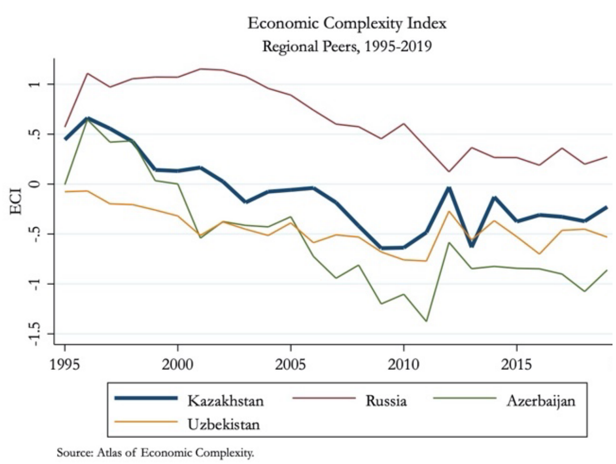
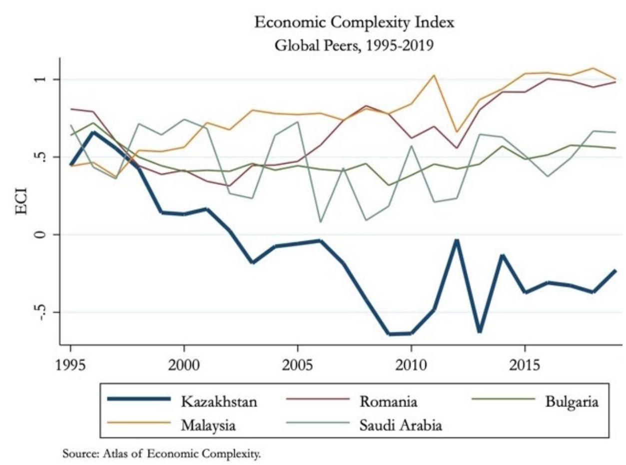


Figure 3.5





Kazakhstan's low ECI and the poor performance over time can be partially explained by the concentration of minerals in its export basket. Just as the measures of diversity and ubiquity can be used to convey information on the complexity of a country's economy, these same measures also convey information on the complexity of products. By examining the tendencies of products to be traded by few or many countries—and the complexity of those country's economies—we arrive at the Product Complexity Index (PCI) that communicates the quantity and sophistication of knowhow needed to produce and export a particular product. Petroleum oils and gases, along with many other extractive exports, are among the least complex products by PCI and occupy more than half of all of Kazakhstan's exports. Meanwhile, Kazakhstan's most complex exports (batteries, parts of motor vehicles, tantalum) tend to constitute very small shares of the country's total export basket (Fig. 3.6).[9] At the sector level, Kazakhstan's exports in agriculture, minerals and textiles have the lowest

[9] Some of the most complex products that Kazakhstan produces include: Amine-function compounds (HS 2921; PCI 1.298), Tantalum (HS 8103; PCI 1.153), Mica and reconstituted mica (HS 6814; PCI 1.121), Rail locomotives (HS 8601; PCI 0.897), slag and rock wool (HS 6806; PCI 0.886).

average complexity, while those in transportation and stone and glass have the highest average complexity (Fig. 3.7).

Figure 3.6

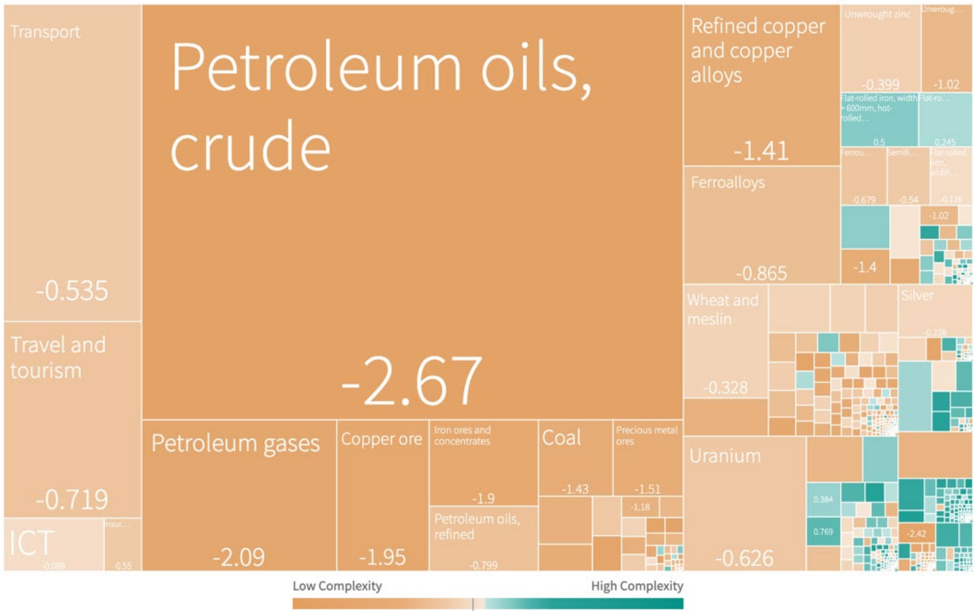
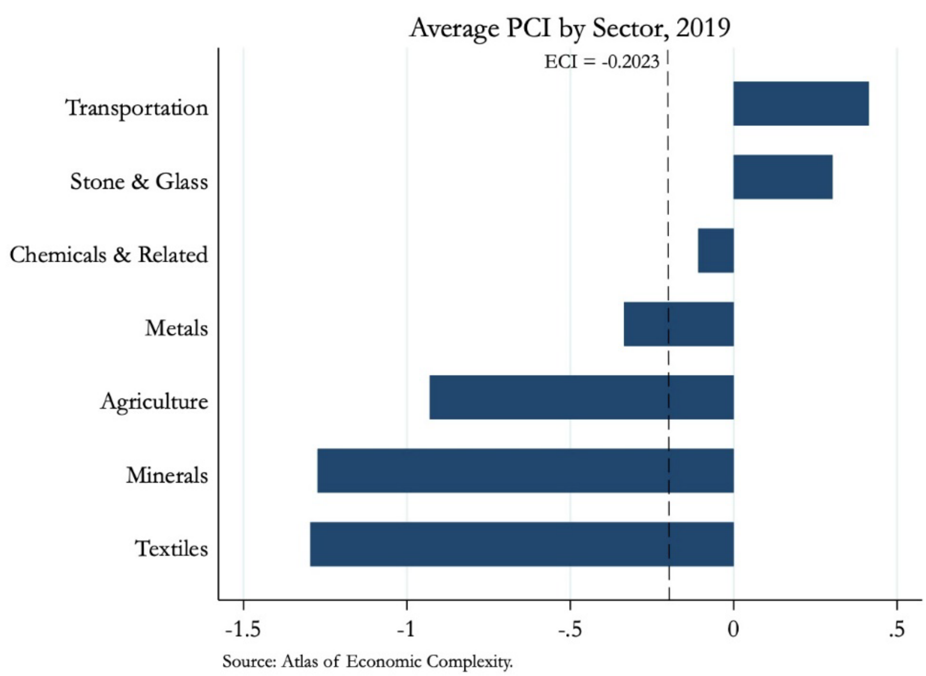
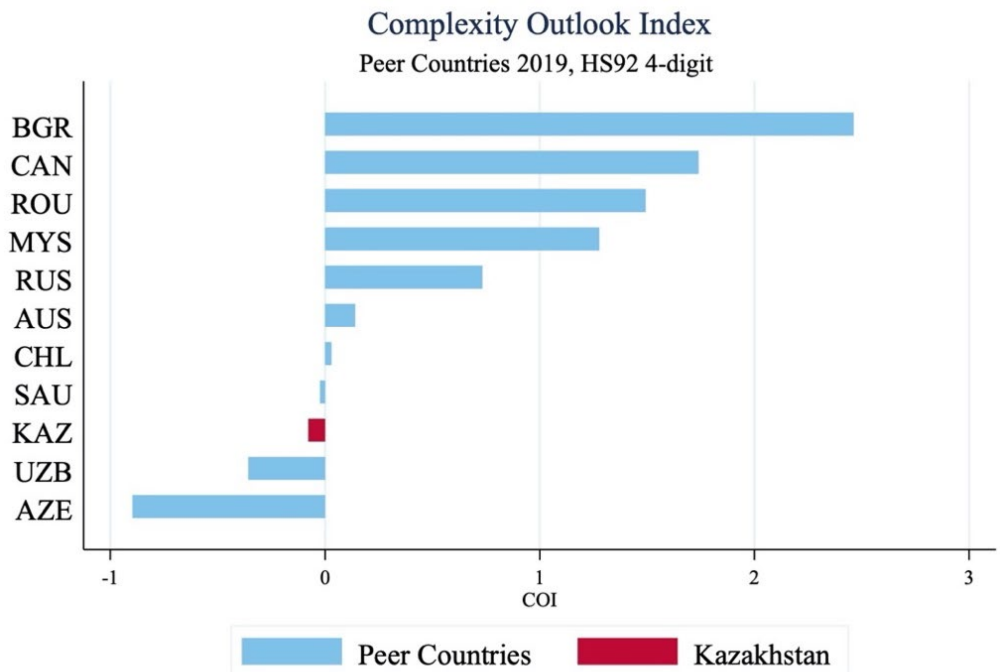


Figure 3.7



The Complexity Outlook Index (COI) employs this measure of PCI to communicate how close a country is positioned to diversify into higher complexity products. In Kazakhstan, the closest diversification opportunities appear to have limited strategic value in contributing to complexification of the economy. Kazakhstan's COI in 2019 lags behind regional, global, and OECD peers (Fig. 3.8).

Figure 3.8



Source: Atlas of Economic Complexity, 2019.

The Challenge of Oil Producers

Kazakhstan's difficulty in diversifying its economy is a challenge that is familiar to other major mineral exporters. The set of capabilities present in these economies tend to be further away from the knowhow required to jumpstart new industries, meaning that more coordination is often required to be competitive in other sectors. In describing the process of diversification and economic complexity, this can be explained as the country not having enough current capabilities—or not being able to coordinate them—to expand beyond its existing production basket to new activities. The problem for Kazakhstan is not that it cannot introduce new products to its export basket, but rather that it has few opportunities nearby that have strategic value in developing complementary engines of growth. Though there have been new products and activities added to its export basket in the last decade, these have been of low product complexity and thus far do not constitute a significant share of the total export basket.

This is a pervasive challenge in countries that concentrate in extractives. This challenge is reflected in the complexity metrics of countries that are particularly intensive in oil and gas. Controlling for GDP per capita, we observe that multiple measures of economic complexity are negatively correlated with the share of a country's GDP in oil and gas rents across the world, including Economic Complexity Index (ECI), Complexity Outlook Index (COI), and export diversity (Fig. 3.9 and 3.10). This result is the same when examining the share of a country's GDP in natural resource rents at large, and when controlling for GDP per capita excluding oil and natural resource rents. For each of these complexity metrics, Kazakhstan performs as we would expect given its endowment. Therefore, Kazakhstan's challenges in diversification are not unique; rather, it is partly explained by the trends systematically observed in resource-intensive countries around the world.

Figure 3.9

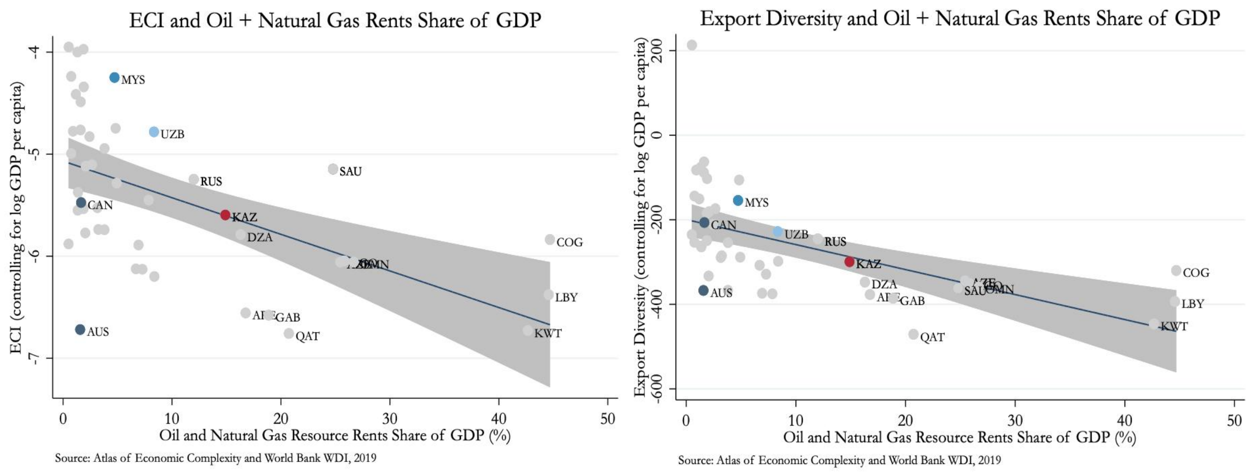
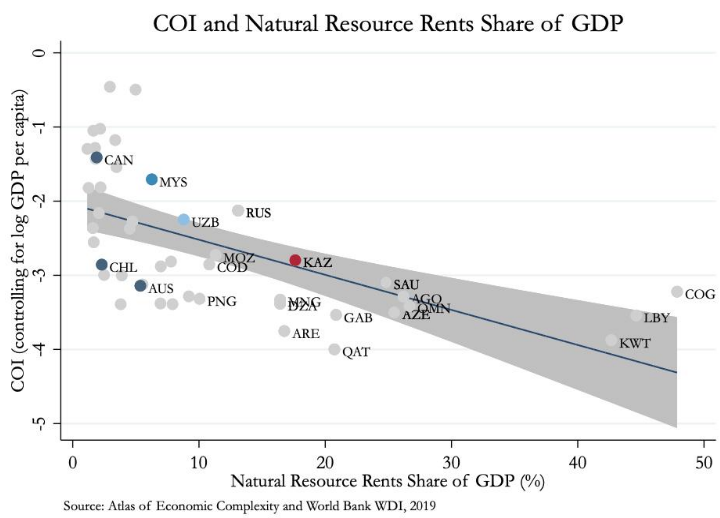


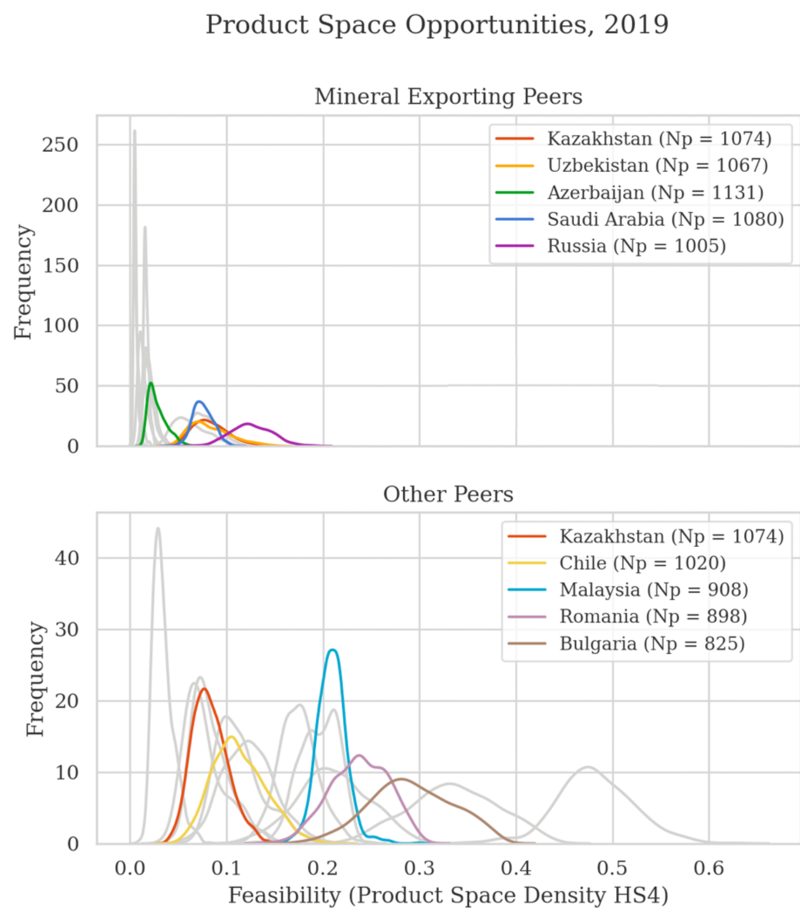
Figure 3.10



In general, countries whose export baskets are concentrated in minerals have fewer feasible opportunities to diversify into. An analysis of Kazakhstan's peer countries reveals that those countries which are more intensive in minerals[10] have, on average, more products that they can diversify into that are *less* feasible, given the country's position in the product space. These countries face additional difficulties acquiring the knowhow required to produce new products (Fig. 3.11). Subsequent parts of this report will employ an alternative version of the measure used here (density) to refine how the distance to opportunities are defined.

Figure 3.11

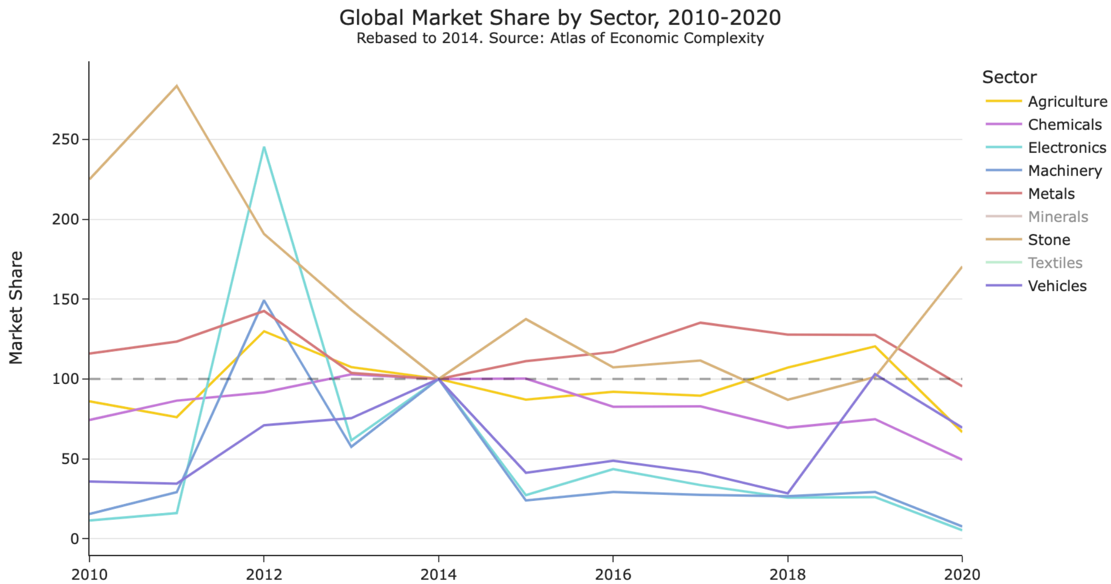
[10] Defined as having more than 5% of GDP composed of oil and gas rents.



Source: Atlas of Economic Complexity and own calculations.
Note: Includes non-mineral product space opportunities only (excludes HS 2500-2799, 2844, 7106 and 7108).
Grey traces in top plot indicate additional countries with oil and gas rents above 10% of GDP in 2019.
Grey traces in bottom plot indicate additional countries with real GDP per capita within 30% of Kazakhstan's value in 2019.

This gap in productive capabilities may help to explain why changes in macroeconomic conditions in 2014 did not make way for stronger growth in non-oil tradables in the years thereafter. Kazakhstan's economy suffered an economic downturn with the global oil price shock in 2014 and the country subsequently switched to a floating exchange rate regime. This led to devaluation of the Tenge, which should have loosened conditions for the growth of tradables; however, after 2015, Kazakhstan's global market share in most sectors failed to improve (Fig. 3.12).

Figure 3.12



Source: Atlas of Economic Complexity

Product Appearances

Despite limited nearby opportunities, Kazakhstan has been able to diversify into a few products which are adjacent to its existing capabilities over the previous decade. Examining the introduction of new products — their type, rate of survival, and frequency with which they appear — can signal crucial information about a country's development trajectory. Since 1997, 50% of products that appeared in Kazakhstan's export basket for the first time had disappeared by 2019. Of the top 50 products that were not competitively exported in 2010 (RCA < 1) with the closest proximity to Kazakhstan's current capabilities, only 11 of them began to be competitively exported in subsequent years, and only 7 of these were still in Kazakhstan's export basket by 2019. In the last 10 years, Kazakhstan has introduced 35 new products that have remained competitive exports until 2019.[11] The largest share of these new products were animal, vegetable, or other agricultural products (42.9%), with smaller shares of products in metals, chemicals, and minerals (Table 3.1). However, in 2019, only 10.4% of the value of Kazakhstan's gross goods exports came from products that were introduced in the past two decades (Fig. 3.13).

Table 3.1

[11] A product appearance is defined here as the first year in which a country finally attains a comparative advantage in the product in global trade (RCA > 1). This measure reduces some noise when compared to using absolute export values (such as the first time a product is exported at all [> $0]) but also implies that the export activity of other countries can influence the product appearances that are measured in a country's export basket.

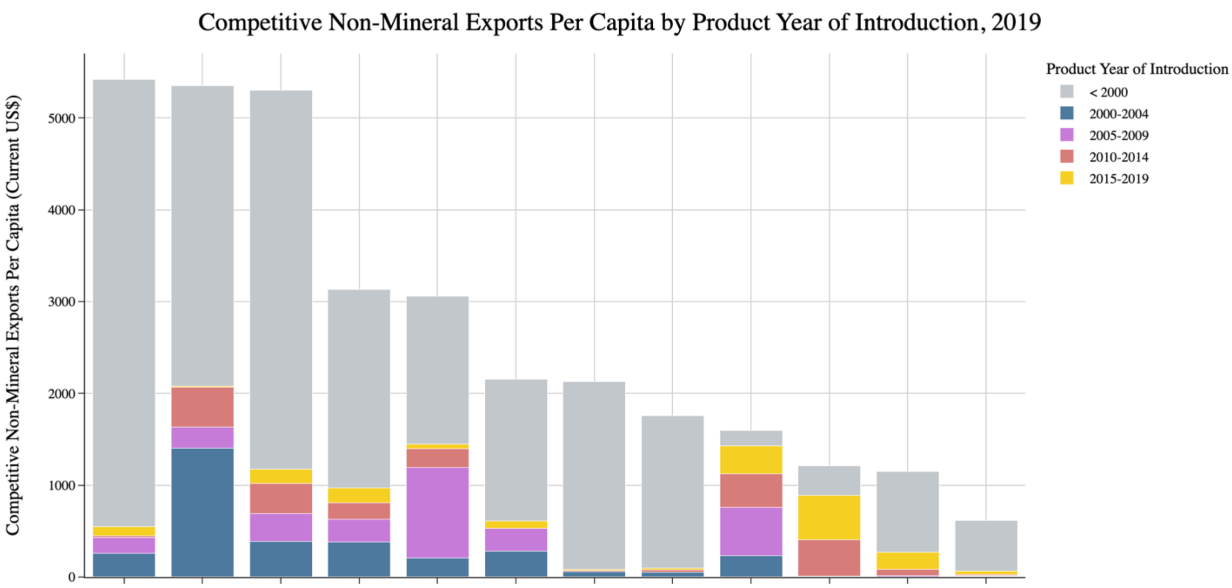
New Products by Product Group: 2010-2019

Product Group	Number of New Products
Animal, Vegetable & Other Food Products	15
Metals	5
Chemicals & Related Industries	4
Mineral Products	4
Transportation Products	3
Stone & Glass	2
Textiles & Garments	2

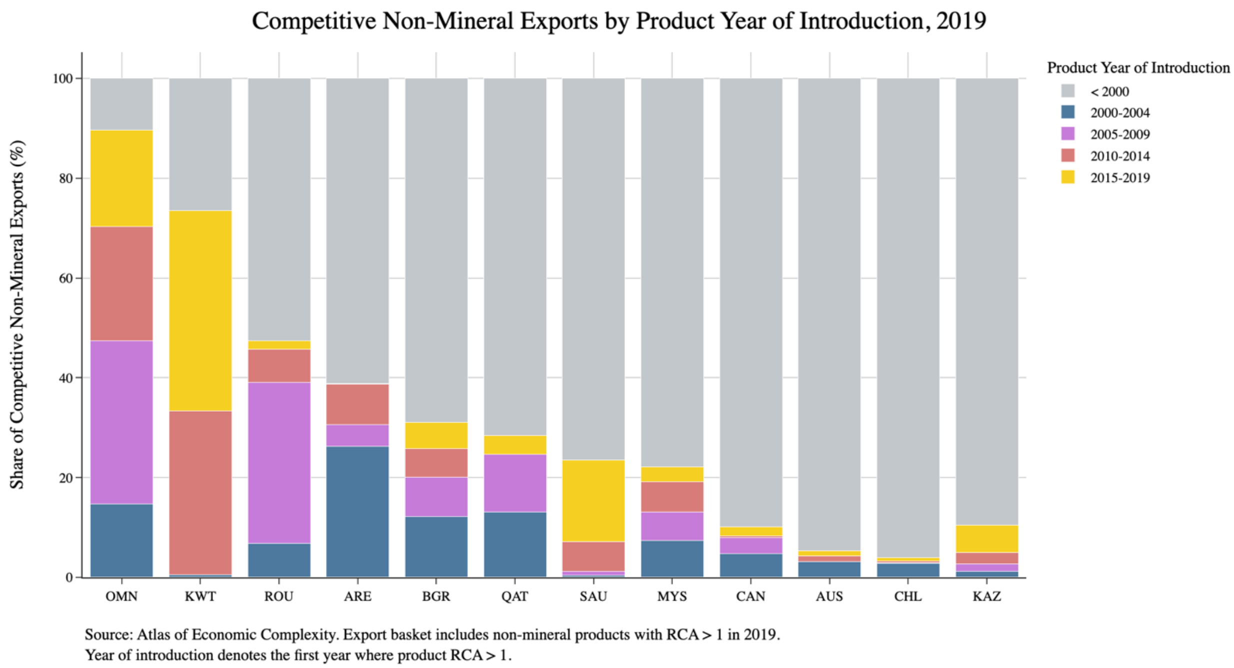
*Note: only includes products that have remained in the export basket since their year of introduction with RCA >=1.

Source: Atlas of Economic Complexity.

Figure 3.13



Source: Atlas of Economic Complexity. Export basket includes non-mineral products with RCA > 1 in 2019.
Year of introduction denotes the first year where product RCA > 1.



Overall, Kazakhstan's efforts to introduce new products have been relatively unsuccessful at materializing into diversification of the nonmineral economy and promoting new engines of growth. Government-driven efforts to promote certain industries have had varying degrees of success in attracting investments, bolstering sectoral productivity and expanding exports. The next section provides an overview of several broad-based diversification plans spearheaded by the government in the last decade and identifies limitations that may have prevented them from being more successful.

An Overview of Previous Diversification Efforts

The most recent period of Kazakhstan's active industrial policy began in 2010, with the initiation of the first State Program for Industrial and Innovative Development (SPIID). The SPIIDs are five-year programs spanning 2010-2014, 2015-2019, and 2020-2025. They are designed and implemented principally by the Ministry of Industry and Infrastructure Development, with roles for institutions and line ministries such as the Ministry of Economic Development of Trade, Ministry of Education, Ministry of Labour and Social Protection, Baiterek, and DAMU described alongside complementary plans (e.g. Business Roadmap 2020, Productivity 2020, Employment 2020, Industrialization Maps).

The SPIID 2010-2014[12] was a broad-based program focused on the development of a wide range of sectors including agriculture, mining and metallurgical industries, tourism, pharmaceuticals, petrochemicals, ICT, biotechnology, alternative energies, and space technologies. Within each broad sector, the program specified a combination of target indicators to increase sectoral GVA, export volumes, labor productivity, and more. Overall program targets aimed to increase real GDP by 15%, non-raw material share of exports to 40%, and productivity of the manufacturing sector by 50% between 2008-2015. Major investments in air, rail, road, and port infrastructure were anticipated to be the drivers behind much of this growth; within the program's budget of 4.2 trillion Tenge, approximately 46.4% was allocated to the development of such projects. Over the span of the

[12] Office of the President of the Republic of Kazakhstan, "State Program for Industrial and Innovative Development 2010-2014," March 19, 2010, https://afmrk.gov.kz/en/activity/strategy-and-program/the-state-program-of-industrial-innovative-develo.html.

program, the economy saw an increase of real GDP and productivity within the manufacturing sector, but efforts fell short to stimulate non-raw-material exports to represent a larger role in the economy. While the program introduced a range of new non-oil tradables that were not previously produced in Kazakhstan,[13] these products were not exported enough to spur stronger growth.

Exogenous factors in 2014, primarily fueled by the fall in commodity prices, exposed that the economy was still vulnerable to commodity price shocks after a decade of strong growth and the efforts to counteract such consequences outlined in the SPIID. Before the global financial crisis of 2007-2008, nominal GDP and Brent oil prices were already closely correlated, but medium-term growth dynamics appeared especially closely correlated with oil prices after 2010. After the fall in oil prices in 2014, the resulting macroeconomic adjustments were necessary but ultimately insufficient in sparking the diversification into non-oil tradables.

Subsequently, the SPIID 2015-2019[14] aimed at improving upon the 2010-2014 program by narrowing its focus to only 14 manufacturing sectors, including ferrous metallurgy, non-ferrous metallurgy, oil refining, food production, agrochemistry, agricultural machinery, automotives, and building materials. Within each of these 14 priority sectors, the program specified a combination of target indicators to increase sectoral GVA, employment, labor productivity, and export volumes by 2019. Overall targets spanning the entire manufacturing industry aimed at increasing gross value added (GVA) and labor productivity by 40%, nonprimary exports by 10%, reducing energy intensity of production by 15%, and increasing sectoral employment by 5%. The program also identified specific target products. However, the selection of these products appears to have been largely driven by an import substitution agenda and an evaluation of the demand in China, Russia, and the CIS region.

The program was largely effective in achieving broad targets. An evaluation of the program by the Asia Development Bank concluded that the 2015-2019 program achieved 99% of the real GVA targets and 113% of the aggregate export targets, though targets to increase sectoral employment fell short by 20% and investments in fixed assets fell short by 2%.[15] This program was particularly successful in stimulating investments in the manufacturing sector, though the majority of investments were in the same products in which the country already had a comparative advantage.[16]

Currently, the SPIID 2020-2025[17] is underway and considers the experience of the two previous programs in order to develop new approaches for further industrialization. The plan introduces a new emphasis on supporting digitalization and technology industries and describes the importance of "moving from disparate instruments of state support to a system of comprehensive development stimulation in exchange for counter obligations and business responsibility."[18] Key performance

[13] Office of the President of the Republic of Kazakhstan, "Technological Upgrade, Investment Inflow, Innovation Boom, or How Kazakhstan Made an Industrial Breakthrough," December 12, 2019.; Gulaikhan Kubayeva and Alibek Konkakov, "Progress in Diversification of the Economy in Kazakhstan" (The Rise of Eurasia: New Perspectives on East-West Business and Economic Relations, The Netherlands, 2016).

[14] Office of the President of the Republic of Kazakhstan, "State Program for Industrial and Innovative Development 2015-2019," August 1, 2014, https://policy.asiapacificenergy.org/node/309.

[15] Asia Development Bank, "Assessing the Results of Industrial Innovation Development in Kazakhstan," June 11, 2020, https://development.asia/summary/assessing-results-industrial-innovation-development-kazakhstan.

[16] Office of the President of the Republic of Kazakhstan, "Technological Upgrade, Investment Inflow, Innovation Boom, or How Kazakhstan Made an Industrial Breakthrough."

[17] Office of the President of the Republic of Kazakhstan, "State Program for Industrial and Innovative Development 2020-2025," February 15, 2018, https://baiterek.gov.kz/en/programs/gosudarstvennaya-programma-industrialno-innovacionnogo-razvitiya-respubliki-kazakhstan-na-2020-2025.

[18] Office of the President of the Republic of Kazakhstan, "Technological Upgrade, Investment Inflow, Innovation Boom, or How Kazakhstan Made an Industrial Breakthrough."

indicators for the overall program are organized around similar economic objectives as previous plans. The program sets out to increase labor productivity by 60%, manufacturing exports by 90%, and investment in fixed assets by 60% by 2025. The program introduces improving the country's ECI as an explicit objective, with a target for Kazakhstan to become the 55th most complex economy by 2025.

In parallel to the SPIIDs set at the national level, the government implemented concurrent subnational development plans such as the Strategy for Territorial Development 2015 and Scheme for Spatial-Territorial Development 2020. The latter came to replace the Strategy and it was further extended through 2030. The Scheme outlines similar growth objectives that aim to be more tailored to the socio-economic potential of each region.[19] These plans were designed and implemented principally by the Ministry of Economy, with collaboration between regional and central governments for specific needs.

The plans focus on the development of regions to reduce cross-regional inequalities and promote export activity. They are oriented towards supporting urbanization in designated growth centers (most often the capitals of each region) and identifying areas of specialization with potential in different regional centers.[20] The Scheme for Spatial-Territorial Development set targets to increase the share of processing activities in all industries by 2020, ranging from 30% in Mangystau, Atyrau, and West Kazakhstan to 95% in Almaty City. The program also designates specific state support to help single-industry towns and populations in areas of low economic potential to relocate to these growth centers.

Subnational plans with more details on implementation are the Programme on Development of Regions (PDR) and Industrialization Maps. The PDR—under the Ministry of Development and Trade—provides detailed metrics, financial arrangements, and programs for infrastructure development within the regions. They borrow instruments for regional development that have been effective elsewhere, including structural funds like those in the EU, and programs for infrastructure development similar to those implemented in Canada.[21] Hundreds of specific capital-intensive projects were implemented under the companion Industrialization Maps, beginning with the first SPIID, which lay out project details, locations, and partners, with varying degrees of success.[22] Much of these earlier projects revolved around construction or modernization of processing plants in agrochemicals, metals, and mineral refinement with SOEs as the main project applicant,[23] and more recent Industrialization Maps have included more diverse industries and more locations throughout the country.

Our Understanding of Past Efforts' Limitations

This section synthesizes the limitations of the previous national and subnational diversification strategies that prevented the programs from being more effective.

1. National and subnational plans lacked policy customization to fit regional contexts. Previous plans were limited in their ability to assess future growth opportunities for regions, linking growth

[19] Office of the President of the Republic of Kazakhstan, "Experts on the Address: Forecast Scheme Will Allow Regions to Use Competitive Advantages and Identify Reserves of Economic Growth," October 20, 2018, https://primeminister.kz/en/news/17323.
[20] Ibid.
[21] OECD, "OECD Territorial Reviews: Kazakhstan," June 15, 2017, https://doi.org/10.1787/9789264269439-en.
[22] Office of the President of the Republic of Kazakhstan, "Technological Upgrade, Investment Inflow, Innovation Boom, or How Kazakhstan Made an Industrial Breakthrough."; Dinara Bekbolaeva, "Billions of Wasted Budget Funds Spent on Industrialization Were Revealed by the Accounts Committee," *Baige News*, August 14, 2019, https://baigenews.kz/milliardy-vpustuyu-potrachennyh-na-industrializatsiyu-byudzhetnyh-sredstv-vyyavil-schetnyy-komitet_136825/.
[23] Office of the Prime Minister of the Republic of Kazakhstan, "On Industrialization Map," December 31, 2014, https://adilet.zan.kz/eng/docs/P1400001418.

prospects to current economic performance rather than on demonstrated capabilities.[24] Our methodology aims to address this by building the consideration of productive capabilities directly into the framework to identify opportunities (see Section 6).

2. Past SPIIDs and Industrialization Maps often had limited ability to assess innovative diversification opportunities beyond what the regions are already intensive in. While the later SPIIDs (2015-2019 and 2020-2025) were more specific in their sectoral focus, they are still somewhat limited in their capacity to identify new and innovative opportunities that build on existing capabilities in the regions, beyond strategies that emphasize beneficiation of existing raw materials.[25]

3. Opportunity identification criteria may have focused too much on the size of domestic and regional markets, rather than factors that point to the feasibility of implementing such recommendations. While domestic and regional demand is important to assess whether the product has a consumer base that is sufficiently large and nearby, it does not suggest the viability of the product being produced and exported competitively (such as whether the location has the necessary skills, inputs to production, or infrastructure).

4. Both the national and subnational development plans focused on simultaneous development of many sectors of the economy at the same time. While this broad scope reflects the ambitions of the government to transform the economy and promote a more diverse export base, the process of structural transformation requires prioritization and targeted efforts towards sectors that leverage existing capabilities.

5. The effectiveness of previous plans may have been inhibited by the sheer number of concurrent initiatives pursued at different levels of government and a lack of coordination between them. The OECD describes how the overlap of previous plans and initiatives across agencies has led to inefficiency in the use of government funds, and that the multitude of programs and overwriting of them may have weakened institutional commitment to them.[26]

6. Finally, there were additional exogenous factors outside control of the programs that may have hindered past efforts to promote certain industries. As with many countries, Kazakhstan is subject to commodity price shocks, geopolitical tension, and global financial risks that contribute to economic and political uncertainty. Despite the efforts of the national plans and the devaluation of the currency that should have improved conditions for exports, most export-oriented activities remained constrained after the global commodity shock in 2014.

[24] Office of the Prime Minister of the Republic of Kazakhstan, "Forecast Scheme for Spatial-Territorial Development of Kazakhstan Till 2020," July 21, 2011.
[25] OECD, "Regional Policies to Support Diversification and Productivity Growth in Kazakhstan," OECD Eurasia Competitiveness Programme, April 2020, https://www.oecd.org/eurasia/competitiveness-programme/central-asia/Regional-Policies-to-Support-Diversification-and-Productivity-Growth-in-Kazakhstan-ENG.pdf.
[26] OECD, "OECD Territorial Reviews: Kazakhstan."

4. Motivation and Data Considerations for a Sub-National Complexity Analysis

Overview of Main and Complementary Economic Complexity Analyses for Kazakhstan

This report contains two types of complexity analysis for diversification opportunities in Kazakhstan: a main approach based on export data and a complementary approach based on employment data. The main analysis leverages export data to determine which goods constitute appropriate diversification opportunities for Kazakhstan. Export data for goods is typically very high quality; data in the Harmonized System (HS) classification, which we use herein, covers more than 1,200 goods traded by all world countries from 1995 – present. This data is also available subnationally for Kazakhstan at the level of each region. However, the main analysis does not cover service exports because the data quality of international traded services is typically very poor. Not all countries are covered, and service data is highly aggregated into just 5 categories covering all services (finance, ICT, transport, travel and tourism, and miscellaneous).

Nevertheless, service diversification opportunities are arguably important for Kazakhstan. According to the World Bank, services account for more than 50% of GDP and more than 60% of employment in Kazakhstan. We thus address this shortcoming of the main analysis by using a complementary analysis that leverages employment data rather than export data. Specifically, this complementary analysis uses firm registry data from Kazakhstan's Bureau of National Statistics in combination with the United States' County Business Patterns dataset in order to determine how many people are employed in each industry in a given location. While this data covers all industries, its particular coverage of services is especially useful. We thus apply economic complexity methodology[27] to this data to determine which industries could constitute valuable diversification opportunities. However, this is a complementary rather than main analysis because the quality of employment data is not as good as that of export data. In order to mitigate data quality issues, employment-based analyses focus on Astana and Almaty City.[28]

Rationale for a Subnational Approach

The rationale for pursuing a subnational approach to complexity lies in the heterogeneity of export baskets across the regions of Kazakhstan, both in types of goods and the share of total exports they contribute to the national export basket. Decomposing the national export basket reveals that a handful of regions are responsible for most exports in the country. The value of total exports across regions in 2019 ranged from $236.8 million (North Kazakhstan) to $21.272 billion (Atyrau). Furthermore, the export baskets of many regions are heavily concentrated in one sector and tend to remain close to their natural resource endowments. As such, many regions display export basket compositions like that of nearby regions.

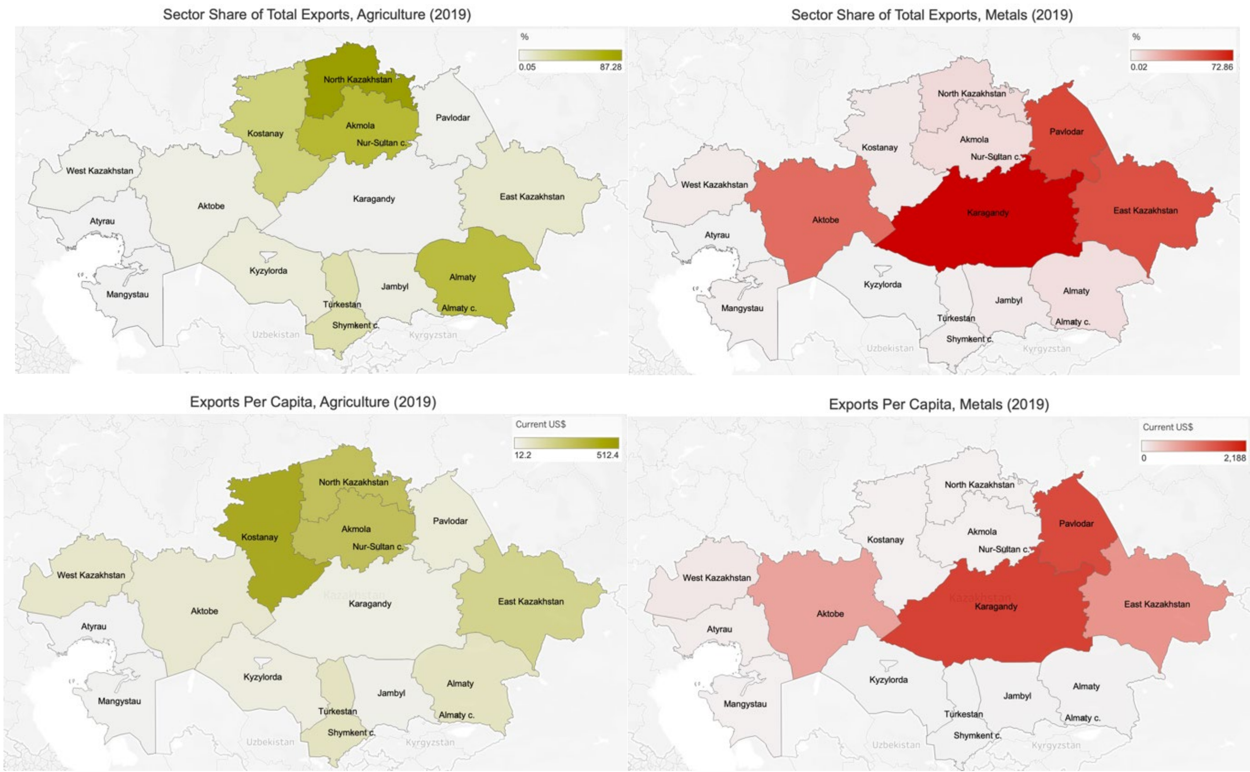
Minerals (and particularly crude petroleum oils) dominate over 96% of the total exports in the western regions of Atyrau, Mangystau, and West Kazakhstan; metals comprise at least 50% of total exports in the eastern regions of Karagandy, Pavlodar and East Kazakhstan; various agricultural goods comprise the majority of exports in the northern regions of North Kazakhstan and Akmola, and the uranium is the predominant export of the southern regions of Jambyl and Turkestan, where many of

[27] We apply economic complexity methodology to industries of employment insofar as we consider how the current production of a location informs what it could produce next. However, some economic complexity metrics like PCI and ECI only function well with international trade data, not with subnational employment data; we thus do not compute these metrics in the complementary analysis of industries.

[28] Note that Almaty City is distinct from the region of Almaty, which are both their own administrative divisions within Kazakhstan.

Kazatomprom's uranium deposits are located. For example, Fig. 4.1 illustrates the export concentration of agriculture and metals across regions.

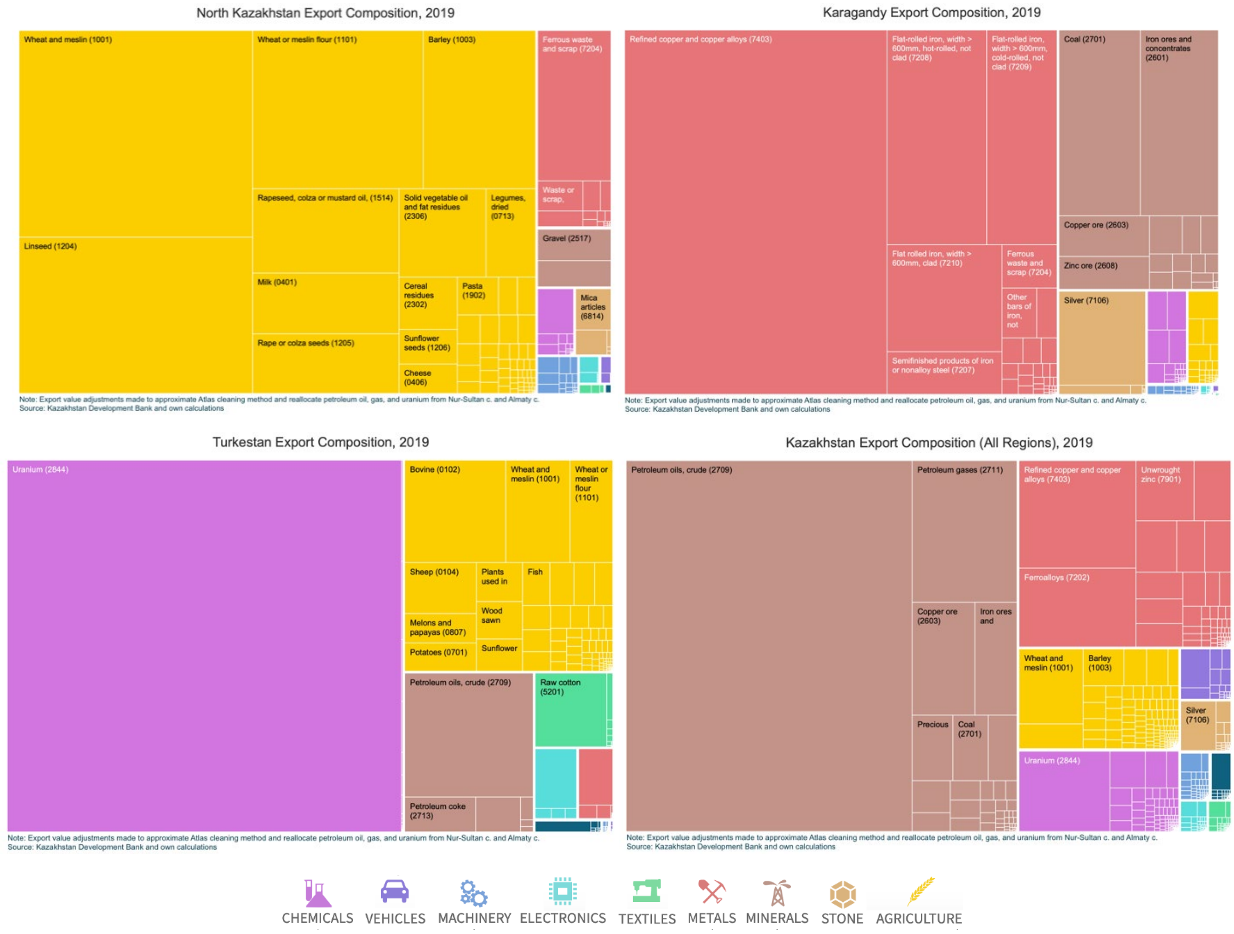
Figure 4.1



Source: Development Bank of Kazakhstan.

Many capabilities of these regions would not be apparent in a national level analysis of the export basket. In evaluating the export competitiveness of Kazakhstan as a country, the sheer value of oil and gas tends to overshadow many non-mineral products that otherwise comprise very significant shares of their region's export basket. In consideration of Kazakhstan's potential constraints to agglomeration, this means a diversification strategy that only views pathways at the national level might miss some opportunities altogether. Conversely, opportunities might not be feasible if the capabilities assessed at the national level exist in different corners of the country and face difficulties coming together. Fig. 4.2 demonstrates the heterogeneity in export basket compositions across regions and compared to the country-level export basket in 2019.

Figure 4.2



Thus, we pursue a complexity analysis at the subnational level that organizes regions into several "macro regions" based on the similarity of their export basket compositions. The grouping of individual regions together into larger units helps to reduce noisy and less significant recommendations that may result from separate analyses of individual regions, while also addressing the concerns of an analysis at the national level. Since the most similar regions in terms of export composition tend to be neighbors geographically, agglomerating the required knowhow to pursue diversification opportunities may be less of a barrier within the context of macro regional strategies. Lastly, thinking of "macro regions" might also strike a better balance in terms of feasibility and effectiveness of implementation between pursuing dozens of independent strategies each region and a wholly aggregated strategy. The end of this section describes the approach to construct the macro regions.

Data and Methodology Adjustments for the Main Complexity Analysis with a Subnational Approach

Adjustments to Data

To pursue the complexity analysis with a subnational approach, we leverage a dataset from the Kazakhstan Development Bank (KDB) on exports and imports of goods at the regional level, recorded in thousands of USD on a monthly basis from January 2012 through April 2021 using

Harmonized System (HS) product classifications. The exclusive focus on goods means that the core complexity analysis is conducted without consideration of services and the capabilities and opportunities they imply.

Cleaning was needed to make the dataset comparable to the country-level export data in the Atlas with which we leverage to calculate comparative advantage and complexity metrics for the regions. The Atlas implements a cleaning procedure of raw UN COMTRADE data to account for inconsistent reporting practices across countries. This procedure re-estimates exports and imports using the following steps:

1. Import values are corrected (which are reported including the costs of freight and insurance – CIF) to compare with the same flows reported by exporters (which are reported free on board – FOB).
2. An index of reliability in each country's reporting of trade flows is constructed, based on the consistency of trade totals reported by all exporter and importer combinations over time.
3. Trade values are re-estimated using the data reported by exporters and importers, by considering how reliable each country is.

To implement this procedure for the regions of Kazakhstan, we compare Kazakhstan's national-level exports in the cleaned Atlas data to the sum of regional exports in the KDB data. We take the ratio of these exports at the year-product-trade partner level and adjust the KDB export values proportionally by this ratio across all regions where there are recorded exports of that product to that trade partner. This procedure results in the net reduction of total exports across regions from $58.065 billion to $43.838 billion in 2019, with over $11.3 billion of the original discrepancy between the data occurring in crude petroleum oils alone. This follows that the regions with the most significant change in total exports after the adjustments are the main oil-exporting regions of Atyrau, Mangystau and West Kazakhstan.

A second adjustment was made to partially correct the "headquarter problem", whereby a share of exports may be recorded where the headquarter of an establishment is located rather than where the activity takes place—and where arguably a large share of the productive capabilities are concentrated. Since we use exports to signal the presence of knowhow in a particular place, it is crucial for exports to be recorded where productive activities happen. In the KDB data, Astana records higher than expected exports of crude petroleum oils, petroleum gases and uranium potentially due to its role as the headquarter city of KazMunayGas, Kazatomprom, and other companies that are the country's major exporters of these commodities. In 2019, the export of these three commodities exceeded US$ 5.5 billion dollars in Astana as over 88% of the city's total export basket. This problem exists to a lesser extent in Almaty city and is largest for petroleum gases.

To partially address this problem, we reallocate crude petroleum oil and uranium exports from Astana to each other region proportionally by their reported shares of exports in the commodity in 2019. For petroleum gases, we observe significant exports recorded in Almaty city and less than expected exports from West Kazakhstan, where the major onshore Karachaganak gas field is situated.[29] To address this, we first reallocate 16% of Astana and Almaty City's exports to West Kazakhstan based on sources of its production shares in the country. Then, we reallocate the remaining petroleum gas exports from the cities to the remaining regions proportionally by their reported shares of exports in the commodity. Table 4.1 details the reallocation of petroleum oils, gases, and uranium, including each region's shares

[29] The Karachaganak field is responsible for approximately 16% of the country's daily output of petroleum gases, according to https://www.offshore-technology.com/marketdata/karachaganak-conventional-gas-field-kazakhstan/.

of exports in the commodity pre-allocation (the reported shares with which the reallocation is based) and post-allocation. We do acknowledge that other "headquarter problems" may persist, however, further corrections require more granular information regarding firm ownership structures, production, sales, or similar data. This remains as an opportunity for improvement for future research efforts.

Table 4.1

Petroleum oils, crude (2709)

Region	*Pre-reallocation* Export Value (million US$)	*Pre-reallocation* Export Share of Product (%)	*Post-reallocation* Export Value (million US$)	*Post-reallocation* Export Share of Product (%)
Aktobe	378.84	1.83	406.69	1.96
Almaty	0.03	0.00	0.03	0.00
Almaty c.	56.02	0.27	60.14	0.29
Atyrau	14110.16	67.99	15147.29	72.99
East Kazakhstan	0.00	0.00	0.00	0.00
Kyzylorda	341.29	1.65	366.37	1.77
Mangystau	1834.61	8.84	1969.46	9.49
Nur-Sultan c.	1420.89	6.85	0.00	0.00
Turkestan	71.55	0.35	76.81	0.37
West Kazakhstan	2538.70	12.23	2725.30	13.13

Petroleum gases (2711)

Region	*Pre-reallocation* Export Value (million US$)	*Pre-reallocation* Export Share of Product (%)	*Post-reallocation* Export Value (million US$)	*Post-reallocation* Export Share of Product (%)
Aktobe	78.01	2.67	259.24	8.88
Almaty	1.58	0.05	5.24	0.18
Almaty c.	502.71	17.22	0.00	0.00
Atyrau	583.82	19.99	1940.12	66.44
East Kazakhstan	43.15	1.48	143.41	4.91
Jambyl	3.38	0.12	11.24	0.39
Kyzylorda	18.28	0.63	60.76	2.08
Nur-Sultan c.	1619.88	55.47	0.00	0.00
Pavlodar	12.78	0.44	42.46	1.45
Shymkent c.	24.19	0.83	80.40	2.75
Turkestan	2.28	0.08	7.59	0.26
West Kazakhstan	30.11	1.03	369.72	12.66

Uranium (2844)

Region	Pre-reallocation		Post-reallocation	
	Export Value (million US$)	Export Share of Product (%)	Export Value (million US$)	Export Share of Product (%)
Akmola	12.42	0.84	39.86	2.70
Almaty c.	50.57	3.43	162.29	11.01
Atyrau	1.48	0.10	4.74	0.32
East Kazakhstan	8.38	0.57	26.90	1.83
Jambyl	142.70	9.68	457.99	31.07
Kyzylorda	26.73	1.81	85.78	5.82
Mangystau	0.00	0.00	0.00	0.00
Nur-Sultan c.	1014.77	68.84	0.00	0.00
Turkestan	217.01	14.72	696.49	47.25

Adjustments to Calculations of Complexity Metrics

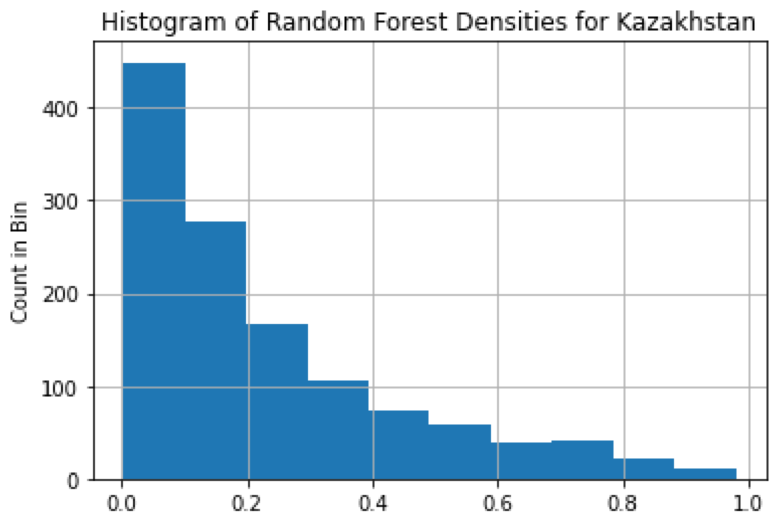
To conduct an analysis at the sub-national level, benchmarks are needed to measure the relative intensity of exports in Kazakhstan's macro regions. A standard complexity analysis at the country level measures a country's intensity in a product as relative to the product's presence in global trade, and this is also how macro regions will be benchmarked in this analysis. The latest year available for global trade data in the Atlas is 2019, so this is the year on which the complexity analysis is based.

Each macro region's Revealed Comparative Advantage (RCA) in each product is calculated as a ratio between the product's share of the macro region's exports (by summing the exports of the regions within the macro region) and the product's share of global trade in 2019, using all 243 countries and territories in the Atlas data. The M_{cp} matrix on which most complexity metrics are based is constructed by binarizing these RCAs, whereby $M_{cp}=1$ indicates country c has an RCA in product p that is >= 1, and 0 otherwise. Proximities between product pairs and PCIs are calculated using a subset of this M_{cp} matrix including 133 countries that meet minimum thresholds for population (> 1 million), annual trade flows (> $1 billion), and reporting reliability. ECI is computed with these measures following the same form detailed in Appendix C.

This analysis uses an alternative density measure that employs a Random Forest (RF) machine learning approach rather than the traditional approach that leverages co-export probabilities to infer the similarities—and thus, proximities—between products. This alternative measure of density is particularly useful in sparse product spaces—for instance, in oil economies like Kazakhstan, where the knowhow employed in the commodity sectors they are intensive in is far from knowhow required by other industries. For a given product, the traditional density measure considers its proximity to every other product in the product space with which the country is exporting competitively (M=1). In oil economies there may be very few products with M=1, meaning the density of their product space is extremely low. Fig. 3.11 plots the traditional product densities at the extensive margin of Kazakhstan and its peers, distinguishing between its mineral exporting peers (those with more than 8% of GDP in oil and gas rents) and other peers.

The RF approach to density works by answering the following prediction problem: given knowledge of all your country's M_{cp}s except for a target product p, does your country have an M_{cp} in product p? The algorithm divides countries into a set of 'training' countries and 'validation' countries. It examines data among the training countries and uses the RF algorithm to learn how to answer this prediction problem one product at a time. It then tests the accuracy of the prediction rules it learns by seeing how well they perform on validation countries. In practice, this approach provides a modest boost in predictive performance for countries generally and a large boost in predictive power for countries intensive in natural resources such as Kazakhstan. The RF algorithm also greatly expands the variation of density across products for countries intensive in natural resources (Fig. 4.3). Whereas the standard density measure indicates that countries like Kazakhstan are consistently far from all potential product opportunities, the RF algorithm helps to differentiate between products at different levels of feasibility. For more details see Appendix C.

Figure 4.3



The calculations for COG and COI use the densities produced from the RF approach. The product selection process in Section 4 will leverage the complexity metrics of distance, COG, PCI and more to identify regional export opportunities.

Clustering of Regions into Macro Regions

Diversification opportunities are to be identified at the level of several regional groupings—or "macro regions". To construct the groups, a dimension reduction algorithm called UMAP is applied to the export baskets (in terms of M_{cp}) of the regions of Kazakhstan and world countries. In this context, the dimension reduction algorithm compresses complex non-linear relationships among the export profiles of the regions and world countries into a two-dimensional space that places countries immediately next to each other if they have highly similar export capabilities (Fig. 4.4). This allows us to ascertain which regions of Kazakhstan are most like each other in terms of their export M_{cp}, and thus to group them together into cohesive units for later analysis (Fig. 4.5). For additional details see Appendix C.

The clustering algorithm yields clear groupings of the regions of Kazakhstan, mapped in Fig. 4.5. These groupings become the macro regions with which diversification opportunities will be identified, with the exceptions of separating Almaty City and Astana into their own groups as major urban centers with distinct constraints and opportunities. We identify the following macro regions:

- Industrial Belt: Aktobe, East Kazakhstan, Karagandy, Pavlodar, West Kazakhstan
- Caspian Regions: Atyrau, Mangystau
- Northern Regions: Akmola, Kostanay, North Kazakhstan
- Southern Regions: Almaty, Jambyl, Kyzylorda, Shymkent, Turkestan

- Almaty City
- Astana

Exports across each region within the macro region are summed to form the macro region's export basket, which will serve as the main input into the complexity analysis for the macro region. We acknowledge that this grouping is wholly determined by the algorithmic approach, and might be undermining relevant considerations in terms geography, connectedness, history, policy planning and other relevant variables that may indicate more suitable groupings for the identification of opportunities.

Figure 4.4

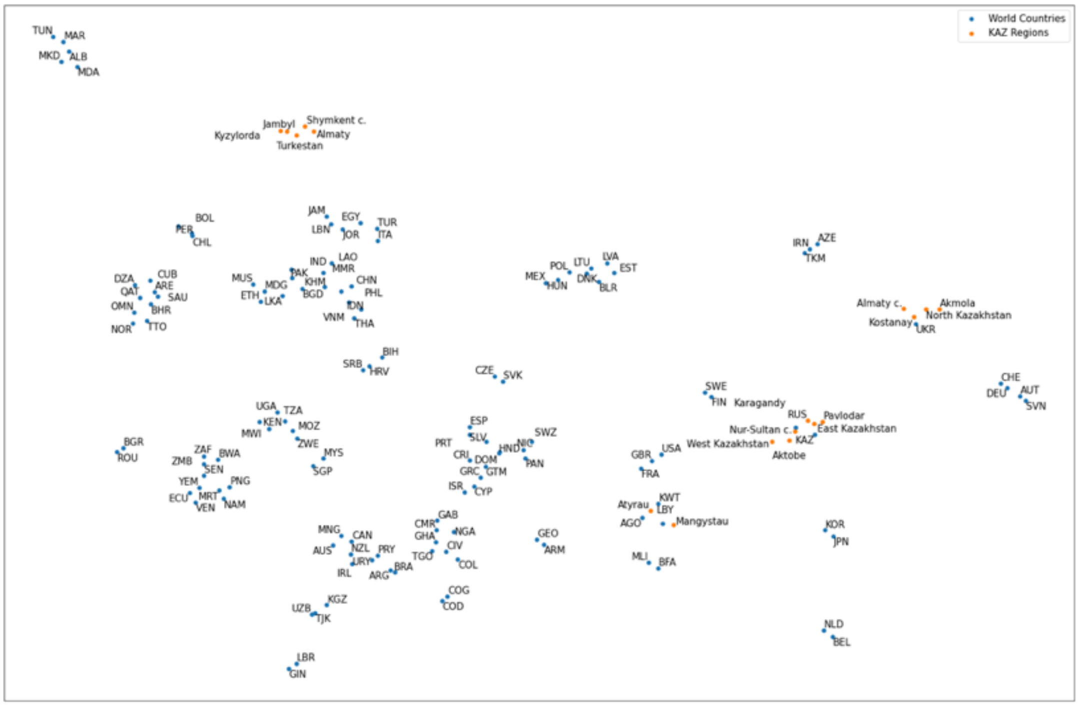
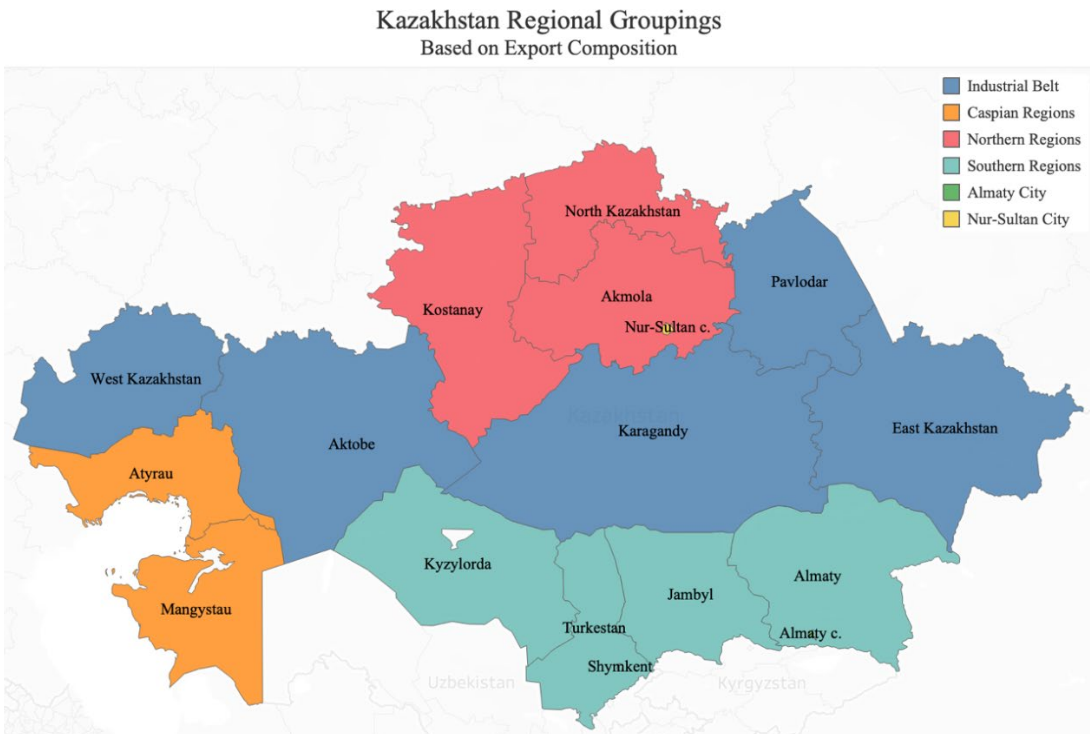


Figure 4.5


Kazakhstan Regional Groupings
Based on Export Composition
Industrial Belt
Caspian Regions
Northern Regions
Southern Regions
Almaty City
Nur-Sultan City
North Kazakhstan
Kostanay
Akmola
Nur-Sultan c.
Pavlodar
West Kazakhstan
Aktobe
Karagandy
East Kazakhstan
Atyrau
Mangystau
Kyzylorda
Turkestan
Shymkent
Jambyl
Almaty
Almaty c.
Uzbekistan
Kyrgyzstan

5. Sub-National Analysis of Capabilities

Section 3 developed an analysis of Kazakhstan's export trajectory and economic complexity at the national level. It leveraged global export data to explain Kazakhstan's comparative advantage in products and compared the development of the nonmineral economy against peers and other major oil exporters. This section expands upon the regional dynamics underlying the national export picture that was introduced in Section 3. We briefly characterize the export capabilities of Kazakhstan's regions. We then describe main regional complexity metrics. The analysis in this section includes information for individual regions but is largely focused on the macro regions introduced in the previous section.

Characterization of Regional Export Capabilities

The regions of Kazakhstan have unique productive capabilities. This is conveyed in the significant range in exports per capita across the country and large differences in the composition of their exports. In 2019, exports per capita were largest in Atyrau, exceeding $27,450, and lowest in Almaty at $204 (Fig. 5.1). Though the magnitude of the gap fluctuates with the country's exports in petroleum oils, the gap has persisted over time. In 2016, when Kazakhstan's exports of petroleum oils were the lowest since 2004, Atyrau's total exports per capita exceeded $11,200 and Almaty's stood at $172. Furthermore, nationally, exports per capita were $2,255 in 2019, but only five regions of 17—concentrating 10.8% of the population—had exports per capita above this amount. In exports of nonmineral products, the variation between regions is much smaller, but still, only five regions—foremost the Industrial Belt—have exports per capita above the national average of $795 in 2019 (Fig. 5.3). These results are consistent when aggregating for macro regions (Fig. 5.2 and 5.4).

Figure 5.1

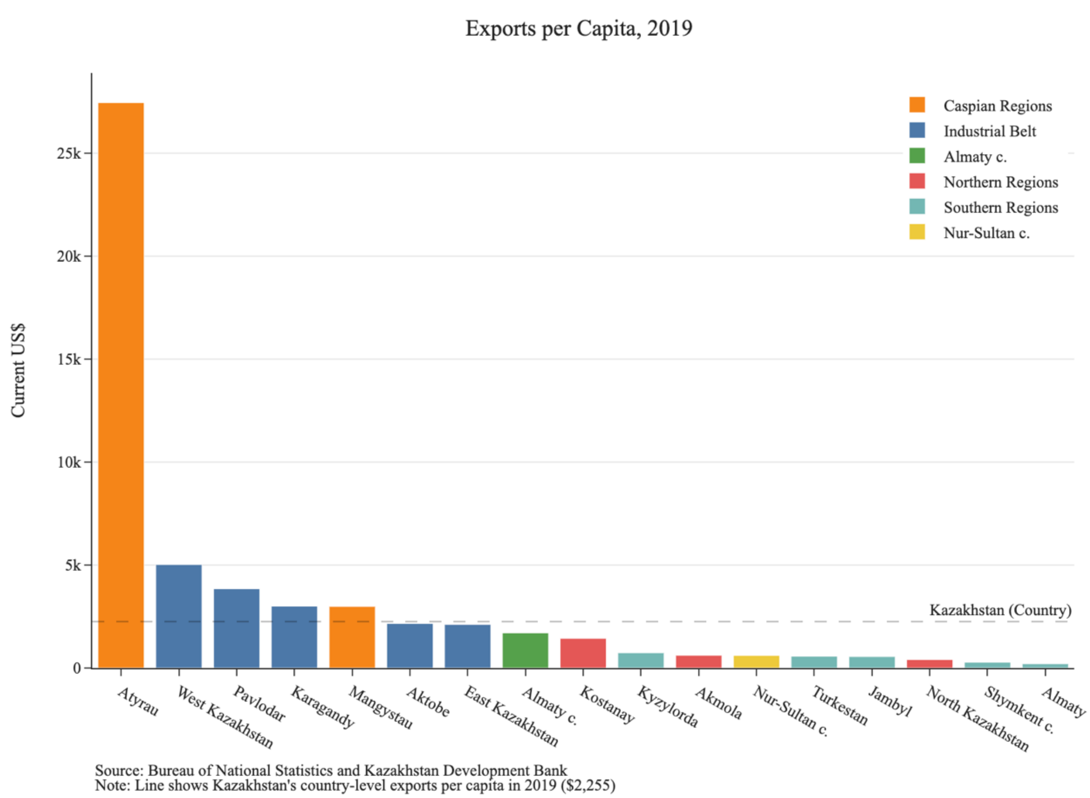


Figure 5.2

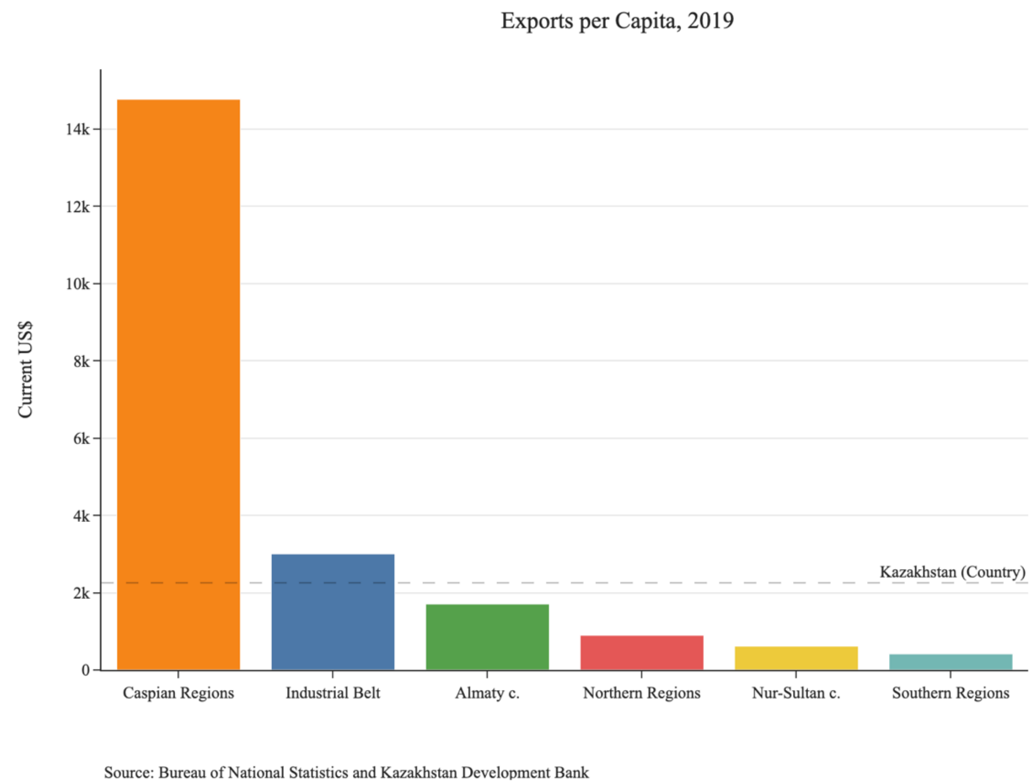


Figure 5.3

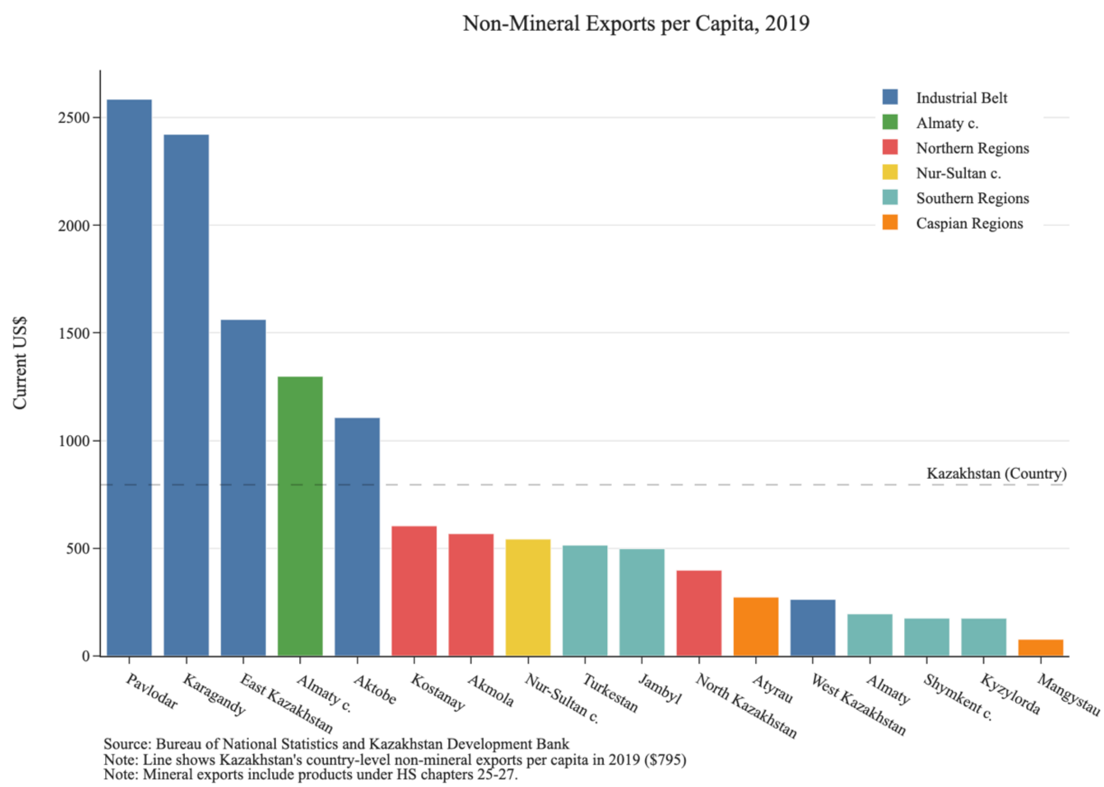
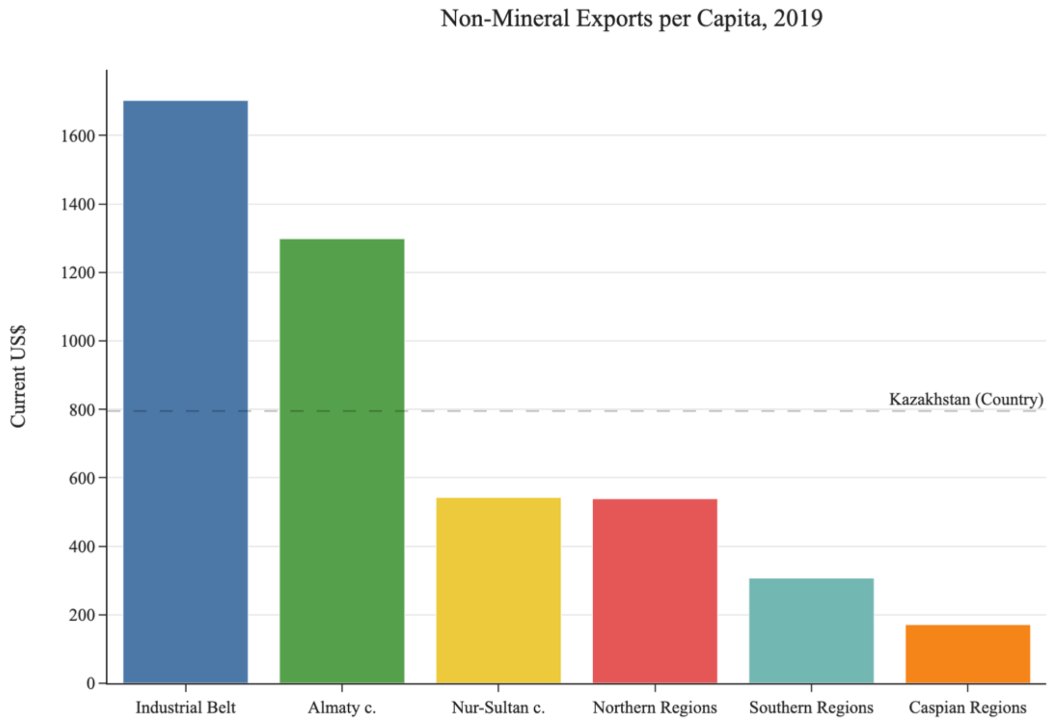


Figure 5.4



Source: Bureau of National Statistics and Kazakhstan Development Bank
Note: Mineral exports include products under HS chapters 25-27.

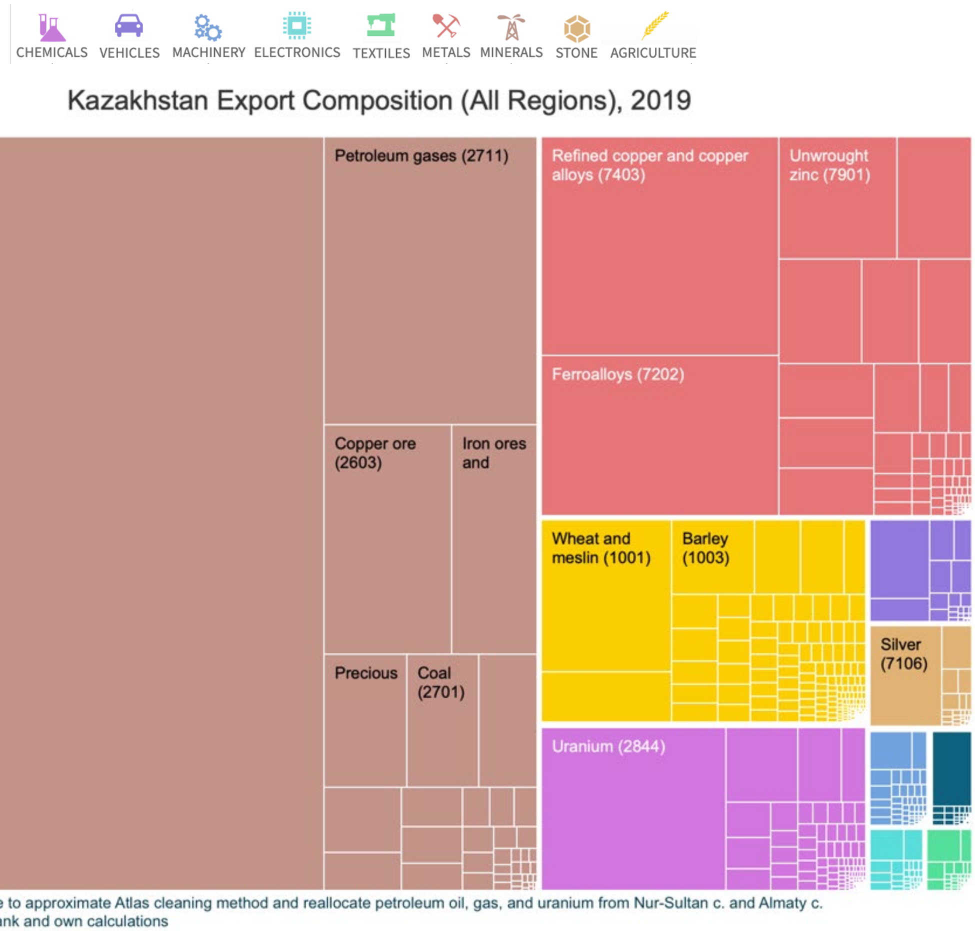
The composition of goods exported by each region reveals both their differential endowments and productive knowledge. The export compositions of Kazakhstan and each macro region constructed in this report are included here (Fig. 5.5). Appendix D includes the export compositions of all regions in 2019 (Fig. D1).[30]

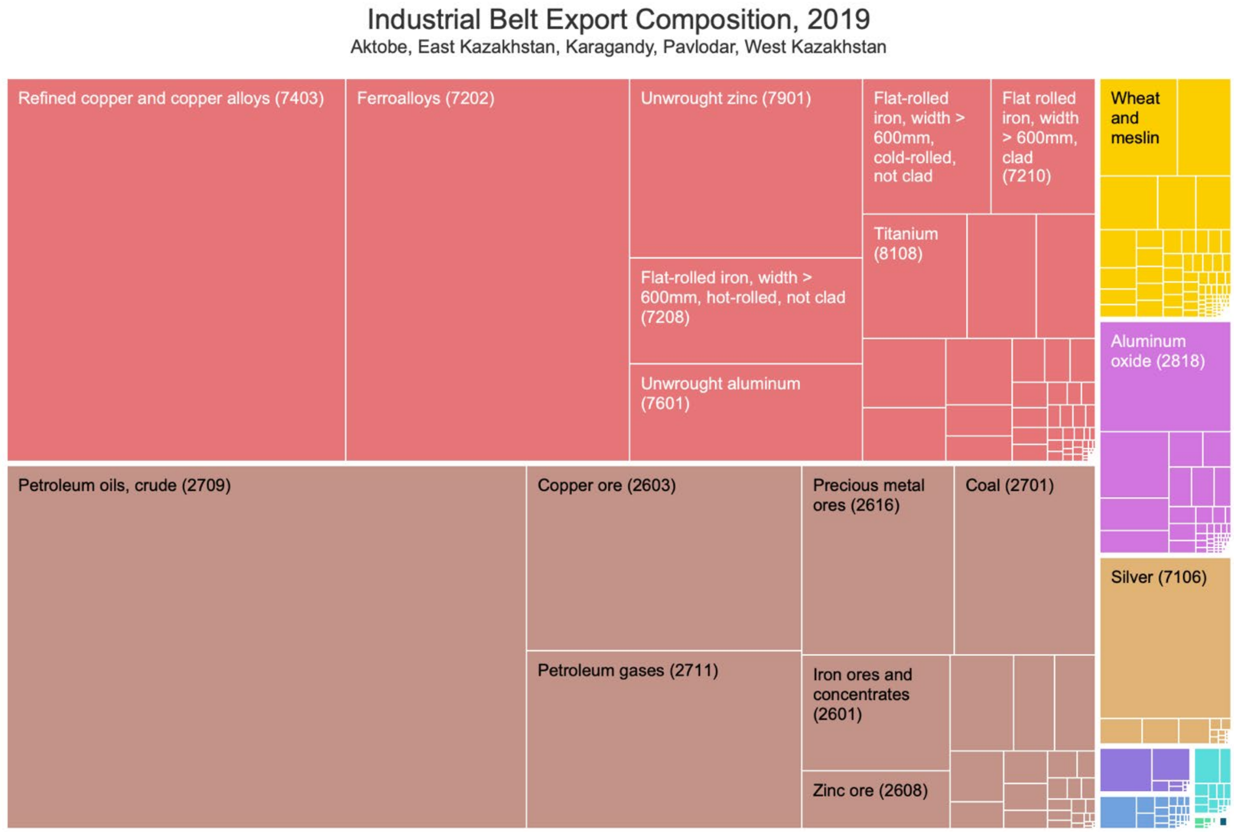
The Industrial Belt macro region demonstrates a concentration of exports in metals (~45%) and minerals (~43%), particularly in products such as Crude Petroleum (~21%), Refined Copper (~14%), Ferroalloys (~12%) and Copper Ore (~6%). Exports from the Caspian Regions are almost entirely concentrated in minerals (~98%), particularly in Crude Petroleum (~87%). In the Northern Regions, exports are concentrated in agricultural products (~47%) and minerals (40%), with a particular focus on Iron Ore and Concentrates (~31%), Wheat and Meslin (~16%), Wheat and Meslin Flour (~9%) and Linseed (~5%). For the Southern Regions, the main export categories are chemicals (~42%) and minerals (~26%), with the top products being Uranium (~38%) and Crude Petroleum (~14%). The macro region also has a relatively important presence of agricultural products (~21%). Almaty City displays a significant share of exports in minerals (~24%), agriculture (~20%), metals (~17%) and chemicals (~17%). The products that concentrate the highest share of exports are Copper Ore (~15%), Refined Copper and Copper Alloys (~10%), Other Aircraft and Spacecraft (~7%), Wheat and Meslin (~7%) and Uranium (~5%). Lastly, Astana displays a significant concentration of exports

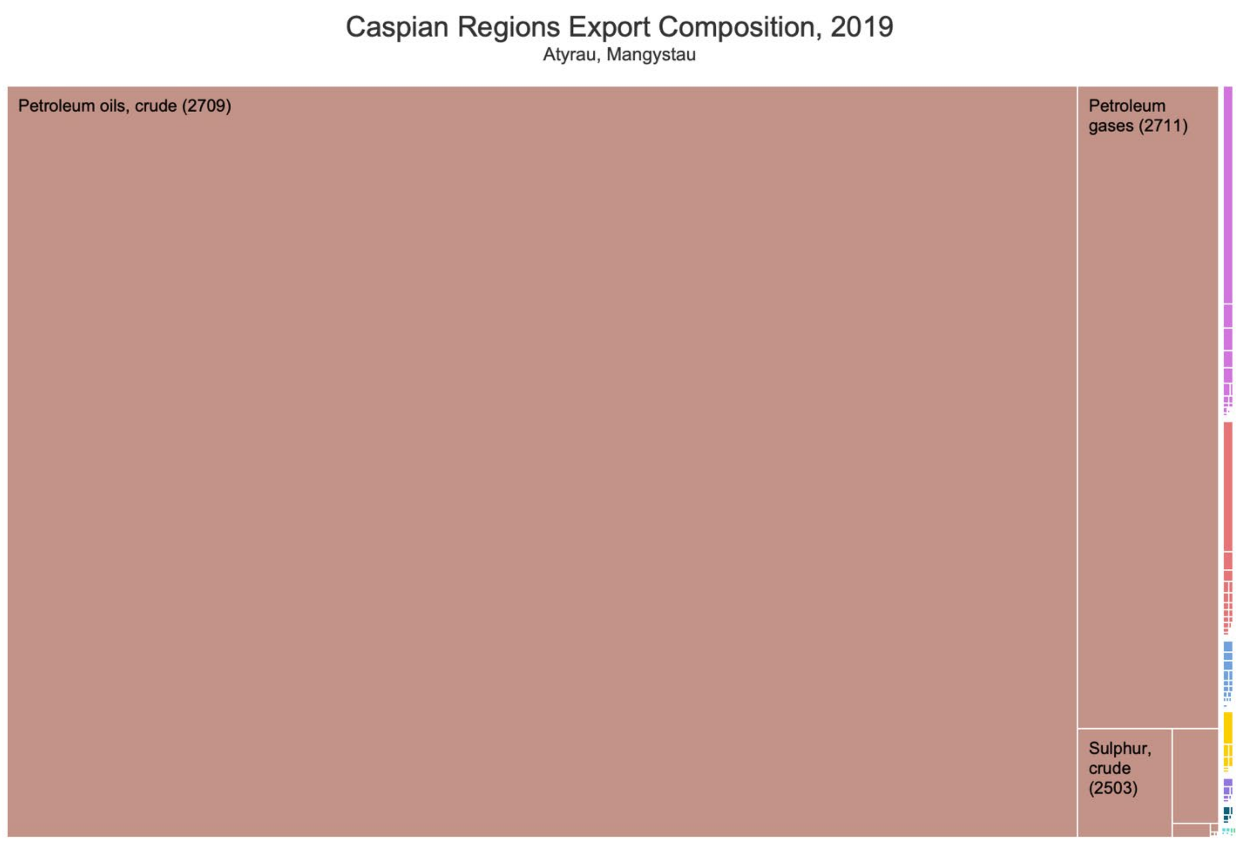
[30] Though the information available for regional exports extends back to 2012, we focus most of our analysis on 2019. The reason being that upon further inspection the export values for individual products in individual regions or macro regions is very volatile over time. In several instances we observe exports for individual products valued in several million dollars for non-consecutive years with values of zero exports in between. We also infer a potential re-allocation of product exports over time across regions and a classification change during the span of the series. All in all, these factors would introduce substantial noise to complexity metrics and may lead to a suboptimal assessment of regional productive capabilities.

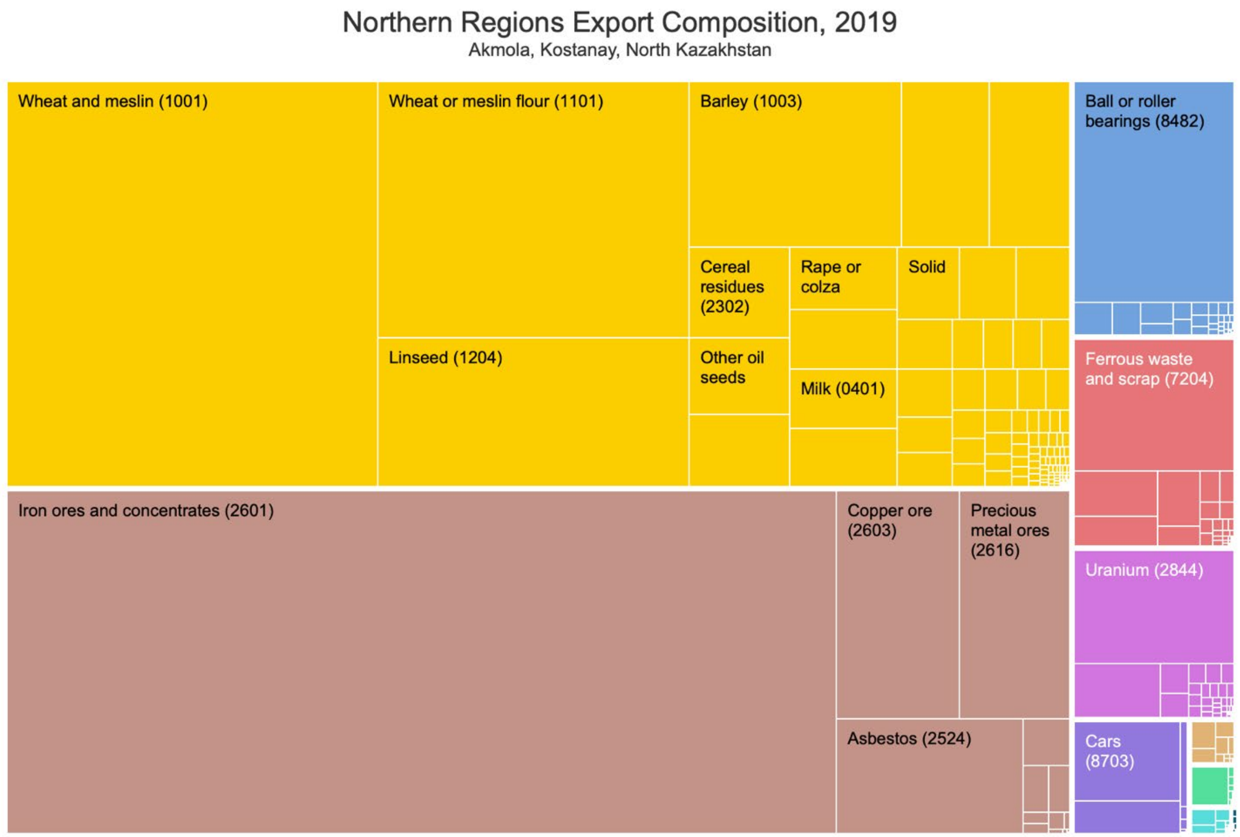
in agricultural products (~57%) and metals (~17%). The products that concentrate the largest share of exports in Astana are Wheat and Meslin (~31%), Barley (~12%), Unwrought Aluminum (~8%) and Electric Trains (~5%). It should be noted that the heavy concentration of exports in agricultural products and metals in Almaty City and Astana reinforces the concern of headquarter problems highlighted earlier.

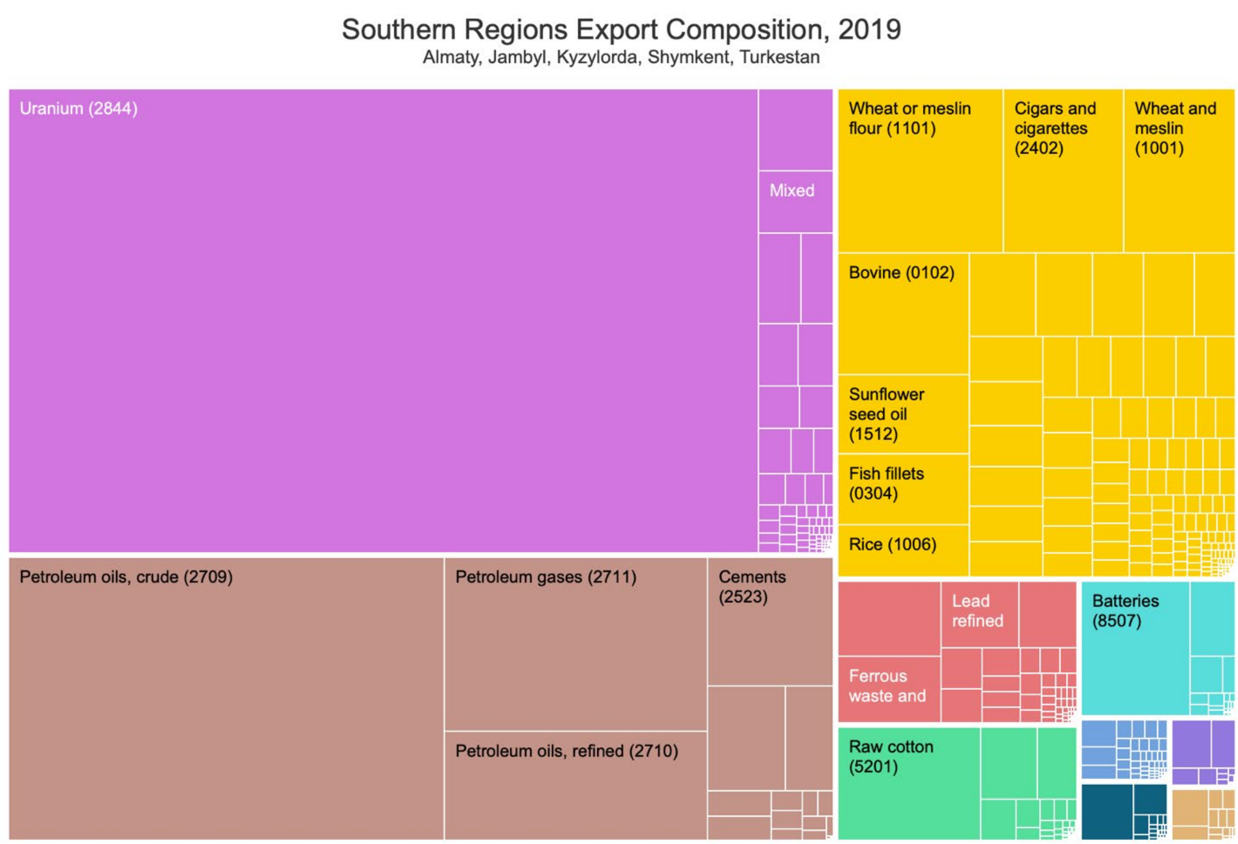
Figure 5.5



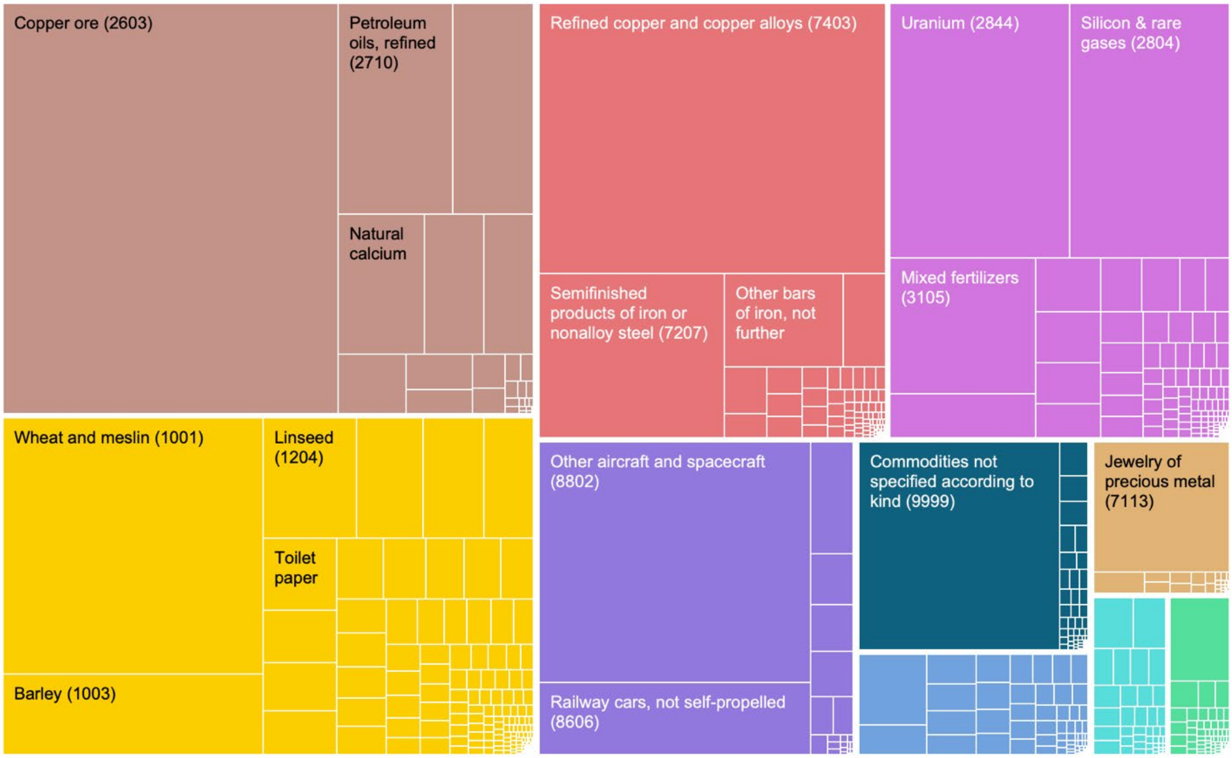

Industrial Belt Export Composition, 2019
Aktobe, East Kazakhstan, Karagandy, Pavlodar, West Kazakhstan
Refined copper and copper alloys (7403)
Ferroalloys (7202)
Unwrought zinc (7901)
Flat-rolled iron, width > 600mm, cold-rolled, not clad
Flat rolled iron, width > 600mm, clad (7210)
Wheat and meslin
Titanium (8108)
Flat-rolled iron, width > 600mm, hot-rolled, not clad (7208)
Unwrought aluminum (7601)
Aluminum oxide (2818)
Petroleum oils, crude (2709)
Copper ore (2603)
Precious metal ores (2616)
Coal (2701)
Silver (7106)
Petroleum gases (2711)
Iron ores and concentrates (2601)
Zinc ore (2608)


Caspian Regions Export Composition, 2019
Atyrau, Mangystau
Petroleum oils, crude (2709)
Petroleum gases (2711)
Sulphur, crude (2503)

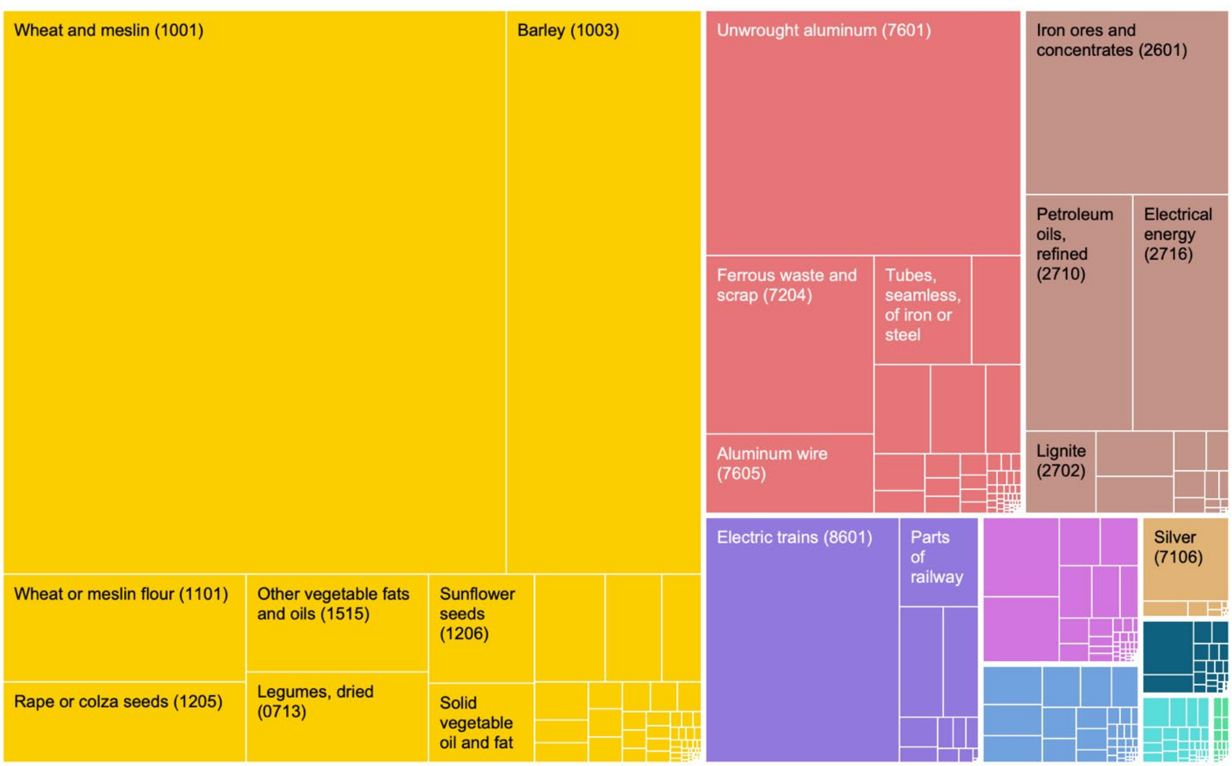

Northern Regions Export Composition, 2019
Akmola, Kostanay, North Kazakhstan
Wheat and meslin (1001)
Wheat or meslin flour (1101)
Barley (1003)
Cereal residues (2302)
Rape or colza
Solid
Linseed (1204)
Other oil seeds
Milk (0401)
Ball or roller bearings (8482)
Ferrous waste and scrap (7204)
Iron ores and concentrates (2601)
Copper ore (2603)
Precious metal ores (2616)
Uranium (2844)
Asbestos (2524)
Cars (8703)


Southern Regions Export Composition, 2019
Almaty, Jambyl, Kyzylorda, Shymkent, Turkestan
Uranium (2844)
Mixed
Wheat or meslin flour (1101)
Cigars and cigarettes (2402)
Wheat and meslin (1001)
Bovine (0102)
Sunflower seed oil (1512)
Fish fillets (0304)
Rice (1006)
Petroleum oils, crude (2709)
Petroleum gases (2711)
Cements (2523)
Petroleum oils, refined (2710)
Lead refined
Ferrous waste and
Batteries (8507)
Raw cotton (5201)

Almaty c. Export Composition, 2019

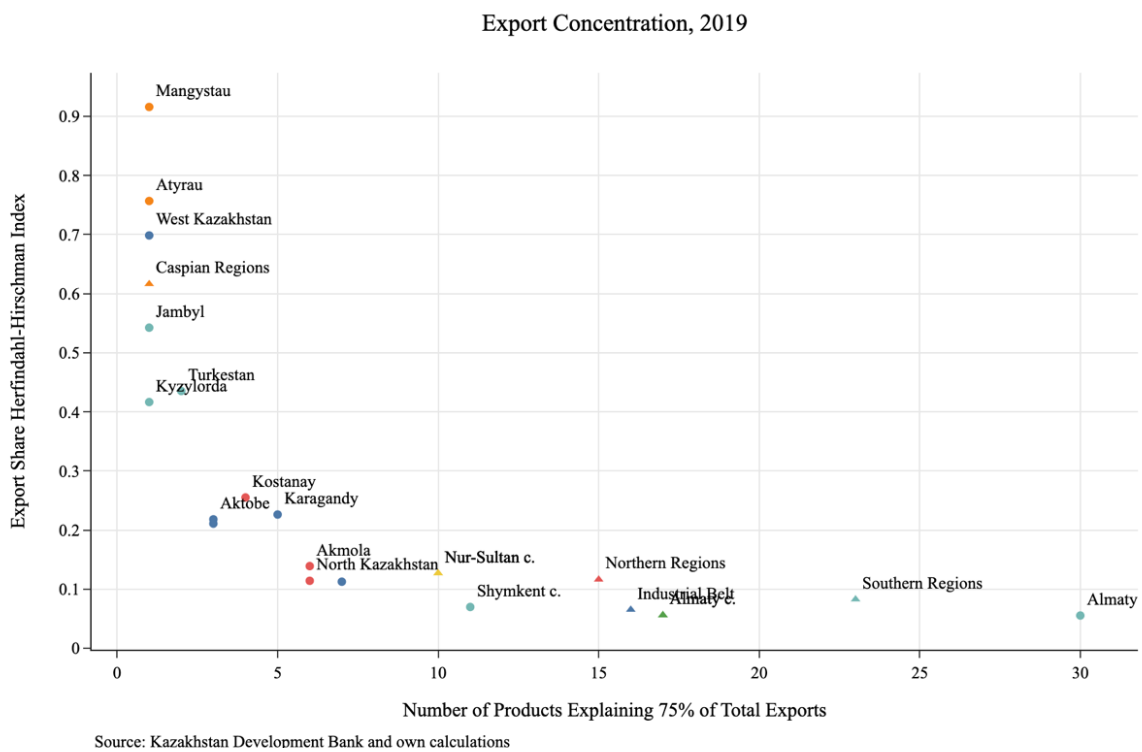


Nur-Sultan c. Export Composition, 2019



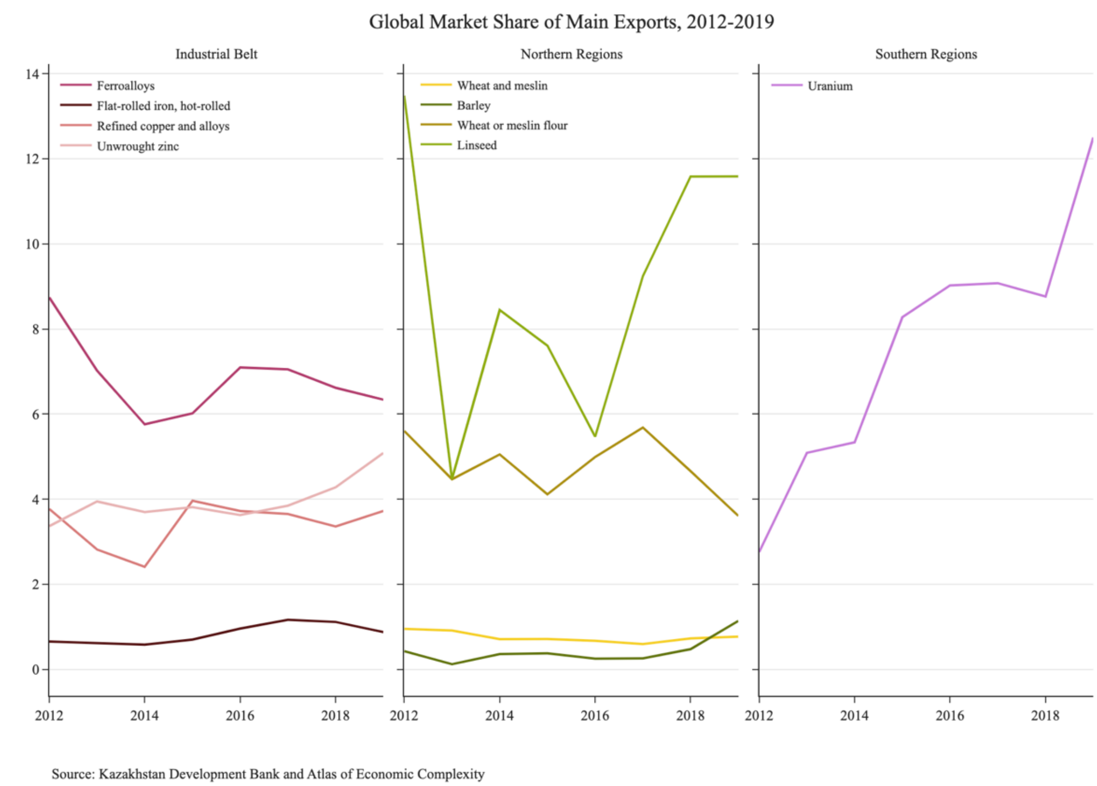
As the above figures illustrate, the export baskets of the macro regions are often intensive in a particular sector and dominated by a handful of products within it. However, it should be noted that this concentration tends to be even more pronounced in individual regions. In ten of the 17 individual regions, just 5 products or less explain 75% of the region's total export basket in 2019 (Fig. 5.6). This is only the case for the Caspian Regions. The Herfindahl-Hirschman index can also be leveraged to express how concentrated the export baskets are. The index is calculated as the sum of squared export shares of each product in the basket. Regions including Mangystau, Atyrau, West Kazakhstan, Kyzylorda and Jambyl that have over 75% of total exports concentrated in one product have consequently very high scores on the index (ranging from 0 to 1). Other regions with more diversified export baskets such as East Kazakhstan, Almaty and North Kazakhstan and the cities of Astana, Almaty and Shymkent have lower concentrations of their export basket in a small number of products. Overall, the export basket of macro regions is less concentrated than individual regions, as can be seen in Fig. 5.6. This might lead to more attractive diversification opportunities than what may be possible when considering only individual regions.

Figure 5.6



Many products tend to be exclusively or overwhelmingly exported from one macro region, and the macro regions capture significant global market shares of their main products. For example, the Industrial Belt plays an outsized role in the export of Refined Copper and Ferroalloys. These two products had a combined export value in the macro region of \$3.95 billion, accounting for 86.4% and 99.0% of Kazakhstan's total exports in the products, respectively. In 2019, the macro region's global market shares in the products were 3.7% and 6.3%, respectively. The Northern Regions represent a major player in the exports of Linseed and Wheat and Meslin Flour. They represent over 50% of Kazakhstan exports in these products and capture 11.5% and 3.6% of the global market, respectively. Likewise, in the Southern Regions Uranium comprised over 37.9% of total exports, representing 84.1% of Kazakhstan's total exports in the product and capturing 12.5% of its global market in 2019 (Fig. 5.7).

Figure 5.7



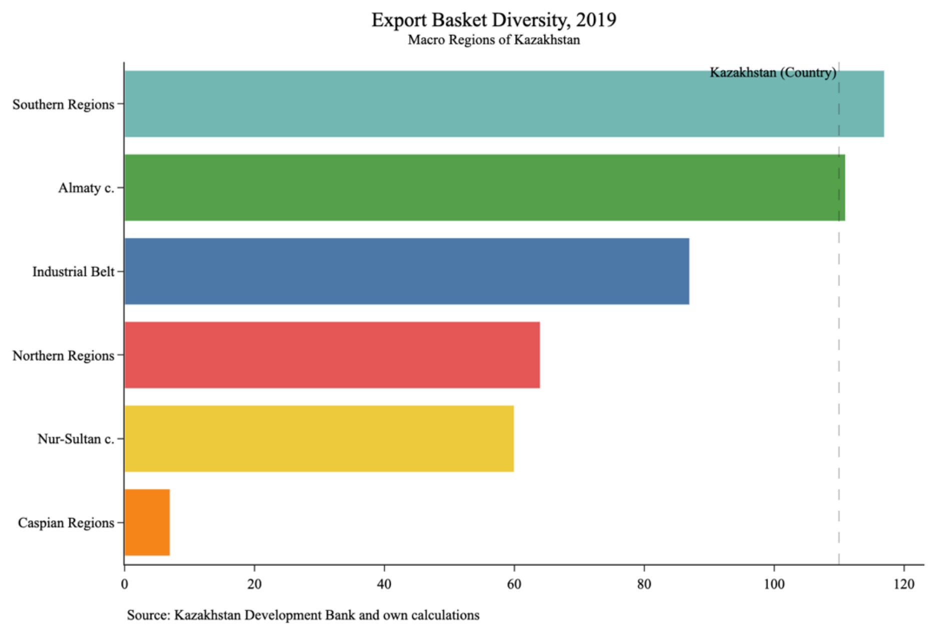
The significant competitive advantage that macro regions display in some products is one motivation for pursuing a strategy that considers both the intensive and extensive margins (i.e. building on an existing export versus developing a new one). The identified diversification opportunities for each macro region will include options at the intensive margin that targets expansion of activities it already has success in. The intensive margin recommendations are paired with extensive margin recommendations, which may present opportunities in other sectors that show promise based on their proximity to current capabilities among other factors. These complementary approaches will be further detailed in the following sections.

Subnational Complexity Metrics

The differences in endowments and productive capabilities highlighted so far similarly translate into varying performance on measures of economic complexity across macro regions. We calculated complexity metrics such as diversity, ubiquity and ECI at the subnational level in a similar manner as the country level. These measures were calculated for both individual regions and macro regions. This section will focus on the complexity measures for macro regions, while Appendix D includes the complexity measures at the level of individual regions (Fig. D2).

In terms of diversity, RCAs are used to measure whether regions have a comparative advantage in a product in relation to global trade. The threshold of RCA=1 is applied to determine a product's presence in the macro region's set of competitive exports (M_{cp}=1). The sum of products for which a macro region has RCA higher or equal to 1 represents the diversity of the macro region. We observe a large variance in terms of diversity across macro regions, whereby the Southern Regions and Almaty City—the most diverse macro regions—are ten times as diverse as the Caspian Regions (Fig. 5.8).

Figure 5.8



With regards to average ubiquity, the differences are less stark. However, less diverse macro regions tend to display a higher average ubiquity—the products they specialize in are products in which many other places also specialize in—while more diverse places tend to display a lower average ubiquity (Fig. 5.9). The performance on this metric is particularly salient for the Industrial Belt, which is the only region that displays a lower average ubiquity than Kazakhstan as a whole, and one of the lowest levels of average ubiquity for its levels of diversity when compared to all countries in the world (Fig. 5.10). This might imply that the Industrial Belt is able to leverage a relatively uncommon type of productive knowhow for its level of diversity.

Figure 5.9

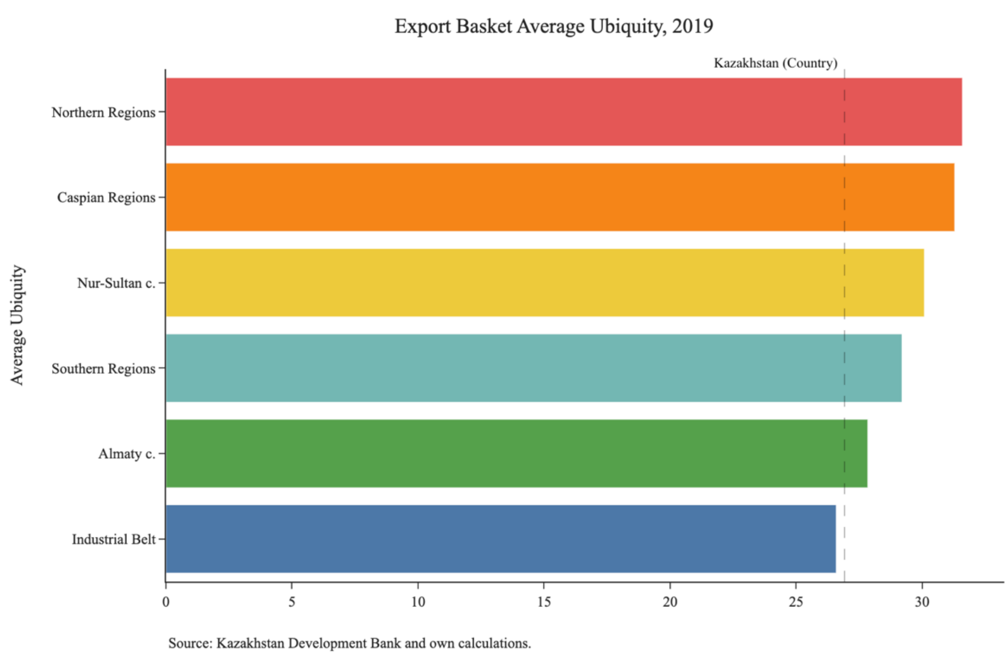
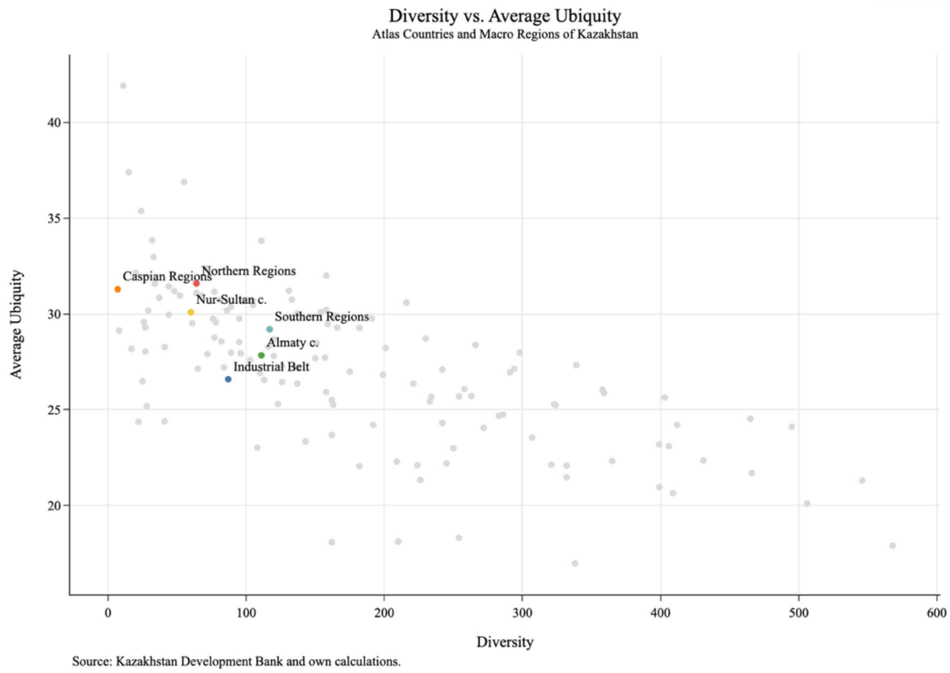
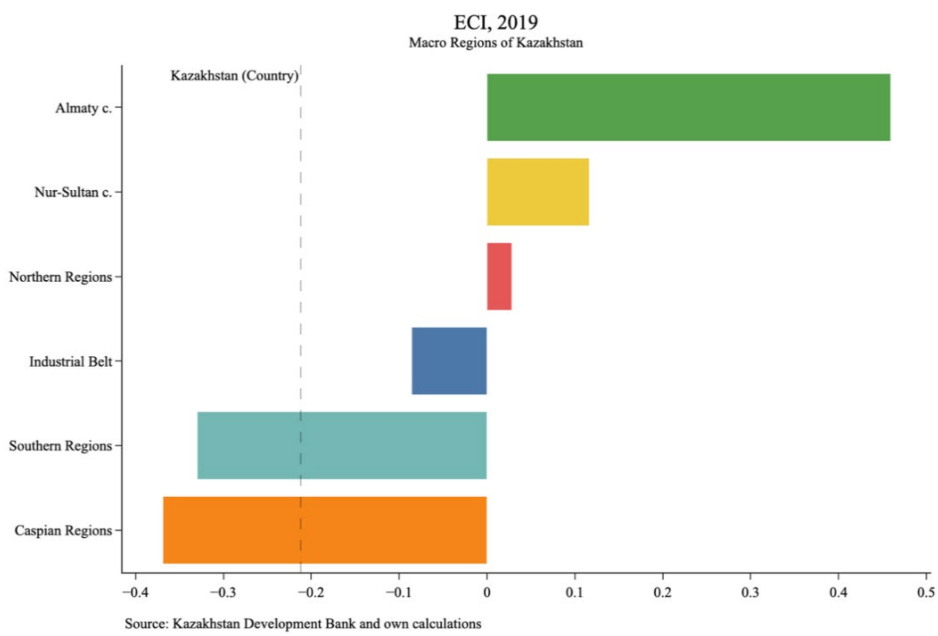


Figure 5.10



The combination of higher diversity and lower average ubiquity is generally associated with higher economic complexity (ECI). However, this relationship does not hold for most of the macro regions of Kazakhstan. Almaty City and the Caspian Regions do have an ECI consistent with expectations. But it is notable that Astana and the Northern Regions display a positive ECI despite its low levels of diversity and high average ubiquity. Similarly, it is surprising that the Southern Regions and the Industrial Belt display negative ECI—in the case of the Southern Regions even lower than Kazakhstan as a whole—despite respectively displaying the highest levels of diversity and lowest levels of average ubiquity (Fig. 5.11).

Figure 5.11



Decomposing the macro regions' ECI enables us to better understand the region's overall performance (Fig. 5.12 and 5.13). First, we assess whether the macro region concentrates most of its diversity in relatively complex types of products. Second, we consider whether the products in which a macro region specializes are, on average, higher or lower complexity products within each type of product. From this analysis we can better understand the relative underperformance of the Southern Regions and the Industrial Belt. In the case of the Southern Regions, close to 60% of its diversity is concentrated in low complexity categories such as agriculture and textiles. Furthermore, it displays the lowest average PCI in both categories. In other words, the Southern Regions are indeed very diverse, but are diverse in low complexity products within low complexity product categories. With regards to the Industrial Belt, close to 50% of its diversity is concentrated in agriculture and minerals, where it displays an average negative PCI. This overshadows the fact that a large share of its diversity (30%) is concentrated in metals and that it displays higher than average PCIs in vehicles.

Conversely, the positive performance of Astana could be explained by its positive average PCIs for 7 of 9 product categories and being among the two top performers in average PCI for 4 product categories. In the case of the Northern Regions, it displays a positive average PCIs for 5 of 8 product categories and is among the two top performers in average PCI for 4 product categories. Furthermore, even though over half of its diversity is concentrated in agricultural products, the average PCI of these products is relatively high for the sector.

Figure 5.12

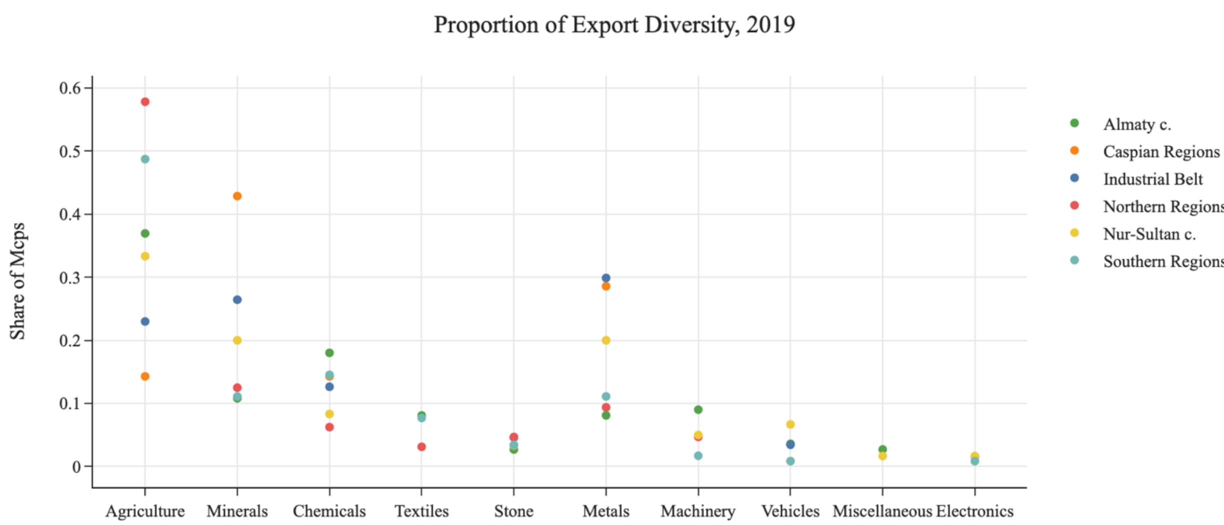
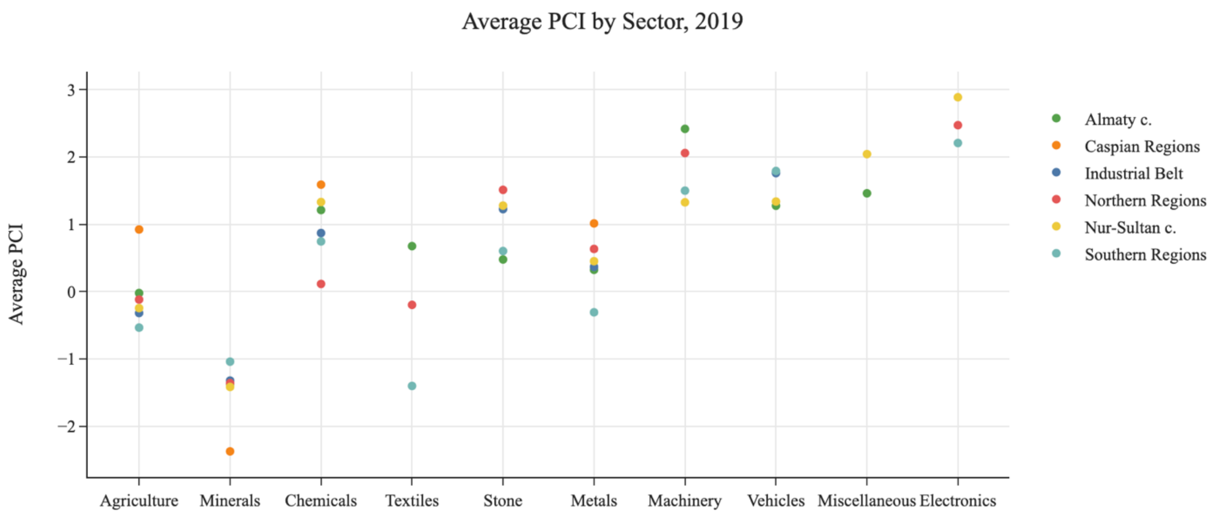


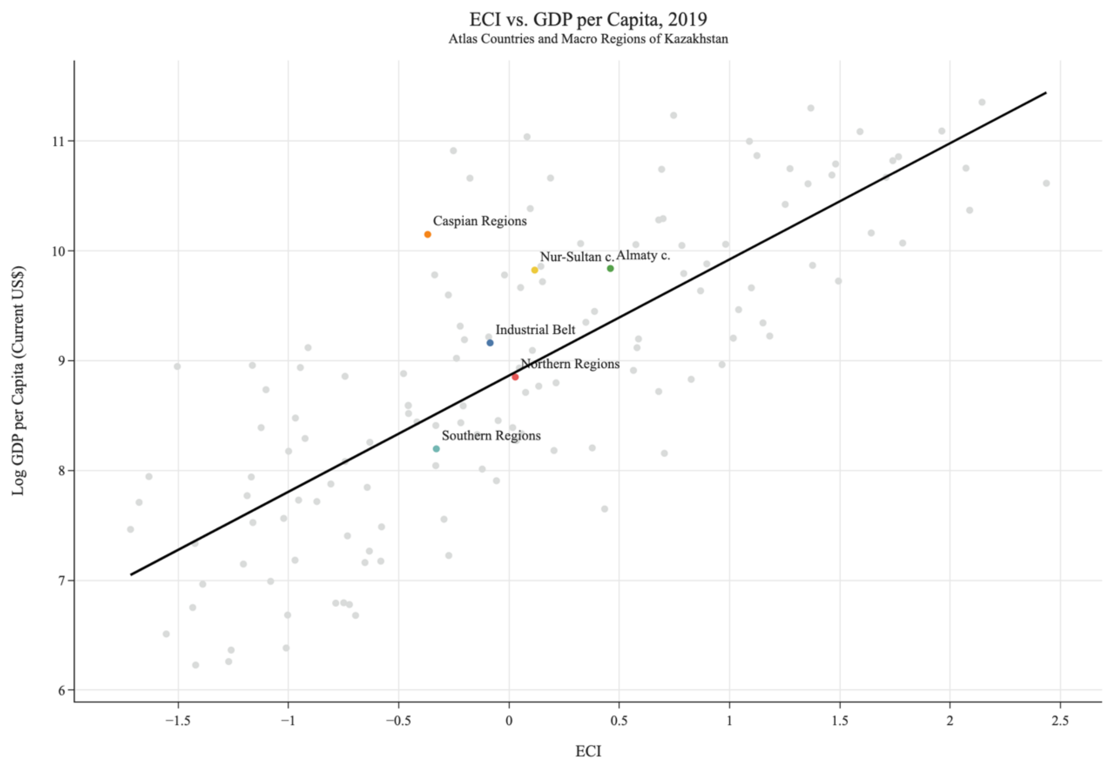
Figure 5.13



Section 2 described the empirical positive relationship between found globally between ECI and GDP per capita. Similar observations within this relationship at the country level can generally be applied at the subnational level. Locations with low complexity that are intensive in oil or other natural resources may have high current incomes because of temporary positive shocks; unless the booms can be leveraged to increase complexity of the economy, their growth can be particularly volatile to global commodity prices. In a subnational context, regions particularly intensive in commodities can suffer from labor market imbalances that are not as acute in other areas of the country. These can include wage stagnation, underutilization, lack of inclusion, and recovery that relies on flows of investments into the commodity sector.[31] On the other hand, locations with high complexity relative to their current incomes may be well-positioned for periods of sustained growth. If existing constraints can be overcome, the existing knowhow can translate into new exports and higher output in the future. Fig. 5.14 illustrates the relationship between ECI and GDP per capita for 133 countries and the macro regions of Kazakhstan. It should be noted that these intuitions may not extend fully to major cities within a country. This is because major cities can have a deep and globally competitive base of productive knowhow in services that are not captured in our current measure of ECI. Additionally, they might be uniquely suited to benefit from the multiplier effect of commodity booms, consumption and investment booms, expansionary fiscal cycles, and other positive exogenous shocks.

Figure 5.14

[31] Hausmann, R., et al., 2021. Growth Perspective on Western Australia.



Source: Kazakhstan Development Bank, Bureau of National Statistics, Atlas of Economic Complexity and World Bank WDI

This analysis of the existing productive capabilities of Kazakhstan's macro regions describes the wide variance in terms of export performance—proxied by exports per capita—and focus of specialization —proxied by export composition and diversity. Similarly, it outlines the differential capacity of macro regions to pursue diversification opportunities across different sectors and levels of sophistication. Lastly, it highlights the importance of leveraging existing globally competitive industries along the intensive margin. These findings serve as further motivation to explore diversification strategies at a macro region level and inform the product identification framework that will be explained in the next section.

6. Identification of Diversification Opportunities – Main Export-Based Analysis

This section outlines export diversification opportunities for each macro region of Kazakhstan. We first describe our understanding of the objectives of such an exercise. We then highlight the variables that we believe are most crucial to consider in the specific case of Kazakhstan and present a customized systematic framework to jointly examine these variables. Finally, we present the identified opportunities for each macro region and provide elements of interpretation.

Objectives and Limitations

The general objective of such an opportunity identification exercise is to uncover activities, products, services, or industries that may have potential to drive Kazakhstan's economic diversification. The slightly narrower objective of the methodology we outline below is to identify products and product themes deemed as attractive and viable export diversification opportunities based on Kazakhstan's existing productive capabilities.

Going from general considerations to an actual list of opportunities products comes at the cost of a few hypotheses, simplifications, and limitations. Some relate to data sources and data quality; others concern the general approach based on the concepts of productive capabilities and economic complexity. Regarding data, the approach we outline below is based on regional and national exports of goods. As such, it only captures tradable activities and leaves out non-tradable activities that are produced and consumed locally.[32] Furthermore, it only captures information about capabilities that are apparent in international trade and is hence largely oblivious of the production of tradable goods if they are mainly consumed on the domestic market. Finally, regional trade data is likely to be reported with a number of errors and biases that we could only partially correct for.

More fundamentally, the economic complexity approach in which this exercise is anchored is only one possible approach to understand the economic structure and its possible evolution. Hence, the fact that some products do not appear in the list of identified opportunities does not imply they must be excluded from future efforts, as there may be other evidence to substantiate their value in a diversification process.

In view of these limitations and other necessary methodological choices, the results of this identification exercise should not be viewed as a definitive set of recommendations, but rather as one internally consistent attempt in a broader and iterative process to discover and prioritize sectors with high potential for successful diversification, export growth and investment promotion.

Product Identification Framework

Outlining the Framework

Our framework was designed to be applied at a macro regional level. Given the differences in productive capabilities across regions and the potential constraints to cross-regional agglomeration of capabilities, the product identification process is conducted for internally cohesive macro regions. Macro regions seem to present the necessary granularity to highlight these unique productive capabilities, while mitigating against noise in the source data.

The product identification framework is in fact made up of two parallel frameworks, described in Fig. 6.1 and 6.2: one framework for new or nascent products that may be feasible (the "extensive margin"), and one for existing competitive exports that can be further scaled up (the "intensive margin"). The

[32] There however are several reasons to rely on tradable rather than non-tradable activities to drive economic diversification, especially in a country of 19M inhabitants.

RCA is the metric that is used to classify products in these two groups for each macro region. An extensive margin product, classified as such with an RCA < 1, indicates that while the macro region might export some of the product, it is not yet globally competitive in it and exports less of it than the average country in share of the macro region's total exports. Conversely, a product with an RCA>=1 indicates that the macro region is globally competitive in the product, comprising proportional or higher shares of its export basket than the average country.

Figure 6.1

Identification framework for extensive margin products

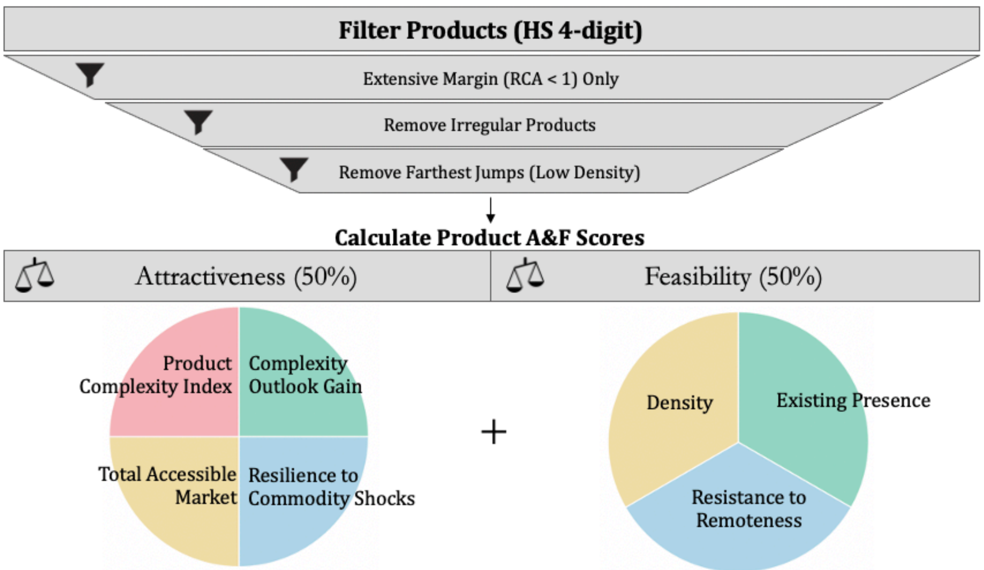
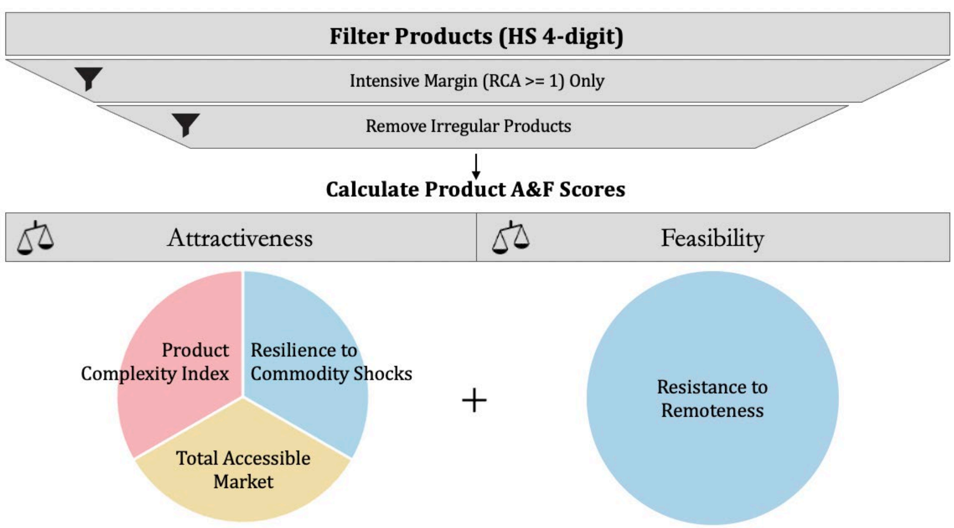


Figure 6.2

Identification framework for intensive margin products



Initial Filtering

An initial filtering of the universe of HS 4-digit products is undertaken to remove irregular products that do not present genuine diversification opportunities. Irregular products are classified as those:

- In the HS 2-digit categories of either "ores, slag, and ash" or "mineral fuels, oils, and waxes"; the location is either endowed with it or it isn't.
- In the HS 2-digit category of "art".
- Categorized as any kind of waste or scrap.
- Exported at positive volumes by fewer than 30 countries in any year of the data.
- Have a maximum global trade value across all years below $500 million.
- "Monopolized" products in which a single country was responsible for (a) 70% of total exports in any given year or (b) 50% of total exports every year.

The choices of filtering thresholds were made to be reasonable approximations of intrinsically meaningful levels of monopolization. While they were not subject to sensitivity analyses, it is likely that any change in these levels would result only in marginal additions or subtractions of products from the pool being considered.

In the case of the extensive margin products, an additional filter is applied to exclude products that present a long distance to the existing export basket of the macro region. The goal is to limit recommendations that may perform particularly well in "attractiveness" measures but that might be unfeasible because they represent "too long of a jump" given the macro region's current set of productive capabilities. A threshold was set based on a machine learning algorithm, roughly equivalent to excluding products for which, based on global co-export patterns, the macro region at hand would have less than a 20% likelihood of exporting that product competitively.

Attractiveness and Feasibility Factors

The remaining products are evaluated along several dimensions of feasibility and attractiveness. Feasibility factors address qualities of products that may make them more or less likely to thrive in each of the macro regions of Kazakhstan. Attractiveness factors address qualities of products that can make a positive contribution to the local economy, through higher economic complexity, larger demand for the product in question, or other specific policy objectives, such as capturing more of local markets or strengthening engines of growth that are resilient to the type of commodity shocks that in the past may have derailed diversification efforts.

The attractiveness factors include:

- Product Complexity Index (PCI)
- Complexity Outlook Gain (COG) (only for extensive margin products)
- Resilience to commodity shocks
- Size of the total addressable markets (i.e., the serviceable demand for the product), corrected for distance to import markets and product sensitivity to distance

The feasibility factors include:

- Existing presence (only for extensive margin products)
- Resistance to remoteness
- Density, i.e., the relatedness with existing exports (only for extensive margin products)

This attractiveness and feasibility framework is based on our understanding of the specific opportunities and challenges in Kazakhstan. Alternative weighting schemes of factors in the extensive and intensive margin frameworks would change the resulting lists of identified opportunities. The current allocations of factor weights seek to strike the balance between identifying products that are achievable given Kazakhstan's constraints while allowing room for longer jumps with potentially higher payoffs. Further research and other inputs for policy design can inform whether different factors express differentially binding constraints and should be weighed accordingly.

Attractiveness Factors

Product Complexity Index (PCI): PCI captures the amount and sophistication of productive knowhow required to produce a product. It considers the average diversity of locations that make a specific product, and the average ubiquity of the other products that those locations make. All else equal, an export opportunity with a higher PCI should be pursued as it helps the location expand the sophistication of productive capabilities. PCI is a product-specific metric that does not vary by location.

Across the universe of products in global trade except irregular products, those with the highest PCI in 2019 include Apparatus and equipment for photographic laboratories (9010), Halides of nonmetals (2812) and Self-propelled railway coaches (8603), while products with the lowest PCI include Cocoa beans (1801), Natural rubber (4001), and Tea (0902).

Complexity Outlook Gain (COG): COG considers the fact that diversification can happen in several steps. It captures the likelihood that diversifying into a particular product can help unlock opportunities of higher complexity in the future. Thus, opportunities with high COGs can optimally support further export diversification. COG considers the distance of the specific product to other products the location is not yet exporting, and the complexity of those products. This is a product and location specific metric. For the same product, two different macro regions will have different COGs.

Total Addressable Market (TAM): TAM measures the size of the market that is accessible to export from Kazakhstan, considering the size of each import market, the sensitivity of the product to distance, and the distance to each import market. Products with larger global markets are more attractive as export opportunities, because they represent higher potential earnings. While demand for a product can be indicative of its growth potential as an export from Kazakhstan, TAM is treated as an attractiveness factor as opposed to a feasibility factor because demand for a product does not automatically translate into production, nor signal the economy has the productive capabilities to produce it. Considering the difficulties in exporting from Kazakhstan, rather than considering the total world market, the measure adjusts based on:

- The distance between each exporter and each importer;
- The distance between Kazakhstan and each importer; and
- The gravity model coefficients that indicate how much harder it becomes to trade a product over a longer distance.

Refer to Appendix C for full detail on how the TAM is calculated for each product. It should be noted that this measure of accessible market does not explicitly consider formal and informal barriers to access the market or the competitiveness landscape of these markets. It should also be noted that because product TAMs were computed relying on global trade data, they do not differ across macro regions.

Across the universe of products in global trade except irregular products, those with the largest TAM for Kazakhstan in 2019 include Cars (8703), Transmission apparatus for radio, telephone, and TV (8525), and Gold (7108), while products with the smallest TAM include Word processing machines (8469), Motion-picture film (3706), and Swords (9307).

Resilience to Commodity Shocks: Resilience to oil shocks captures the correlation between year-on-year differences in real crude oil prices with year-on-year differences in real global traded value of each product from 1995-2020. Products that are resilient to shifts in the price of oil may be especially attractive as diversification opportunities in Kazakhstan as they could provide alternative engines of growth during downturns in commodity markets. Because both an inverse correlation and an absence of correlation with the price of oil may prove attractive, our measure only penalizes positive correlation with oil price but does not particularly reward negative correlation. Across the universe of products in global trade except irregular products, those whose export values are most positively correlated with the price of oil include Acyclic hydrocarbons (2901), Ethers (2909), and Carbon (2803).

Feasibility Factors

Existing Presence: Existing presence captures the relative intensity of the location in each product. A prospective product is more likely to thrive in a location if it is already exported with some intensity. Existing presence is operationalized by calculating the product's Revealed Comparative Advantage (RCA), which is the ratio of the location's export share in the product divided by the global export share in the product.

Resistance to Remoteness: Resilience to remoteness measures the sensitivity of each product to physical distance. Given Kazakhstan's geographic distance to many key global markets, exports that are easier to trade over long distances may be more feasible for diversification in Kazakhstan. Each product's sensitivity value is the coefficient calculated from a gravity model of trade that considers export values across all countries and the distance between their centroids.

In practice, products span wide ranges of sensitivity to distance. For example, live animals such as bovine and poultry are highly sensitive to distance, presumably because it is difficult and expensive to safely transport such live animals very far. In contrast, goods like uranium with high ratios of value to weight are very tolerant to being traded over long distances.

Density: Density measures the proximity of a location's current capabilities to the product. Products with a higher density are easier to diversify into in that specific location because the location has already demonstrated it has similar capabilities, inferred through the products it currently exports competitively. In this analysis, we employ a novel machine-learned measure of density. See Appendix C for detail on this methodology.

Box: How Remoteness Affects Kazakhstan's Export Opportunities

Given that Kazakhstan is a country impacted by remoteness, it is important to evaluate potential export opportunities in terms of how easily Kazakhstan can connect to global markets. One way in which we do this is by calculating how sensitive each product is to being traded over long distances. We use a gravity model of trade, described in Appendix C, to quantify how much the trade of a product is expected to decay over a given distance.

In practice, products span wide ranges of sensitivity to distance. For example, live animals such as bovine and poultry are highly sensitive to distance, presumably because it is difficult and expensive to safely transport such live animals very far. In contrast, goods like uranium with high ratios of value to weight are very tolerant to being traded over long distances. As would be expected, Kazakhstan tends to trade distance-sensitive products only in its nearby vicinity — even if the product has a high comparative advantage — and distance-tolerant products much further afield. Central Asia and Russia are major export markets for live bovine from Kazakhstan, for example, and uranium is conversely exported all over the world (Fig. Box 1).

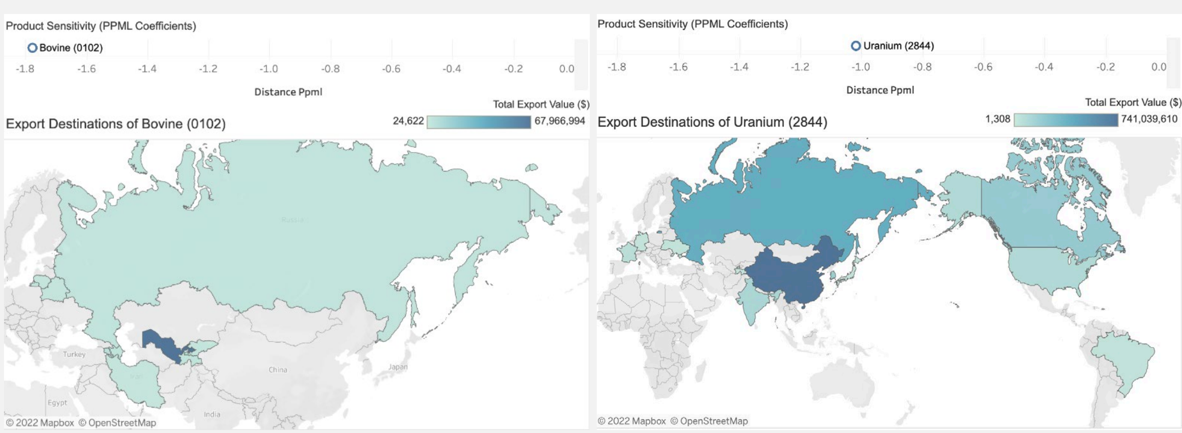
We additionally leverage this measure of distance sensitivity to calculate the Total Addressable Market (TAM) for each product that Kazakhstan faces. Whereas normally the TAM for an export product is comprised of all world trade in that product, we adjust Kazakhstan's TAM depending on 1) how distance-sensitive a product is and 2) how far each exporter is from each importer, versus how far Kazakhstan is from each importer. See Appendix C for details.

We use each product's measure of distance sensitivity as a Feasibility factor to evaluate each export, given that ceteris paribus Kazakhstan will arguably have an easier time entering markets for products that are distance-tolerant. We additionally use the TAM as an Attractiveness factor, given that larger world markets for an export lead to a higher potential payoff for expanding into them.

Figure Box 1

Sensitivity and Destinations of Kazakhstan's Exports of Bovine versus Uranium

Bovine *Uranium*



Normalizing and Combining Factors

Because each of the above factors has a different scale, it is necessary to normalize the scores before combining them in a meaningful way. Values are normalized using a standard z-score normalization with two exceptions. Values for the Total Addressable Market (TAM) factor were taken in logarithms before normalization. Values for Resilience to Commodity Shocks were normalized such that products with negative and nonsignificant correlations between their export value and the price of oil were given the highest score. The penalty on the highest score increases with the magnitude of a positive correlation between export value and the price of oil. For each factor except Resilience to Commodity Shocks, the resulting product scores follow a normal distribution with a mean score of 0.

For extensive margin products, the four attractiveness factors are averaged to obtain a composite attractiveness score and the three feasibility factors are averaged to obtain a composite feasibility score. The overall product score is calculated as the sum (equal weighting) of the composite attractiveness and feasibility scores.

The intensive margin selection process omits density, existing presence, and COG, because the country has already demonstrated significant presence in the product and thereby its feasibility. Furthermore, it has already realized "Complexity Outlook Gains" associated with its presence. Each of the four remaining factors (PCI, TAM, Resilience to Commodity Shocks and Resistance to Remoteness) are averaged to obtain the overall product score (Table 6.1).

It is worth noting that future efforts that leverage this information for policymaking may opt for a different weighting of factors that better reflects policy priorities or relevant constraints. Factor weights can be adjusted in the online tool detailed in Appendix B.

Table 6.1

Summary of Factors by Framework

	Factor	**Extensive Margin**	**Intensive Margin**	**Specific to Macro Region**
Attractiveness	PCI	X	X	
	COG	X		X
	TAM	X	X	
	Resilience to Commodity Shocks	X	X	
Feasibility	Density	X		X
	Existing Presence	X		X
	Resistance to Remoteness	X	X	

From Product Scores to Opportunities

The lists of opportunities for each macro region are constructed by considering the overall scores calculated for products at the extensive and intensive margins.[33] Rather than allocating a fixed number of opportunities per macro region, the number of recommendations per macro region are customized to only retain the most attractive and feasible opportunities. The formula to obtain the number of recommendations per macro region was formed from two core reasonings. The first is that the number of recommendations should scale with the capabilities demonstrated by the location. This ensures that locations with more diverse productive capabilities such as Almaty City and Astana are offered a wider range of recommendations, and locations such as the Caspian Regions are only presented the highest scoring opportunities. The second is that the macro regions should be presented enough recommendations such that diversification themes become identifiable, but not too many recommendations such that the lists are less targeted or require another set of filtering for discussion on a strategy.

The number of recommendations is given by the resulting formula. Given *n*, the number of products that the macro region exports with a comparative advantage, the number of recommendations is 30 products or 30% of *n,* whichever is lower. However, we retain a minimum of 10 products. This process is carried out separately for the extensive and intensive margin product lists, so the maximum number of possible recommendations is 30 at the extensive margin + 30 at the intensive margin = 60 products. The minimum number of possible recommendations is 10 at the extensive margin + 1 at the intensive margin = 11.[34] Table 6.2 presents the number of identified opportunities per macro region.

Table 6.2

Number of Opportunities Identified per Macro Region

Macro Region	Regions	Diversity	Extensive Margin Recommendations	Intensive Margin Recommendations	Total Recommendations
Industrial Belt	Aktobe, East Kazakhstan, Karagandy, Pavlodar, West Kazakhstan	96	28	29	57
Caspian Regions	Atyrau, Mangystau	8	10	5	15
Northern Regions	Akmola, Kostanay, North Kazakhstan	74	22	22	44
Southern Regions	Almaty, Jambyl, Kyzylorda, Shymkent, Turkestan	145	30	29	59
Almaty City	Almaty c.	127	30	28	58
Nur-Sultan	Nur-Sultan c.	72	22	22	44

Results and Potential Groupings of Diversification Opportunities

This section describes the resulting lists of diversification opportunities for each macro region and potential themes across product recommendations.

Identification of Themes

[33] While product scores for factors existing presence, density and COG are unique to each macro region, product scores for PCI, TAM, resistance to remoteness, and resilience to commodity shocks are calculated using global data at the country level. Thus, the scores for a given product in these four factors are the same across every macro region. As a result, the factors existing presence (RCA), density and COG are the source of variation in recommendations across the macro regions, after filtering out the furthest products from each macro region's capabilities.

[34] By construction of RCAs in global trade there will always be at least one product with M_{cp}=1.

To identify themes across individual product recommendations, a clustering algorithm[35] was applied to cluster products into useful groupings[36]. Manual adjustments were applied to move products that were originally unassigned by the clustering algorithm into the nearest cluster. Then, the clusters were consolidated, named and several outlier products at the periphery of the space were removed. In total, this process yielded 172 unique product recommendations across all macro regions, grouped in 9 broader categories and 29 themes (Table 6.3 and Fig. 6.3) See Appendix D for a mapping of the product space illustrating all recommendations and their assigned clusters (Fig. D3 and Fig. D4).

Table 6.3
Broad Categories and Themes

Broad Category	Number of Products	Theme	Number of Products
Chemicals and Allied Industries	29	Amino-compounds and other organic chemicals	10
		Inorganic chemicals	8
		Fertilizers, pesticides and chemicals used for their production	4
		Plastic articles	4
		Pharmaceutical products	3
Transportation	28	Aircarft, surveilance and defense equipment	8
		Trains, train parts and train equipment	8
		Vehicle and vehicle parts	7
		Vessels	5
Meat, Dairy, and Other Processed Agricultural Products	26	Meat, animal, and dairy products	9
		Processed animal or agricultural products	8
		Furskins and animal hair	3
		Sugar, cocoa and tobacco	3
		Paper and paper products	3
Metals	26	Soft metals and articles of soft metal	13
		Hard metals and byproducts	5
		Precious metals and copper	5
		Rare earth metals, metal oxides, uranium and related equipment	3
Fruits, Vegetables, and Cereals	21	Cereals, honey and oil seeds	10
		Vegetables and legumes	6
		Fruits and nuts	5
Construction	17	Construction materials and products	17
Machinery	14	Metalworking and other industrial machinery	6
		Agricultural machinery	3
		Turbines and generators	3
		Electrical machinery and equipment	2
Professional Equipment	8	Laboratory apparatuses	5
		Office machines	3
Precious Stones and Jewelry	3	Precious stones and jewelry	3

35 A UMAP dimension reduction algorithm and HDBSCAN clustering algorithm were used.
36 The point of using such an algorithm is to create groupings that are driven by objective similarities in which products co-occur with which others. This stands in contrast to human-chosen groupings, which are ultimately subjective.

Figure 6.3
Illustrating Broad Categories and Themes

National Patterns in Identified Products

While recommendations are specific to each macro region, certain products are recommended across several regions (Table 6.4). Barley (1003) is an identified opportunity at the intensive margin in all six macro regions. In 2019, Kazakhstan's exports of barley exceeded $295 million as 0.54% of the country's export basket and 4.58% of global trade in the product. Sunflower seeds (1206) and Nickel unwrought (7502) are identified opportunities across all macro regions except Caspian Regions

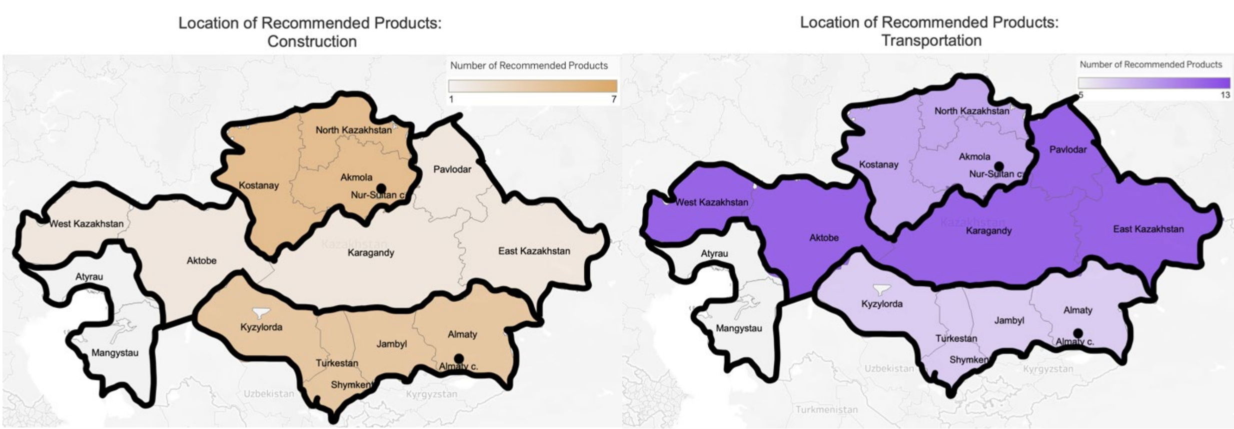
(Atyrau and Mangystau). Of the 172 unique recommended products, 14% are recommended across at least three of the macro regions and 61% are recommended to only one macro region.

Table 6.4
Products Identified in 3 or More Macro Regions

Product (HS Code)	Theme	Broad Category	Number of Macro Regions
Barley (1003)	Cereals, honey and oil seeds	Fruits, Vegetables, and Cereals	6
Sunflower seeds (1206)	Cereals, honey and oil seeds	Fruits, Vegetables, and Cereals	5
Nickel unwrought (7502)	Hard metals and byproducts	Metals	5
Uranium (2844)	Inorganic chemicals	Chemicals and Allied Industries	4
Worked cereal grains (1104)	Cereals, honey and oil seeds	Fruits, Vegetables, and Cereals	4
Rape or colza seeds (1205)	Cereals, honey and oil seeds	Fruits, Vegetables, and Cereals	4
Legumes, dried (0713)	Vegetables and legumes	Fruits, Vegetables, and Cereals	4
Other aircraft and spacecraft (8802)	Aircarft, surveilance and defense equipment	Transportation	4
Gas turbines (8411)	Aircarft, surveilance and defense equipment	Transportation	4
Floating structures for scrapping (8908)	Vessels	Transportation	4
Silicon & rare gases (2804)	Inorganic chemicals	Chemicals and Allied Industries	3
Pumps for liquids (8413)	Construction materials and products	Construction	3
Wheat and meslin (1001)	Cereals, honey and oil seeds	Fruits, Vegetables, and Cereals	3
Onions, shallots, garlic (0703)	Vegetables and legumes	Fruits, Vegetables, and Cereals	3
Butter (0405)	Meat, animal, and dairy products	Meat, Dairy, and Other Processed Agricultural Products	3
Animal feed (2309)	Processed animal or agricultural products	Meat, Dairy, and Other Processed Agricultural Products	3
Tubes, seamless, of iron or steel (7304)	Soft metals and articles of soft metal	Metals	3
Ferroalloys (7202)	Soft metals and articles of soft metal	Metals	3
Unwrought aluminum (7601)	Soft metals and articles of soft metal	Metals	3
Surveying instruments (9015)	Laboratory apparatuses	Professional Equipment	3
Tanks and other armored fighting vehicle (8710	Aircarft, surveilance and defense equipment	Transportation	3
Parts of railway locomotives (8607)	Trains, train parts and train equipment	Transportation	3
Railway track fixtures (8608)	Trains, train parts and train equipment	Transportation	3
Other vessels (8906)	Vessels	Transportation	3

We can observe macro regional specialization patterns in the recommendations generated for each macro region. (Fig. 6.4). Almost 50% of all recommended products in metals and more than 25% in transportation appeared in the Industrial Belt. A higher proportion of recommended products in construction appear in the Northern Regions and Southern Regions. Within opportunities in agriculture, Meat, Dairy, and Other Processed Agricultural Products are concentrated in the Northern Regions and Southern Regions, with a fewer number of product recommendations in Almaty City and Astana. On the other hand, every macro region has at least one opportunity in Fruits, Vegetables, and Cereals; every macro region excluding the Caspian Regions has at least 8 product recommendations in the broader category.

Figure 6.4



Detailed Results for the Industrial Belt of Kazakhstan

Aktobe, East Kazakhstan, Karagandy, Pavlodar, West Kazakhstan

The most prominent categories for the Industrial Belt of Kazakhstan are transportation and metals. Opportunities for new or nascent products tend to cluster around metals, aeronautics, trains, rare earth metals and metal products. Opportunities to scale up existing products revolve around metal products, both steel and precious, and agricultural products. The identified opportunities broadly reflect the adjacencies for regions combining heavy industry and a solid agricultural base.

Figure 6.5

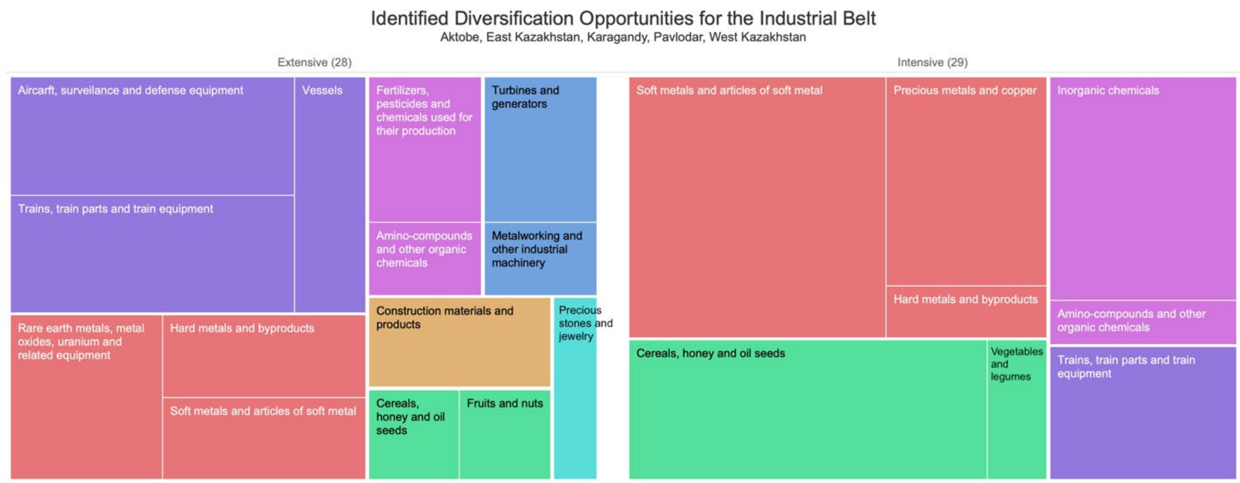


Table 6.5

Identified Diversification Opportunities: Extensive Margin Products

Product	Theme	Broad Category	Feasibility	Attractiveness	Overall Score
Nickel unwrought (7502)	Hard metals and byproducts	Metals	1.14	2.10	3.24
Other tubes and pipes, diameter > 406.4 (7305)	Soft metals and articles of soft metal	Metals	3.03	-0.06	2.97
Nuclear reactors and related equipment (8401)	Rare earth metals, metal oxides, uranium and related equipment	Metals	1.67	1.20	2.86
Railway cars, not self-propelled (8606)	Trains, train parts and train equipment	Transportation	2.89	-0.17	2.73
Clays (2508)	Construction materials and products	Construction	2.97	-0.33	2.64
Binoculars and telescopes (9005)	Aircarft, surveilance and defense equipment	Transportation	1.62	0.96	2.58
Tanks and other armored fighting vehicle (8710)	Aircarft, surveilance and defense equipment	Transportation	1.06	1.36	2.42
Steam turbines (8406)	Turbines and generators	Machinery	-0.01	2.38	2.37
Gas turbines (8411)	Aircarft, surveilance and defense equipment	Transportation	0.39	1.86	2.24
Railway coaches, not self-propelled (8605)	Trains, train parts and train equipment	Transportation	1.59	0.62	2.22
Floating structures for scrapping (8908)	Vessels	Transportation	2.13	0.07	2.20
Self-propelled railway coaches (8603)	Trains, train parts and train equipment	Transportation	0.54	1.64	2.18
Precious stones (7103)	Precious stones and jewelry	Precious Stones and Jewelry	2.26	-0.11	2.15
Hydraulic turbines, water wheels and reg (8410)	Turbines and generators	Machinery	1.74	0.35	2.09
Tomatoes, prepared or preserved (2002)	Fruits and nuts	Fruits, Vegetables, and Cereals	2.44	-0.39	2.05
Other vessels (8906)	Vessels	Transportation	0.23	1.82	2.05
Artificial graphite (3801)	Construction materials and products	Construction	1.15	0.84	1.99
Angles of iron or nonalloy steel (7216)	Soft metals and articles of soft metal	Metals	2.15	-0.18	1.97
Nitrogenous fertilizers (3102)	Fertilizers, pesticides and chemicals used for their production	Chemicals and Allied Industries	2.11	-0.14	1.97
Nickel mattes (7501)	Hard metals and byproducts	Metals	1.97	-0.05	1.93
Rare-earth metals (2805)	Rare earth metals, metal oxides, uranium and related equipment	Metals	2.38	-0.51	1.87
Lathes for removing metal (8458)	Metalworking and other industrial machinery	Machinery	-0.27	2.10	1.83
Other aircraft and spacecraft (8802)	Aircarft, surveilance and defense equipment	Transportation	0.83	0.96	1.79
Glycosides (2938)	Amino-compounds and other organic chemicals	Chemicals and Allied Industries	0.76	1.03	1.78
Other cereals (1008)	Cereals, honey and oil seeds	Fruits, Vegetables, and Cereals	2.61	-0.85	1.76
Electric trains (8601)	Trains, train parts and train equipment	Transportation	0.56	1.18	1.74
Phosphates (2835)	Fertilizers, pesticides and chemicals used for their production	Chemicals and Allied Industries	1.97	-0.29	1.68
Metal base oxides, n.e.c. (2825)	Rare earth metals, metal oxides, uranium and related equipment	Metals	1.94	-0.29	1.65

Table 6.6

Identified Diversification Opportunities: Intensive Margin Products

Product	Theme	Broad Category	Overall Score
Copper mattes (7401)	Precious metals and copper	Metals	0.76
Unrefined copper (7402)	**Precious metals and copper**	**Metals**	**0.62**
Amine-function compounds (2921)	Amino-compounds and other organic chemicals	Chemicals and Allied Industries	0.61
Uranium (2844)	**Inorganic chemicals**	**Chemicals and Allied Industries**	**0.54**
Silver (7106)	Precious metals and copper	Metals	0.52
Railway track fixtures (8608)	**Trains, train parts and train equipment**	**Transportation**	**0.52**
Salts of oxometallic acids (2841)	Inorganic chemicals	Chemicals and Allied Industries	0.52
Unwrought zinc (7901)	**Hard metals and byproducts**	**Metals**	**0.51**
Barley (1003)	Cereals, honey and oil seeds	Fruits, Vegetables, and Cereals	0.46
Ferroalloys (7202)	**Soft metals and articles of soft metal**	**Metals**	**0.45**
Sunflower seeds (1206)	Cereals, honey and oil seeds	Fruits, Vegetables, and Cereals	0.38
Parts of railway locomotives (8607)	**Trains, train parts and train equipment**	**Transportation**	**0.35**
Tubes, seamless, of iron or steel (7304)	Soft metals and articles of soft metal	Metals	0.33
Silicon & rare gases (2804)	**Inorganic chemicals**	**Chemicals and Allied Industries**	**0.31**
Unwrought aluminum (7601)	Soft metals and articles of soft metal	Metals	0.25
Wheat and meslin (1001)	**Cereals, honey and oil seeds**	**Fruits, Vegetables, and Cereals**	**0.24**
Lead refined unwrought (7801)	Soft metals and articles of soft metal	Metals	0.23
Carbides (2849)	**Inorganic chemicals**	**Chemicals and Allied Industries**	**0.22**
Rape or colza seeds (1205)	Cereals, honey and oil seeds	Fruits, Vegetables, and Cereals	0.21
Other bars and rods of other alloy steel (7228)	**Soft metals and articles of soft metal**	**Metals**	**0.17**
Sanitary ware and parts of iron or steel (7324)	Soft metals and articles of soft metal	Metals	0.16
Worked cereal grains (1104)	**Cereals, honey and oil seeds**	**Fruits, Vegetables, and Cereals**	**0.16**
Aluminum oxide (2818)	Inorganic chemicals	Chemicals and Allied Industries	0.09
Potatoes (0701)	**Vegetables and legumes**	**Fruits, Vegetables, and Cereals**	**0.08**
Other rail locomotives (8602)	Trains, train parts and train equipment	Transportation	0.08
Other alloy steel in primary form (7224)	**Soft metals and articles of soft metal**	**Metals**	**0.06**
Refined copper and copper alloys (7403)	Precious metals and copper	Metals	0.06
Sunflower seed oil (1512)	**Cereals, honey and oil seeds**	**Fruits, Vegetables, and Cereals**	**0.01**
Aluminum wire (7605)	Soft metals and articles of soft metal	Metals	0.00

Detailed Results for the Caspian Regions

Atyrau, Mangystau

The most prominent categories for the oil-rich Caspian Regions are chemicals and transportation. Opportunities for new or nascent products are heavily tilted towards fertilizers and vessels. Opportunities to scale up existing products revolve around fertilizers, the existing agricultural production, and vessels. Overall, the low number of identified opportunities and their concentration on a few themes reflect the specific diversification challenges of the oil-rich regions of Kazakhstan, with opportunities largely confined to downstream diversification from oil and gas activities. The notable presence of vessels may point to a different set of opportunities around the Caspian Sea, especially at a time when the importance of the sea as a trade route may be reinforcing. However, it may be important to validate on the ground whether these productive capabilities actually exist or are a byproduct of imperfect data reporting.

Figure 6.6

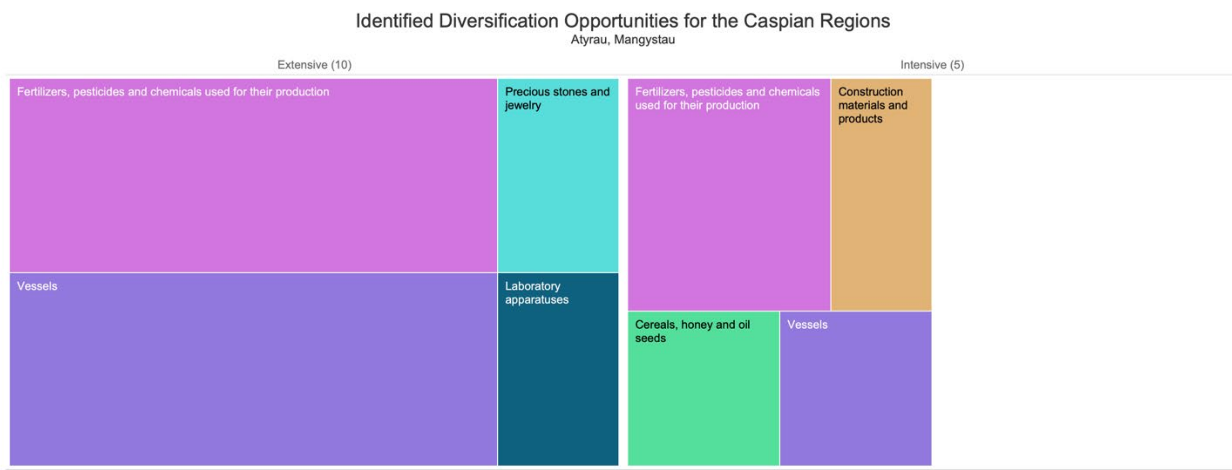


Table 6.7

Identified Diversification Opportunities: Extensive Margin Products

Product	Theme	Broad Category	Feasibility	Attractiveness	Overall Score
Special function vessels, n.e.c. (8905)	Vessels	Transportation	2.03	0.29	2.32
Cargo ships and similar vessels (8901)	Vessels	Transportation	1.20	1.03	2.23
Other vessels (8906)	Vessels	Transportation	0.11	1.67	1.78
Nitrogenous fertilizers (3102)	Fertilizers, pesticides and chemicals used for their production	Chemicals and Allied Industries	1.40	-0.03	1.37
Floating structures for scrapping (8908)	Vessels	Transportation	1.07	0.11	1.18
Sulfur, sublimed or precipitated (2802)	Fertilizers, pesticides and chemicals used for their production	Chemicals and Allied Industries	2.53	-1.63	0.90
Diamonds (7102)	Precious stones and jewelry	Precious Stones and Jewelry	0.65	0.12	0.77
Acyclic alcohols (2905)	Fertilizers, pesticides and chemicals used for their production	Chemicals and Allied Industries	0.21	0.31	0.52
Surveying instruments (9015)	Laboratory apparatuses	Professional Equipment	0.62	-0.29	0.33
Ammonia (2814)	Fertilizers, pesticides and chemicals used for their production	Chemicals and Allied Industries	0.81	-0.56	0.25

Table 6.8

Identified Diversification Opportunities: Intensive Margin Products

Product	Theme	Broad Category	Overall Score
Tugs and pusher craft (8904)	Vessels	Transportation	1.04
Barley (1003)	**Cereals, honey and oil seeds**	**Fruits, Vegetables, and Cereals**	**0.46**
Other articles of copper (7419)	Construction materials and products	Construction	0.40
Cyclic hydrocarbons (2902)	**Fertilizers, pesticides and chemicals used for their production**	**Chemicals and Allied Industries**	**0.06**
Sulphur, crude (2503)	Fertilizers, pesticides and chemicals used for their production	Chemicals and Allied Industries	-0.16

Detailed Results for the Northern Regions

Akmola, Kostanay, North Kazakhstan

The most prominent categories for the Northern Regions are transportation and agricultural products. Opportunities for new or nascent products include several construction materials, cars and processed agricultural products. Opportunities to scale up existing products largely revolve around the existing agricultural production but also include machinery and chemical opportunities. Identified opportunities are overall less concentrated on agriculture than the existing export basket could have suggested.

Figure 6.7

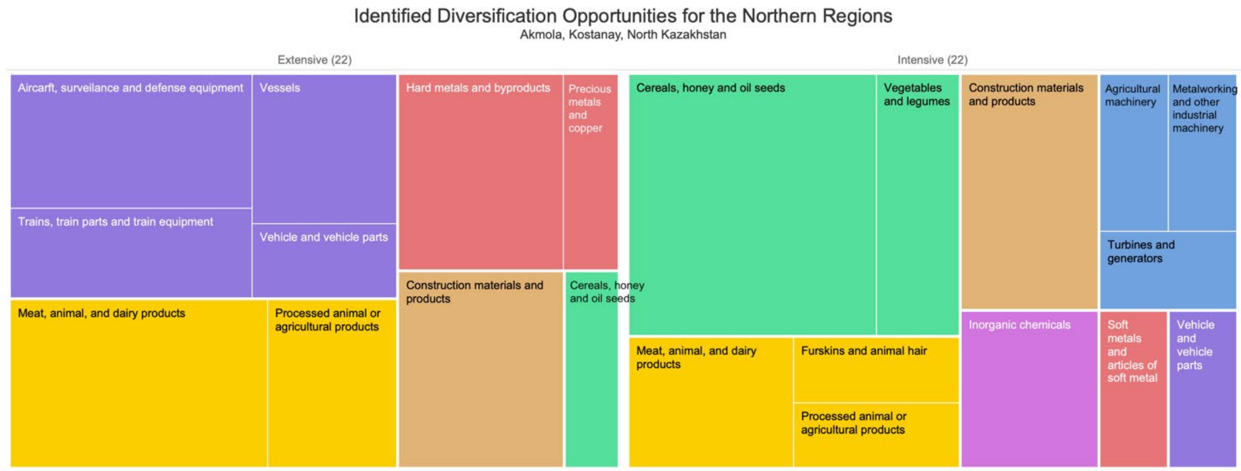


Table 6.9

Identified Diversification Opportunities: Extensive Margin Products

Product	Theme	Broad Category	Feasibility	Attractiveness	Overall Score
Tractors (8701)	Vehicle and vehicle parts	Transportation	0.81	2.41	3.22
Glass fibers (7019)	**Construction materials and products**	**Construction**	**2.09**	**1.01**	**3.10**
Pumps for liquids (8413)	Construction materials and products	Construction	-0.07	3.03	2.96
Nickel mattes (7501)	**Hard metals and byproducts**	**Metals**	**2.70**	**0.26**	**2.96**
Ice cream (2105)	Meat, animal, and dairy products	Meat, Dairy, and Other Processed Agricultural Products	2.63	0.16	2.79
Casein (3501)	**Meat, animal, and dairy products**	**Meat, Dairy, and Other Processed Agricultural Products**	**2.17**	**0.32**	**2.49**
Parts of railway locomotives (8607)	Trains, train parts and train equipment	Transportation	-0.45	2.92	2.47
Pumps, compressors, fans, etc. (8414)	**Construction materials and products**	**Construction**	**-0.67**	**3.04**	**2.37**
Animal feed (2309)	Processed animal or agricultural products	Meat, Dairy, and Other Processed Agricultural Products	1.86	0.48	2.34
Other vessels (8906)	**Vessels**	**Transportation**	**0.52**	**1.75**	**2.28**
Nickel bars, wire etc. (7505)	Hard metals and byproducts	Metals	1.41	0.86	2.27
Munitions of war (9306)	**Aircarft, surveilance and defense equipment**	**Transportation**	**0.07**	**2.16**	**2.22**
Gold (7108)	Precious metals and copper	Metals	0.76	1.41	2.17
Whey (0404)	**Meat, animal, and dairy products**	**Meat, Dairy, and Other Processed Agricultural Products**	**1.23**	**0.88**	**2.10**
Floating structures for scrapping (8908)	Vessels	Transportation	1.97	0.11	2.08
Corn (1005)	**Cereals, honey and oil seeds**	**Fruits, Vegetables, and Cereals**	**2.30**	**-0.22**	**2.08**
Railway track fixtures (8608)	Trains, train parts and train equipment	Transportation	0.58	1.48	2.06
Nickel unwrought (7502)	**Hard metals and byproducts**	**Metals**	**-0.70**	**2.73**	**2.03**
Gas turbines (8411)	Aircarft, surveilance and defense equipment	Transportation	0.01	1.86	1.87
Gelatin (3503)	**Processed animal or agricultural products**	**Meat, Dairy, and Other Processed Agricultural Products**	**0.89**	**0.95**	**1.85**
Other aircraft and spacecraft (8802)	Aircarft, surveilance and defense equipment	Transportation	0.92	0.87	1.79
Horses (0101)	**Meat, animal, and dairy products**	**Meat, Dairy, and Other Processed Agricultural Products**	**1.48**	**0.28**	**1.76**

Table 6.10

Identified Diversification Opportunities: Intensive Margin Products

Product	Theme	Broad Category	Overall Score
Water gas generators (8405)	Turbines and generators	Machinery	0.70
Legumes, dried (0713)	**Vegetables and legumes**	**Fruits, Vegetables, and Cereals**	**0.56**
Uranium (2844)	Inorganic chemicals	Chemicals and Allied Industries	0.54
Ball or roller bearings (8482)	**Vehicle and vehicle parts**	**Transportation**	**0.50**
Edible animal products, n.e.c. (0410)	Meat, animal, and dairy products	Meat, Dairy, and Other Processed Agricultural Products	0.47
Barley (1003)	**Cereals, honey and oil seeds**	**Fruits, Vegetables, and Cereals**	**0.46**
Sunflower seeds (1206)	Cereals, honey and oil seeds	Fruits, Vegetables, and Cereals	0.38
Wheat gluten (1109)	**Cereals, honey and oil seeds**	**Fruits, Vegetables, and Cereals**	**0.31**
Other tanned furskins (4302)	Furskins and animal hair	Meat, Dairy, and Other Processed Agricultural Products	0.29
Butter (0405)	**Meat, animal, and dairy products**	**Meat, Dairy, and Other Processed Agricultural Products**	**0.29**
Harvesting or agricultural machinery (8433)	Agricultural machinery	Machinery	0.28
Bricks, tiles and similar refractory cer (6902)	**Construction materials and products**	**Construction**	**0.25**
Wheat and meslin (1001)	Cereals, honey and oil seeds	Fruits, Vegetables, and Cereals	0.24
Lead refined unwrought (7801)	**Soft metals and articles of soft metal**	**Metals**	**0.23**
Structures and their parts, of iron or s (7308)	Construction materials and products	Construction	0.22
Rape or colza seeds (1205)	**Cereals, honey and oil seeds**	**Fruits, Vegetables, and Cereals**	**0.21**
Worked cereal grains (1104)	Cereals, honey and oil seeds	Fruits, Vegetables, and Cereals	0.16
Other sugars (1702)	**Processed animal or agricultural products**	**Meat, Dairy, and Other Processed Agricultural Products**	**0.11**
Machinery for working minerals (8474)	Metalworking and other industrial machinery	Machinery	0.11
Electrical insulators of any material (8546)	**Construction materials and products**	**Construction**	**0.10**
Aluminum oxide (2818)	Inorganic chemicals	Chemicals and Allied Industries	0.09
Potatoes (0701)	**Vegetables and legumes**	**Fruits, Vegetables, and Cereals**	**0.08**

Detailed Results for the Southern Regions

Almaty, Jambyl, Kyzylorda, Shymkent, Turkestan

The most prominent categories for the Southern Regions are raw and agricultural products, as well as chemicals. Opportunities for new or nascent products include several agricultural products but also a few different industrial products under different themes. Opportunities to scale up also include both industrial and agricultural opportunities, with a notable presence of chemical products. Overall, the set of opportunities identified is remarkably diverse.

Figure 6.8

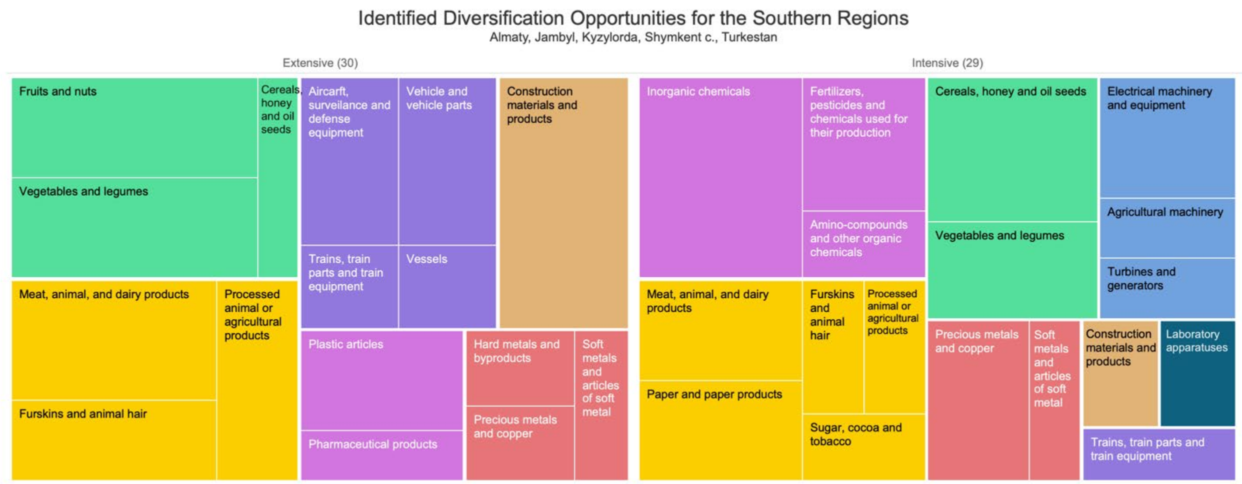


Table 6.11

Identified Diversification Opportunities: Extensive Margin Products

Product	Theme	Broad Category	Feasibility	Attractiveness	Overall Score
Solid soybean residues (2304)	Vegetables and legumes	Fruits, Vegetables, and Cereals	4.40	0.18	4.59
Citrus fruit (0805)	Fruits and nuts	Fruits, Vegetables, and Cereals	3.03	0.60	3.63
Ball or roller bearings (8482)	Vehicle and vehicle parts	Transportation	1.57	1.98	3.55
Articles of cement, of concrete or of ar (6810)	Construction materials and products	Construction	2.52	0.93	3.44
Whey (0404)	Meat, animal, and dairy products	Meat, Dairy, and Other Processed Agricultural Products	2.48	0.90	3.38
Other raw furskins (4301)	Furskins and animal hair	Meat, Dairy, and Other Processed Agricultural Products	3.02	0.15	3.17
Refined copper and copper alloys (7403)	Precious metals and copper	Metals	2.90	0.25	3.15
Fruits, dried (0813)	Fruits and nuts	Fruits, Vegetables, and Cereals	4.18	-1.14	3.05
Animal hair (5102)	Furskins and animal hair	Meat, Dairy, and Other Processed Agricultural Products	3.72	-0.85	2.88
Other plastic plates, sheets etc. (3921)	Plastic articles	Chemicals and Allied Industries	0.85	2.02	2.87
Cheese (0406)	Meat, animal, and dairy products	Meat, Dairy, and Other Processed Agricultural Products	2.52	0.33	2.85
Legumes, dried (0713)	Vegetables and legumes	Fruits, Vegetables, and Cereals	2.40	0.39	2.79
Mineral wools and insulating materials (6806)	Construction materials and products	Construction	0.56	2.22	2.78
Munitions of war (9306)	Aircarft, surveilance and defense equipment	Transportation	0.62	2.12	2.74
Special purpose motor vehicles (8705)	Vehicle and vehicle parts	Transportation	2.64	-0.20	2.44
Fiberboard of wood (4411)	Construction materials and products	Construction	2.35	0.09	2.44
Tanks and other armored fighting vehicle (8710)	Aircarft, surveilance and defense equipment	Transportation	0.95	1.44	2.39
Plastic tubes and fittings (3917)	Plastic articles	Chemicals and Allied Industries	1.68	0.56	2.24
Chocolates (1806)	Processed animal or agricultural products	Meat, Dairy, and Other Processed Agricultural Products	1.63	0.57	2.20
Peptones (3504)	Meat, animal, and dairy products	Meat, Dairy, and Other Processed Agricultural Products	1.45	0.71	2.16
Rape or colza seeds (1205)	Cereals, honey and oil seeds	Fruits, Vegetables, and Cereals	2.07	0.09	2.16
Nickel unwrought (7502)	Hard metals and byproducts	Metals	-0.35	2.51	2.16
Plastic builders' ware (3925)	Construction materials and products	Construction	0.79	1.29	2.08
Self-propelled railway coaches (8603)	Trains, train parts and train equipment	Transportation	0.35	1.68	2.03
Unwrought aluminum (7601)	Soft metals and articles of soft metal	Metals	1.63	0.40	2.03
Vegetables, dried (0712)	Vegetables and legumes	Fruits, Vegetables, and Cereals	1.96	0.02	1.98
Animal feed (2309)	Processed animal or agricultural products	Meat, Dairy, and Other Processed Agricultural Products	1.57	0.39	1.96
Wadding, gauze and bandages (3005)	Pharmaceutical products	Chemicals and Allied Industries	1.64	0.32	1.96
Grapes (0806)	Fruits and nuts	Fruits, Vegetables, and Cereals	2.64	-0.69	1.94
Floating structures for scrapping (8908)	Vessels	Transportation	1.84	0.02	1.86

Table 6.12

Identified Diversification Opportunities: Intensive Margin Products

Product	Theme	Broad Category	Overall Score
Glycosides (2938)	Amino-compounds and other organic chemicals	Chemicals and Allied Industries	1.48
Railway coaches, not self-propelled (8605)	Trains, train parts and train equipment	Transportation	1.33
Copper mattes (7401)	Precious metals and copper	Metals	0.76
Unmanufactured tobacco (2401)	Sugar, cocoa and tobacco	Meat, Dairy, and Other Processed Agricultural Products	0.72
Water gas generators (8405)	Turbines and generators	Machinery	0.70
Horses (0101)	Meat, animal, and dairy products	Meat, Dairy, and Other Processed Agricultural Products	0.70
Modeling pastes (3407)	Construction materials and products	Construction	0.68
Honey (0409)	Cereals, honey and oil seeds	Fruits, Vegetables, and Cereals	0.64
Unrefined copper (7402)	Precious metals and copper	Metals	0.62
Machinery for preparing tobacco (8478)	Agricultural machinery	Machinery	0.61
Batteries (8507)	Electrical machinery and equipment	Machinery	0.60
Pulps of recovered paper fibers (4706)	Paper and paper products	Meat, Dairy, and Other Processed Agricultural Products	0.56
Uranium (2844)	Inorganic chemicals	Chemicals and Allied Industries	0.54
Salts of oxometallic acids (2841)	Inorganic chemicals	Chemicals and Allied Industries	0.52
Other paper cut to size (4823)	Paper and paper products	Meat, Dairy, and Other Processed Agricultural Products	0.47
Barley (1003)	Cereals, honey and oil seeds	Fruits, Vegetables, and Cereals	0.46
Electrical transformers (8504)	Electrical machinery and equipment	Machinery	0.46
Ferroalloys (7202)	Soft metals and articles of soft metal	Metals	0.45
Other manufactured tobacco (2403)	Processed animal or agricultural products	Meat, Dairy, and Other Processed Agricultural Products	0.43
Sunflower seeds (1206)	Cereals, honey and oil seeds	Fruits, Vegetables, and Cereals	0.38
Phosphoric acid etc. (2809)	Fertilizers, pesticides and chemicals used for their production	Chemicals and Allied Industries	0.35
Cyanides (2837)	Inorganic chemicals	Chemicals and Allied Industries	0.34
Surveying instruments (9015)	Laboratory apparatuses	Professional Equipment	0.34
Vegetable products n.e.c. (1404)	Vegetables and legumes	Fruits, Vegetables, and Cereals	0.33
Silicon & rare gases (2804)	Inorganic chemicals	Chemicals and Allied Industries	0.31
Other tanned furskins (4302)	Furskins and animal hair	Meat, Dairy, and Other Processed Agricultural Products	0.29
Butter (0405)	Meat, animal, and dairy products	Meat, Dairy, and Other Processed Agricultural Products	0.29
Mixed fertilizers (3105)	Fertilizers, pesticides and chemicals used for their production	Chemicals and Allied Industries	0.29
Onions, shallots, garlic (0703)	Vegetables and legumes	Fruits, Vegetables, and Cereals	0.27

Detailed Results for Almaty City

Opportunities identified for Almaty City are well-diversified and include several high complexity products. Top opportunities for new or nascent products include pharma, measurement equipment and a range of construction materials. Opportunities to scale up are equally diversified and top opportunities include aeronautics, chemicals, machinery, and equipment. Overall, the set of opportunities identified seem to reflect the set of productive capabilities of a diversified economic capital. Outside of the identified products, opportunities may exist in high-value-added services or creative industries.

Figure 6.9

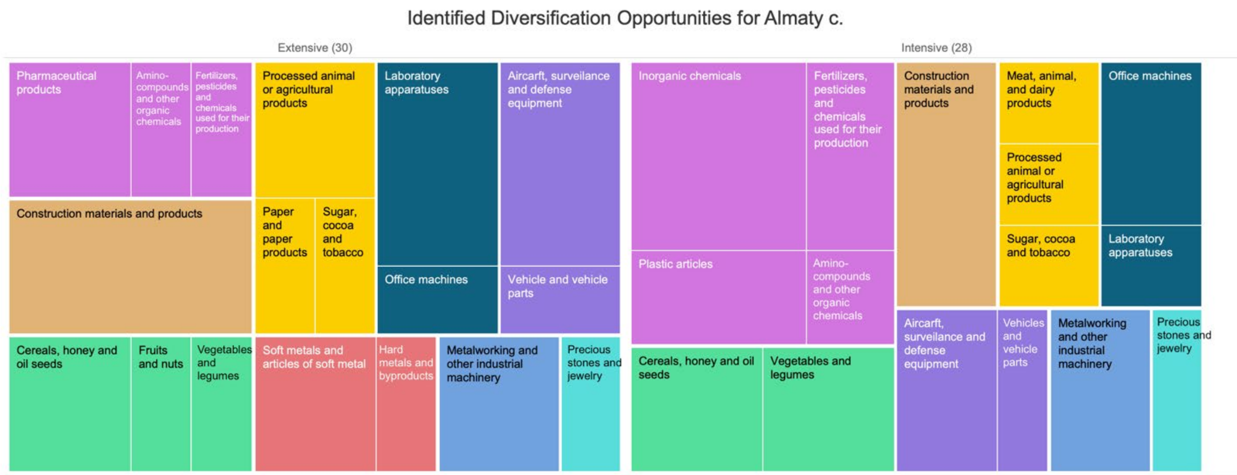


Table 6.13

Identified Diversification Opportunities: Extensive Margin Products

Product	Theme	Broad Category	Feasibility	Attractiveness	Overall Score
Medicaments, packaged (3004)	Pharmaceutical products	Chemicals and Allied Industries	0.86	4.24	5.10
Thermometers, hydrometers etc. (9025)	Laboratory apparatuses	Professional Equipment	2.21	2.74	4.95
Mineral wools and insulating materials (6806)	Construction materials and products	Construction	1.87	2.34	4.21
Other articles of iron or steel (7326)	Soft metals and articles of soft metal	Metals	0.38	3.81	4.19
Serums and vaccines (3002)	Pharmaceutical products	Chemicals and Allied Industries	-0.50	4.31	3.81
Pumps for liquids (8413)	Construction materials and products	Construction	0.74	2.89	3.63
Plastic builders' ware (3925)	Construction materials and products	Construction	2.31	1.32	3.63
Gas turbines (8411)	Aircarft, surveilance and defense equipment	Transportation	1.71	1.73	3.44
Other breathing appliances and gas masks (9020)	Laboratory apparatuses	Professional Equipment	2.10	1.21	3.30
Yeasts (2102)	Processed animal or agricultural products	Meat, Dairy, and Other Processed Agricultural Products	3.82	-0.54	3.28
Worked cereal grains (1104)	Cereals, honey and oil seeds	Fruits, Vegetables, and Cereals	3.32	-0.05	3.28
Tubes, seamless, of iron or steel (7304)	Soft metals and articles of soft metal	Metals	2.38	0.81	3.19
Solid soybean residues (2304)	Vegetables and legumes	Fruits, Vegetables, and Cereals	2.87	0.27	3.14
Electrical insulators of any material (8546)	Construction materials and products	Construction	2.24	0.83	3.07
Sugarcane & sucrose (1701)	Sugar, cocoa and tobacco	Meat, Dairy, and Other Processed Agricultural Products	2.68	0.36	3.04
Glues and adhesives (3506)	Processed animal or agricultural products	Meat, Dairy, and Other Processed Agricultural Products	2.12	0.87	2.99
Radar (8526)	Aircarft, surveilance and defense equipment	Transportation	1.29	1.68	2.97
Phosphoric acid etc. (2809)	Fertilizers, pesticides and chemicals used for their production	Chemicals and Allied Industries	3.16	-0.27	2.89
Rape or colza seeds (1205)	Cereals, honey and oil seeds	Fruits, Vegetables, and Cereals	2.66	0.20	2.86
Sulfonamides (2935)	Amino-compounds and other organic chemicals	Chemicals and Allied Industries	-0.22	3.06	2.83
Apparatus and equipment for photographic (9010)	Laboratory apparatuses	Professional Equipment	-0.14	2.82	2.68
Parts of other aircraft (8803)	Aircarft, surveilance and defense equipment	Transportation	1.15	1.49	2.64
Precious stones (7103)	Precious stones and jewelry	Precious Stones and Jewelry	2.71	-0.09	2.62
Nickel unwrought (7502)	Hard metals and byproducts	Metals	0.06	2.55	2.61
Machinery for working minerals (8474)	Metalworking and other industrial machinery	Machinery	1.86	0.74	2.61
Cars (8703)	Vehicle and vehicle parts	Transportation	-0.57	3.16	2.59
Equipment for temperature change of mate (8419)	Metalworking and other industrial machinery	Machinery	-0.48	3.01	2.54
Other nuts (0802)	Fruits and nuts	Fruits, Vegetables, and Cereals	2.12	0.41	2.53
Newspapers, journals and periodicals (4902)	Paper and paper products	Meat, Dairy, and Other Processed Agricultural Products	1.49	1.03	2.52
Monitors and projectors (8528)	Office machines	Professional Equipment	0.46	2.00	2.46

Table 6.14

Identified Diversification Opportunities: Intensive Margin Products

Product	Theme	Broad Category	Overall Score
Other aircraft and spacecraft (8802)	Aircarft, surveilance and defense equipment	Transportation	1.40
Ion-exchangers based on polymers (3914)	Plastic articles	Chemicals and Allied Industries	1.28
Oxygen-function amino-compounds (2922)	Amino-compounds and other organic chemicals	Chemicals and Allied Industries	1.06
Other office machines (8472)	Office machines	Professional Equipment	0.96
Metal-rolling mills (8455)	Metalworking and other industrial machinery	Machinery	0.84
Navigational instruments (9014)	Aircarft, surveilance and defense equipment	Transportation	0.74
Instruments designed for demonstrational (9023)	Laboratory apparatuses	Professional Equipment	0.73
Auxiliary parts for use with boilers (8404)	Construction materials and products	Construction	0.72
Synthetic monofilament >67 dtex, thickne (5404)	Construction materials and products	Construction	0.69
Legumes, dried (0713)	Vegetables and legumes	Fruits, Vegetables, and Cereals	0.56
Uranium (2844)	Inorganic chemicals	Chemicals and Allied Industries	0.54
Jewelry of precious metal (7113)	Precious stones and jewelry	Precious Stones and Jewelry	0.52
Other engines and motors (8412)	Vehicle and vehicle parts	Transportation	0.52
Cash registers, calculators, etc. (8470)	Office machines	Professional Equipment	0.52
Barley (1003)	Cereals, honey and oil seeds	Fruits, Vegetables, and Cereals	0.46
Other manufactured tobacco (2403)	Processed animal or agricultural products	Meat, Dairy, and Other Processed Agricultural Products	0.43
Cellulose n.e.c. (3912)	Inorganic chemicals	Chemicals and Allied Industries	0.43
Other articles of copper (7419)	Construction materials and products	Construction	0.40
Sunflower seeds (1206)	Cereals, honey and oil seeds	Fruits, Vegetables, and Cereals	0.38
Cocoa powder (1805)	Sugar, cocoa and tobacco	Meat, Dairy, and Other Processed Agricultural Products	0.36
Silicon & rare gases (2804)	Inorganic chemicals	Chemicals and Allied Industries	0.31
Butter (0405)	Meat, animal, and dairy products	Meat, Dairy, and Other Processed Agricultural Products	0.29
Pigments, nonaqueous (3212)	Plastic articles	Chemicals and Allied Industries	0.29
Mixed fertilizers (3105)	Fertilizers, pesticides and chemicals used for their production	Chemicals and Allied Industries	0.29
Prepared pigments (3207)	Inorganic chemicals	Chemicals and Allied Industries	0.29
Insecticides, rodenticides, fungicides, (3808)	Fertilizers, pesticides and chemicals used for their production	Chemicals and Allied Industries	0.29
Onions, shallots, garlic (0703)	Vegetables and legumes	Fruits, Vegetables, and Cereals	0.27
Machinery for making printing components (8442)	Metalworking and other industrial machinery	Machinery	0.27

Detailed Results for Astana

Opportunities identified for Astana span a few categories including transportation, machinery, metals, and agricultural products. Top opportunities for new or nascent products include aeronautics, iron-based construction materials and agricultural machinery. Opportunities to scale up are diversified and top opportunities include lab equipment, trains, or plastic products. Like Almaty City, the set of opportunities identified seem to reflect the diversified existing capabilities, as well as the nearby presence of the northern grain powerhouse. They also point to some successes in industrial policy around trains and railway equipment. Also similar to Almaty City, additional opportunities may exist in tradable services.

Figure 6.10

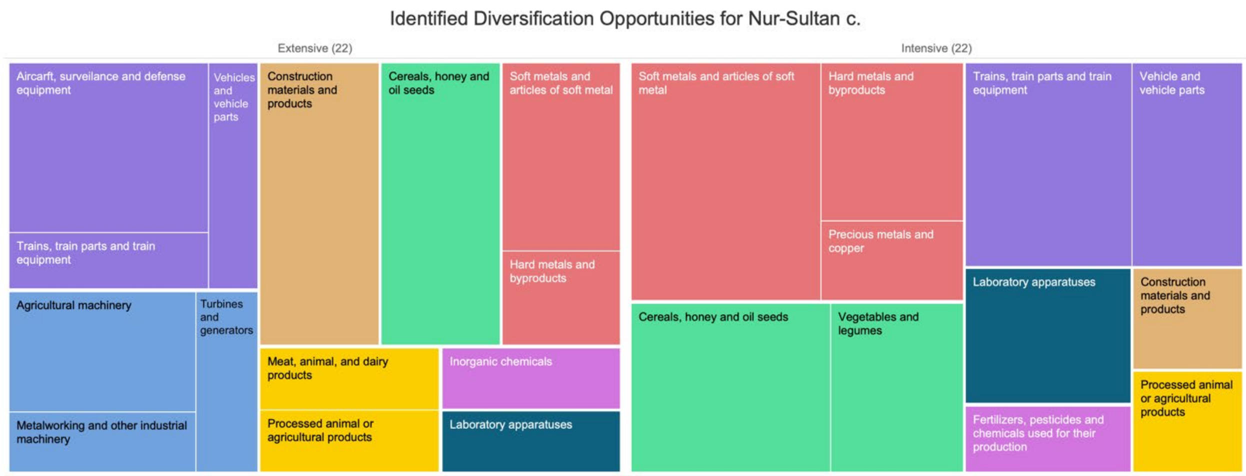


Table 6.15

Identified Diversification Opportunities: Extensive Margin Products

Product	Theme	Broad Category	Feasibility	Attractiveness	Overall Score
Other aircraft and spacecraft (8802)	Aircarft, surveilance and defense equipment	Transportation	4.24	0.96	5.20
Structures and their parts, of iron or s (7308)	Construction materials and products	Construction	1.59	2.70	4.29
Machinery for soil preparation or cultiv (8432)	Agricultural machinery	Machinery	1.15	2.31	3.45
Pumps for liquids (8413)	Construction materials and products	Construction	-0.02	3.34	3.32
Instruments designed for demonstrational (9023)	Laboratory apparatuses	Professional Equipment	2.86	0.43	3.29
Worked cereal grains (1104)	Cereals, honey and oil seeds	Fruits, Vegetables, and Cereals	3.34	-0.05	3.29
Special purpose motor vehicles (8705)	Vehicle and vehicle parts	Transportation	2.98	-0.28	2.71
Gas turbines (8411)	Aircarft, surveilance and defense equipment	Transportation	0.74	1.87	2.60
Nickel unwrought (7502)	Hard metals and byproducts	Metals	-0.27	2.79	2.52
Steam turbines (8406)	Turbines and generators	Machinery	0.23	2.23	2.46
Parts and accessories for metal working (8466)	Metalworking and other industrial machinery	Machinery	-1.25	3.58	2.34
Flat-rolled iron, width < 600mm, not cla (7211)	Soft metals and articles of soft metal	Metals	0.71	1.60	2.32
Other bars and rods of iron or nonalloy (7215)	Soft metals and articles of soft metal	Metals	2.64	-0.40	2.24
Railway track fixtures (8608)	Trains, train parts and train equipment	Transportation	1.11	0.99	2.10
Carbides (2849)	Inorganic chemicals	Chemicals and Allied Industries	1.35	0.68	2.03
Glass fibers (7019)	Construction materials and products	Construction	1.12	0.90	2.03
Horse meat (0205)	Meat, animal, and dairy products	Meat, Dairy, and Other Processed Agricultural Products	2.80	-0.80	1.99
Harvesting or agricultural machinery (8433)	Agricultural machinery	Machinery	-0.45	2.42	1.97
Corn (1005)	Cereals, honey and oil seeds	Fruits, Vegetables, and Cereals	2.20	-0.25	1.96
Tanks and other armored fighting vehicle (8710)	Aircarft, surveilance and defense equipment	Transportation	0.50	1.43	1.93
Wheat gluten (1109)	Cereals, honey and oil seeds	Fruits, Vegetables, and Cereals	2.18	-0.29	1.89
Animal feed (2309)	Processed animal or agricultural products	Meat, Dairy, and Other Processed Agricultural Products	1.39	0.37	1.76

Table 6.16

Identified Diversification Opportunities: Intensive Margin Products

Product	Theme	Broad Category	Overall Score
Other breathing appliances and gas masks (9020)	Laboratory apparatuses	Professional Equipment	1.12
Electric trains (8601)	**Trains, train parts and train equipment**	**Transportation**	**0.78**
Other floating structures (8907)	Vehicle and vehicle parts	Transportation	0.77
Plastic floor coverings (3918)	**Construction materials and products**	**Construction**	**0.70**
Legumes, dried (0713)	Vegetables and legumes	Fruits, Vegetables, and Cereals	0.56
Silver (7106)	**Precious metals and copper**	**Metals**	**0.52**
Unwrought zinc (7901)	Hard metals and byproducts	Metals	0.51
Other tubes and pipes, diameter > 406.4 (7305)	**Soft metals and articles of soft metal**	**Metals**	**0.48**
Ships' derricks; cranes (8426)	Vehicle and vehicle parts	Transportation	0.48
Barley (1003)	**Cereals, honey and oil seeds**	**Fruits, Vegetables, and Cereals**	**0.46**
Ferroalloys (7202)	Soft metals and articles of soft metal	Metals	0.45
Electric signal and traffic controls (8530)	**Trains, train parts and train equipment**	**Transportation**	**0.44**
Sunflower seeds (1206)	Cereals, honey and oil seeds	Fruits, Vegetables, and Cereals	0.38
Parts of railway locomotives (8607)	**Trains, train parts and train equipment**	**Transportation**	**0.35**
Animal or vegetable fats and oils, proce (1518)	Processed animal or agricultural products	Meat, Dairy, and Other Processed Agricultural Products	0.35
Surveying instruments (9015)	**Laboratory apparatuses**	**Professional Equipment**	**0.34**
Tubes, seamless, of iron or steel (7304)	Soft metals and articles of soft metal	Metals	0.33
Insecticides, rodenticides, fungicides, (3808)	**Fertilizers, pesticides and chemicals used for their production**	**Chemicals and Allied Industries**	**0.29**
Onions, shallots, garlic (0703)	Vegetables and legumes	Fruits, Vegetables, and Cereals	0.27
Unwrought aluminum (7601)	**Soft metals and articles of soft metal**	**Metals**	**0.25**
Wheat and meslin (1001)	Cereals, honey and oil seeds	Fruits, Vegetables, and Cereals	0.24
Zinc powders (7903)	**Hard metals and byproducts**	**Metals**	**0.21**

7. Identification of Diversification Opportunities – Complementary Employment-Based Analysis

This section details a complementary, employment-based economic complexity analysis of diversification opportunities for Astana and Almaty City. While the main analysis based on goods export data is appropriate for most regions of Kazakhstan where diversification opportunities are unlikely to come from services, for these two leading cities it is arguably important to conduct analysis which can address diversification opportunities in tradable services.

Objectives and Limitations

The overarching goal of this section is to provide an analysis of diversification opportunities in tradable sectors of employment, including but not limited to service sectors, for Astana and Almaty City. This approach is complementary to the exports approach outlined above. To draw relevant policy conclusions, it is important to understand the objectives and limitations of both approaches, the complementarity of the goods-based and industry-based data, and the ways to make sense of the two sets of results.

The employment-based analysis comes with several important data-related limitations which collectively indicate that employment-based industry recommendations should be manually vetted before implementation. First, the underlying data on employment in Kazakhstan, sourced from the firm registry by Kazakhstan's Bureau of National Statistics, is necessarily approximate. In the raw data the employment of a given area in a given industry is binned with many possible values. We take the harmonic mean of each band to estimate the true employment count, but this may prove an under- or over-estimate. Second, we base recommendations for Astana and Almaty City on diversification patterns among different U.S. Commuting Zones (CZ), which are conceptually similar to the metro areas of cities but cover both urban and rural locations in the United States. High-quality data on how many people are employed in a given industry in a given location is generally unavailable from other countries, and thus using the US as a benchmark is crucial. However, there may be certain inputs to production that are available in the US that are inaccessible in Kazakhstan, so any recommendations must be carefully considered. Third, because industry data in the US and Kazakhstan are recorded in different classification systems, they need to be concorded, which results in a somewhat aggregated system of industries that is ultimately used for recommendations.

With these caveats in mind, the results for employment diversification opportunities are nevertheless arguably valuable. Despite the data limitations, the out-of-sample prediction accuracy of the algorithm's recommendations is good[37]. These recommendations should thus be taken as a useful starting point, subject to more detailed industry-specific exploration.

Industry Identification Framework

Outlining the Framework

As in the main analysis of exports, this complementary analysis of employment analyzes diversification opportunities on both the extensive and intensive margins and scores opportunities in terms of feasibility and attractiveness factors. However, due to the different data sources used, these factors are slightly different. Notably, they do not include measures of industry complexity analogous to PCI or COG in the export-based analysis; this is because subnational measures of industry complexity with

[37] After training on US data, the out-of-sample F1 score for Astana and Almaty City is 64%. This is substantially higher than the average out-of-sample prediction accuracy for a country's exports (as detailed in Appendix C), likely because cities have high concentrations of productive capabilities with predictable structures.

employment tend not to work well in practice[38]. The filtering procedure is also slightly different, in that we specifically consider industries that are considered tradable by Delgado, Porter and Stern (2014). The same filter that excludes opportunities of far distance (low density) in the product-based analysis is also applied to this industry analysis. The processes and variables used are visualized below in Fig. 7.1 and 7.2.

Figure 7.1

Identification framework for extensive margin industries

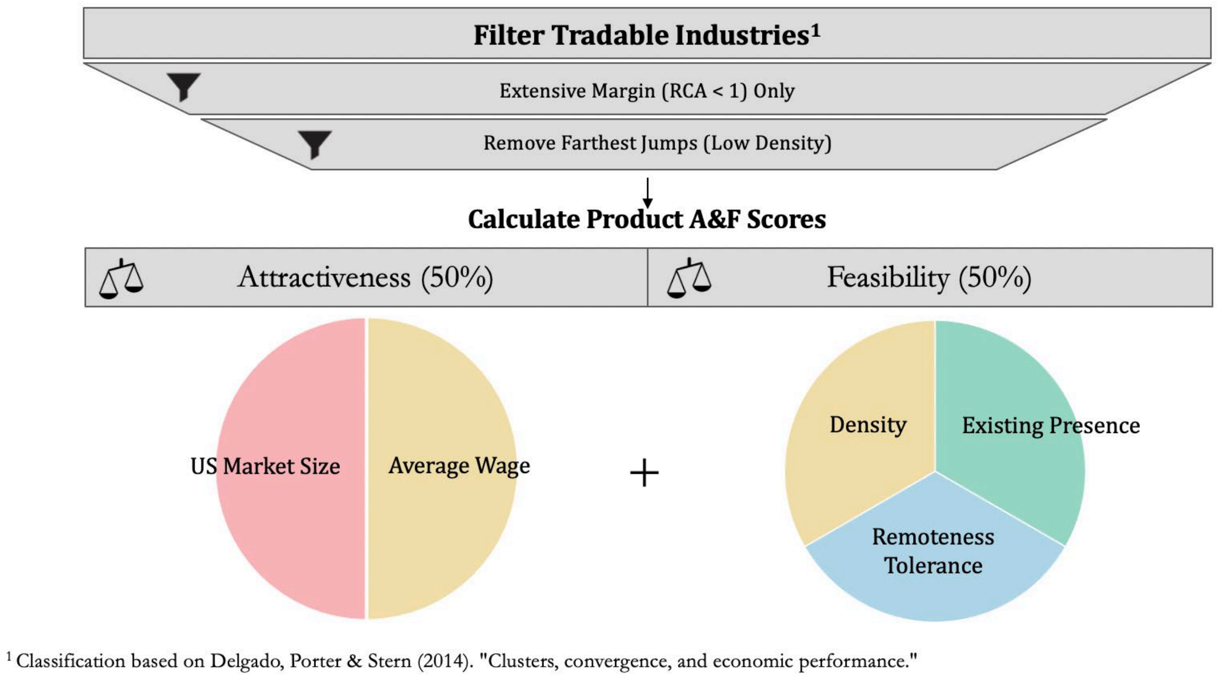
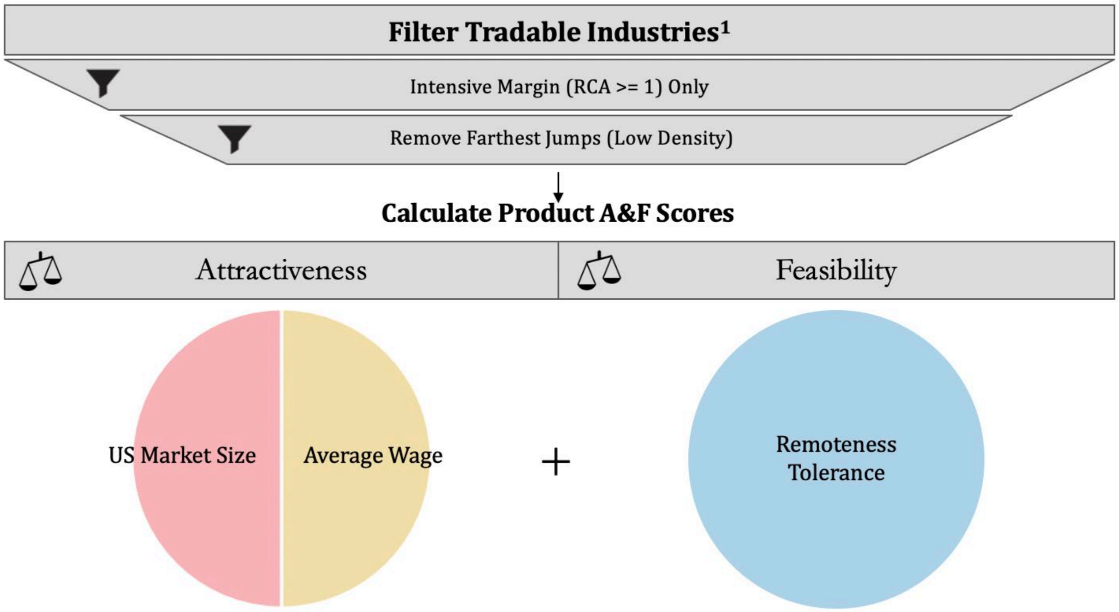


Figure 7.2

Identification framework for intensive margin industries

[38] While it is still possible to make good predictions about opportunities, complexity metrics tend not to work well with subnational data for several reasons. One is that there can be a lack of sufficiently different places in subnational data; whereas countries differ dramatically in their productive capabilities, for example, areas within the US are comparatively more uniform. In addition, subnational data can be affected by the extremely small size of some places. Certain US Commuting Zones only participate in a handful of industries and are heavily affected by idiosyncrasies in production patterns. These irregularities can skew results for complexity metrics.



[1] Classification based on Delgado, Porter & Stern (2014). "Clusters, convergence, and economic performance."

Attractiveness and Feasibility Factors

The attractiveness factors include:

- Size of the total US market for the industry measured in terms of total compensation
- Average wages paid in the industry in the US, as a proxy for industry sophistication

The feasibility factors include:

- Density, i.e. the relatedness to existing industries of employment (only for extensive margin industries)
- Existing presence (only for extensive margin industries)
- Remoteness tolerance

Attractiveness Factors

Size of the US Market: opportunities that face more demand are likely to be more lucrative. We proxy for the relative size of possible markets in Kazakhstan by looking at how large the total market is, in terms of total compensation, in the US (as reported in the Quarterly Census of Employment and Wages). Although this is a very different setting, it is arguably advantageous because the US can be considered on the technological frontier. The relative sizes of its industries thus represent relative sizes at each industry's full potential.

Average Wages: industries that pay higher average wages can provide superior livelihoods for Kazakh citizens. We examine average wages among different industries in the US, as reported in the Quarterly Census of Employment and Wages, to determine how well-paid workers may be.

Feasibility Factors

Density: as in the main analysis of exports, we calculate density here via a machine learning approach. The algorithm determines how proximate a location's productive capabilities are to a certain target industry, by examining the other industries the location participates in.

Existing Presence: as in the main analysis of exports, we use the existing presence of an industry captured by Revealed Comparative Advantage to establish whether a location has an already-existing foothold from which it could build.

Remoteness Tolerance: given that Kazakhstan is a remote country, it should pursue opportunities that are proven to thrive in remote places. However, we do not use the same measure of remoteness as in the main analysis of goods exports; this is because the coverage of each data source is very different. For example, the employment data covers numerous service industries that are wholly absent from export data. Instead, we calculate an index of subnational remoteness for each US Commuting Zone based on driving times to population centers and ports, and then determine the remoteness tolerance of each industry by relating it to the remoteness of the locations in which it is produced. For full details see Appendix C.

Normalizing and Combining Factors

Scores for each industry are combined as follows. First, the natural logarithms of market size and average wages are calculated. This is to express these variables in terms of orders of magnitude, so that extremely large values do not drown out any possible signal from smaller values. Second, each variable is z-normalized. Third, we average across each relevant variable to calculate a feasibility or attractiveness score as needed. Table 7.1 indicates which factors are applicable to which type of recommendation.

Table 7.1

Summary of Factors by Framework

	Factor	**Extensive Margin**	**Intensive Margin**	**Specific to Macro Region**
Attractiveness	Market Size	X	X	
	Average Wage	X	X	
Feasibility	Density	X		X
	Existing Presence	X		X
	Remoteness Tolerance	X	X	

Results

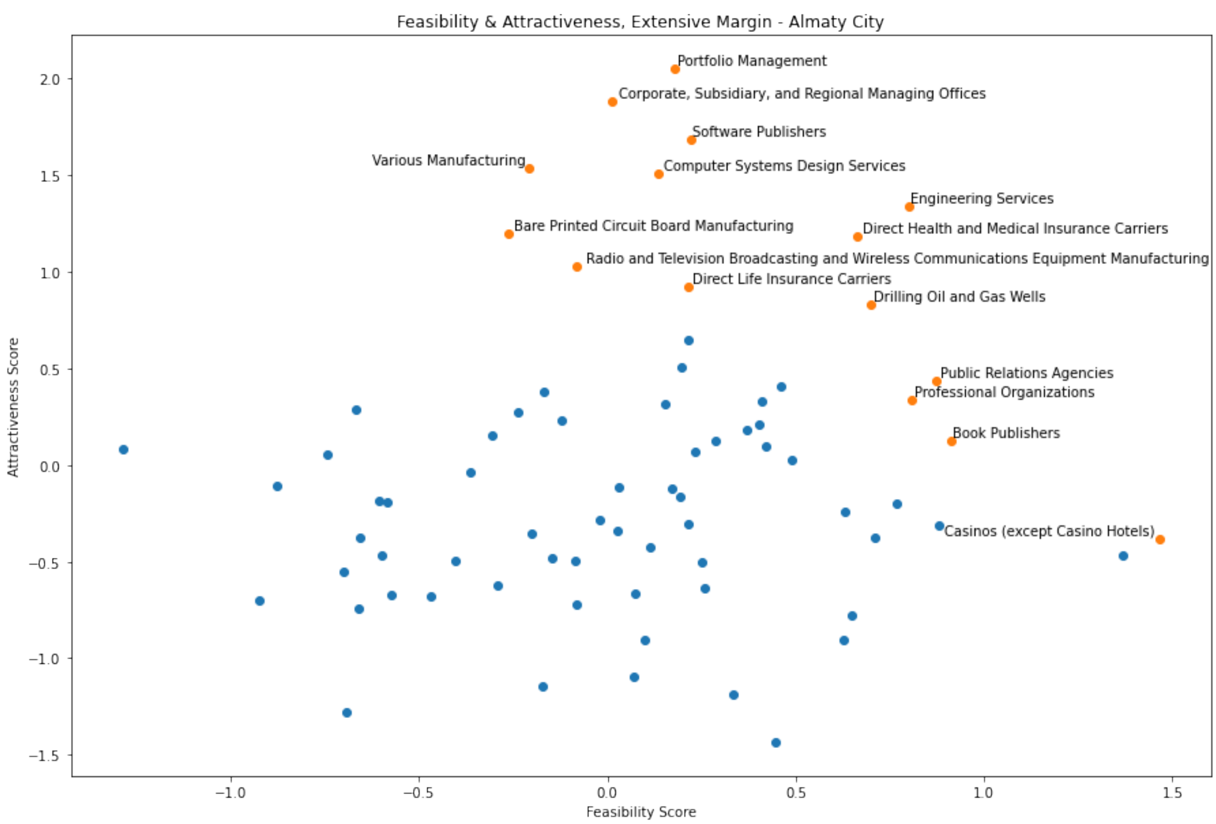
The industries with the highest overall scores are presented below, both on the extensive and intensive margin, for Almaty City and Astana. While in many cases the selected industries represent services or manufacturing, in certain instances there are natural resource industries related to oil or forestry. At a glance this may seem unusual, given that natural resource production typically does not occur in large cities. One explanation for the presence of such industries is the "headquarter" effect whereby Astana or Almaty City may be the headquarter location of companies that produce natural resources in other parts of the country. While we include these industries in the results for transparency and completeness, they demonstrate the need to critically interpret recommended industries; in these instances, the recommended service and manufacturing industries are likely more suitable.

Detailed Results for Almaty City

Results for Almaty City are showcased below. Many top opportunities on both the extensive and intensive margin are in professional services. There are several service opportunities in STEM that may be closely related; for example, Universities and Custom Computer Programming on the intensive margin and Computer Systems Design, Software Publishing, and Engineering Services on the extensive margin. Another collection of potentially related opportunities is comprised of financial services and management, as seen in Savings Institutions, Credit Card Issuing, and Other Legal Services on the intensive margin in addition to Portfolio Management and Managing Offices on the extensive margin. There are additionally some manufacturing industries, such as Plastics Packaging Manufacturing, Paint & Coating Manufacturing, Circuit Board Manufacturing, and Wireless Communications Equipment Manufacturing.[39]

Extensive Margin

Figure 7.3



[39] Various Manufacturing collapses other broader manufacturing codes.

Table 7.2

Identified Diversification Opportunities: Extensive Margin Industries

Industry	Chapter	Feasibility	Attractiveness	Overall Score
Portfolio Management (523920)	Finance and Insurance	0.18	2.05	2.23
Engineering Services (541330)	**Professional, Scientific, and Technical Services**	**0.80**	**1.34**	**2.14**
Software Publishers (511210)	Information	0.22	1.69	1.91
Corporate, Subsidiary, and Regional Managing Offices (551114)	**Management of Companies and Enterprises**	**0.01**	**1.88**	**1.89**
Direct Health and Medical Insurance Carriers (524114)	Finance and Insurance	0.66	1.19	1.85
Computer Systems Design Services (541512)	**Professional, Scientific, and Technical Services**	**0.14**	**1.51**	**1.65**
Drilling Oil and Gas Wells (213111)	Mining, Quarrying, and Oil and Gas Extraction	0.70	0.83	1.53
Various Manufacturing (31-33)	**Manufacturing**	**-0.21**	**1.54**	**1.33**
Public Relations Agencies (541820)	Professional, Scientific, and Technical Services	0.87	0.43	1.31
Professional Organizations (813920)	**Other Services (except Public Administration)**	**0.81**	**0.34**	**1.15**
Direct Life Insurance Carriers (524113)	Finance and Insurance	0.22	0.92	1.14
Casinos (except Casino Hotels) (713210)	**Arts, Entertainment, and Recreation**	**1.46**	**-0.39**	**1.08**
Book Publishers (511130)	Information	0.91	0.12	1.03
Radio and Television Broadcasting and Wireless Communications...	**Manufacturing**	**-0.08**	**1.03**	**0.95**
Bare Printed Circuit Board Manufacturing (334412)	Manufacturing	-0.26	1.20	0.94

Intensive Margin

Figure 7.4

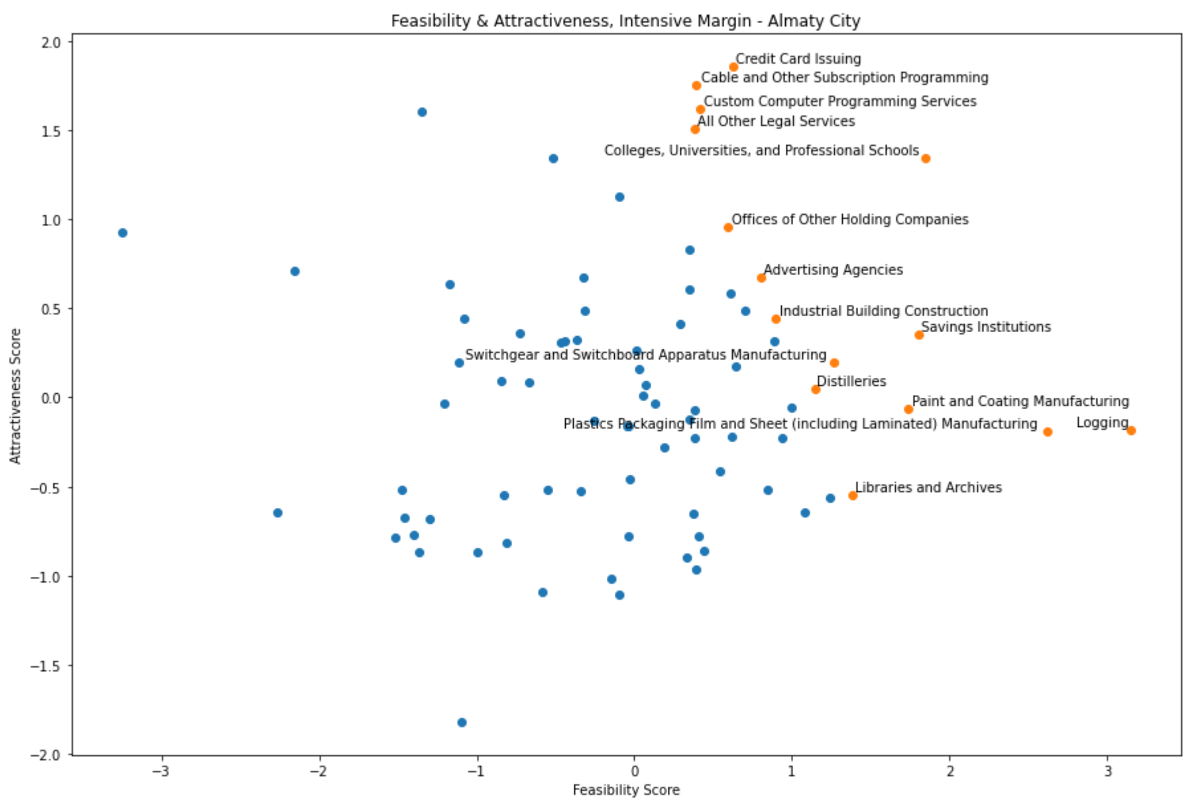


Table 7.3

Identified Diversification Opportunities: Intensive Margin Industries

Industry	Chapter	Overall Score
Logging (113310)	Agriculture, Forestry, Fishing and Hunting	2.03
Plastics Packaging Film and Sheet (including Laminated) Manufacturing (326112)	**Manufacturing**	**1.68**
Colleges, Universities, and Professional Schools (611310)	Educational Services	1.68
Savings Institutions (522120)	**Finance and Insurance**	**1.32**
Paint and Coating Manufacturing (325510)	Manufacturing	1.13
Credit Card Issuing (522210)	**Finance and Insurance**	**1.04**
Switchgear and Switchboard Apparatus Manufacturing (335313)	Manufacturing	0.91
Cable and Other Subscription Programming (515210)	**Information**	**0.85**
Custom Computer Programming Services (541511)	Professional, Scientific, and Technical Services	0.81
Distilleries (312140)	**Manufacturing**	**0.78**
Advertising Agencies (541810)	Professional, Scientific, and Technical Services	0.76
All Other Legal Services (541199)	**Professional, Scientific, and Technical Services**	**0.76**
Industrial Building Construction (236210)	Construction	0.75
Libraries and Archives (519120)	**Information**	**0.74**
Offices of Other Holding Companies (551112)	Management of Companies and Enterprises	0.71

Detailed Results for Astana

Results for Astana show both similarities and differences with those obtained for Almaty City. Similarly, some top opportunities for Astana include STEM and finance & management related opportunities. One notable difference is the comparative sparsity of top opportunities in manufacturing for Astana. That is not to say that Astana ought not to consider manufacturing industries, but simply that the *employment* opportunities in manufacturing for Almaty City could be slightly stronger. That being said, the main analysis of exports offers superior data quality as far as manufactured goods are concerned, and thus should be deferred to in such instances.

Extensive Margin

Figure 7.5

Table 7.4

Identified Diversification Opportunities: Extensive Margin Industries

Industry	Chapter	Feasibility	Attractiveness	Overall Score
Home Furnishing Merchant Wholesalers (423220)	Wholesale Trade	1.07	1.54	2.61
Scheduled Passenger Air Transportation (481111)	Transportation and Warehousing	1.39	0.93	2.32
Portfolio Management (523920)	Finance and Insurance	0.16	2.05	2.21
Computer and Computer Peripheral Equipment and Software…	Wholesale Trade	0.79	1.20	2.00
Engineering Services (541330)	Professional, Scientific, and Technical Services	0.63	1.34	1.97
Computer Systems Design Services (541512)	Professional, Scientific, and Technical Services	0.38	1.51	1.90
Software Publishers (511210)	Information	0.13	1.69	1.82
Corporate, Subsidiary, and Regional Managing Offices (551114)	Management of Companies and Enterprises	-0.08	1.88	1.80
Office Machinery and Equipment Rental and Leasing (532420)	Real Estate and Rental and Leasing	1.99	-0.47	1.52
Hotels (except Casino Hotels) and Motels (721110)	Accommodation and Food Services	1.18	0.13	1.31
Periodical Publishers (511120)	Information	0.85	0.41	1.26
Drilling Oil and Gas Wells (213111)	Mining, Quarrying, and Oil and Gas Extraction	0.41	0.83	1.24
Various Manufacturing (31-33)	Manufacturing	-0.46	1.54	1.08
Fossil Fuel Electric Power Generation (221112)	Utilities	0.15	0.91	1.07
Petroleum Refineries (324110)	Manufacturing	0.17	0.84	1.01

Intensive Margin

Figure 7.6

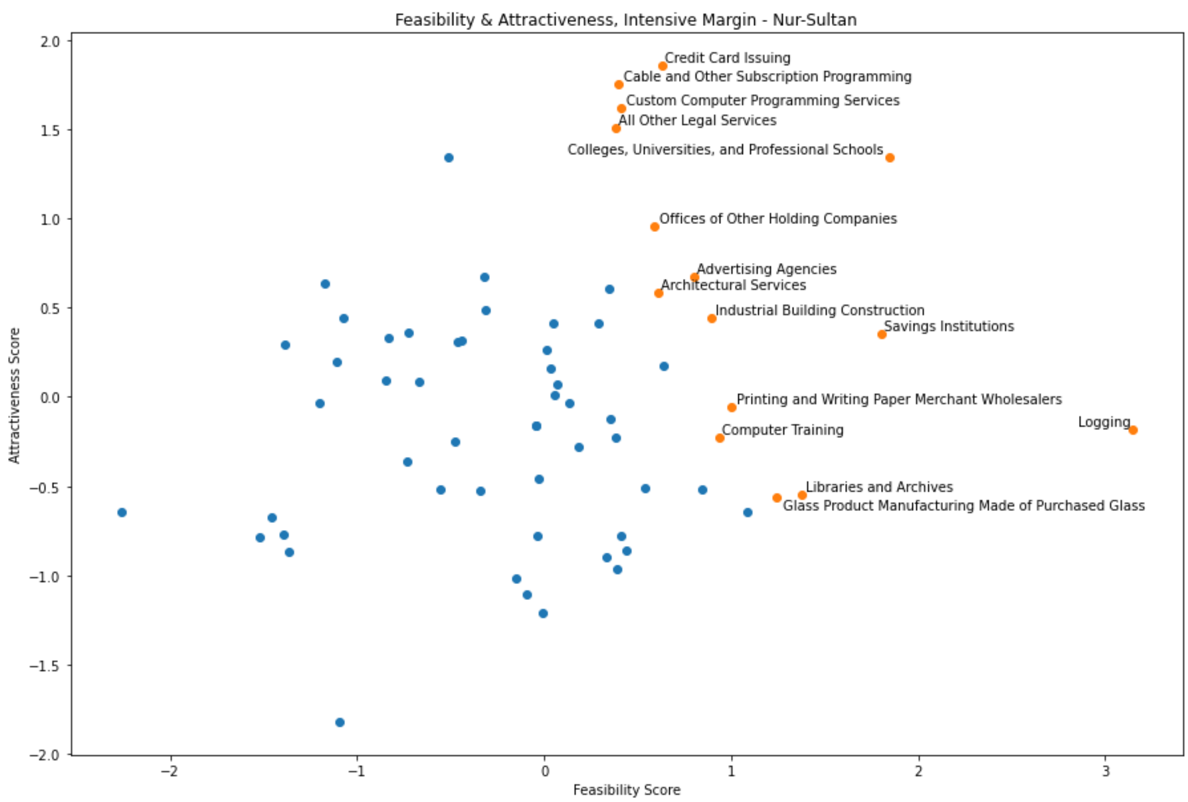


Table 7.5

Identified Diversification Opportunities: Intensive Margin Industries

Industry	Chapter	Overall Score
Logging (113310)	Agriculture, Forestry, Fishing and Hunting	2.03
Colleges, Universities, and Professional Schools (611310)	**Educational Services**	**1.68**
Savings Institutions (522120)	Finance and Insurance	1.32
Credit Card Issuing (522210)	**Finance and Insurance**	**1.04**
Cable and Other Subscription Programming (515210)	Information	0.85
Custom Computer Programming Services (541511)	**Professional, Scientific, and Technical Services**	**0.81**
Advertising Agencies (541810)	Professional, Scientific, and Technical Services	0.76
All Other Legal Services (541199)	**Professional, Scientific, and Technical Services**	**0.76**
Industrial Building Construction (236210)	Construction	0.75
Libraries and Archives (519120)	**Information**	**0.74**
Offices of Other Holding Companies (551112)	Management of Companies and Enterprises	0.71
Printing and Writing Paper Merchant Wholesalers (424110)	**Wholesale Trade**	**0.65**
Glass Product Manufacturing Made of Purchased Glass (327215)	Manufacturing	0.64
Architectural Services (541310)	**Professional, Scientific, and Technical Services**	**0.60**
Computer Training (611420)	Educational Services	0.55

8. Next Steps: Actioning on Identified Opportunities

The pathways to take action on the opportunities identified here are varied and depend on, among other factors, the policy objectives of the institutions responsible for implementation. Those will shape the way the diversification strategy is refined or sequenced. Institutions may reweigh the factors to exclude those that are less relevant for their goals and emphasize those that inform which opportunities best align with their objectives. For example, an institution focused on export promotion may place more weight on a product's Total Addressable Market (TAM). Similarly, policy objectives can inform the way that efforts are sequenced. Institutions focused on rapid job creation may first focus on more labor-intensive industries, while institutions focused on knowledge-based investment attraction may first seek out knowledge-intensive industries whose investors in the region have a high propensity to open local offices and training centers.

Additionally, the process to ultimately implement such efforts will rely on an industry- and region-specific assessment of the binding constraints to industry. Though our methodology to identify opportunities aims to reflect several dimensions of feasibility from a capabilities standpoint, it does not capture other possible feasibility constraints such as challenges in access to finance, availability of necessary inputs, human capital, and infrastructure, navigation of government regulation, means of transporting outputs to target markets, and so on.

We suggest two policy tools to action on suggested opportunities. One targets industries already existing in the country, and the other new or nascent ones. For activities with some existing presence, a specific form of public-private forums called Productivity Taskforces have been effective in identifying and resolving constraints preventing specific industries from being more productive in countries such as Peru, Argentina and Namibia.[40] The taskforces select a focus sector and coordinate the expertise and decision-making power of existing establishments, industry experts, regional and national governments, investment promotion agencies and other actors. These players collaborate to identify economy-wide and industry-specific constraints towards productivity and growth of the sector (coordination problems, inadequate regulations, insufficient public goods, among others), design solutions, and delegate responsibilities towards their implementation. In addition to identifying constraints, these forums offer a space to develop a strategic vision for the sector in the long run, such as targeting new markets or implementing more sustainable practices. The structure of the taskforce varies across country contexts, but in general involves initial brainstorming sessions with existing establishments to discuss their personal challenges to productivity. Then, the taskforce will split into subgroups that will each investigate one challenge and potential solutions that are amenable to policy. The taskforce meets on a regular basis (e.g. monthly) to report on the progress of the subgroups. The taskforce is meant to be an iterative process with an emphasis on experimenting, reviewing, and adjusting to be more effective in alleviating productivity constraints.

For activities without an existing presence in the pertinent regions or country at large, it is not possible to get such feedback from local establishments. Thus, the mechanism to identify constraints must be different. In these cases, targeted investment promotion for a specific sector can help to not only identify suitable foreign investors in the sector but can also enable policymakers to learn what is lacking in Kazakhstan from being a more favorable location for investment in the industry. This entails identifying and approaching previous investors who ended projects in Kazakhstan and prospective

---

[40] For example, see: Martin Obaya and Ernest Stein, "Public-Private Dialogue for the Formulation of Productive Policies: The Experience of Sectoral Roundtables in Argentina," *Inter-American Development Bank*, 2021, https://publications.iadb.org/es/el-dialogo-publico-privado-para-la-formulacion-de-politicas-productivas-la-experiencia-de-las-mesas.

investors who decided to pursue projects elsewhere. The emphasis is on learning from investors the features of contracts and locations that make certain places more attractive investment destinations, and to take stock of which of these things are amenable to policy in Kazakhstan.[41] While certain types of constraints may be limiting to investors anywhere in the country, such as tax incentive structures and regulations on the acquisition of foreign labor, others may be unique to specific regions (such as the quality of required infrastructure or access to affordable transportation).

Conclusion

This *Economic Complexity Report* has explored the productive capabilities of regions of Kazakhstan and identified lists of opportunities for economic diversification. Relying on the economic complexity paradigm, the analysis has also leveraged an array of methodological innovations to better understand productive capabilities, identify the appropriate level of geographic aggregation in which to conduct the analysis, build on barriers to productive growth and diversification discussed in *A Growth Diagnostic of Kazakhstan*, and improve on the predictive power of the product identification process. The report also offers a preliminary analysis of relevant viability and attractiveness factors for the promising industries, which can be leveraged to strategize how to catalyze diversification in each macro region. Among these, it introduces a novel approach to consider the role of remoteness in a diversification strategy. Lastly, it organizes diversification opportunities around internally coherent broad categories and themes that reflect shared required capabilities, and suggest potential pathways for actioning on such opportunities. The *Industry Targeting Dashboard* tool published here allows for flexibility in deciding the factors to be used for product selection and their weights in calculating the final scores. This allows for stakeholders to select the relevant considerations for their needs.

While this report pursues a rigorous identification of products and industries with good prospects for success in Kazakhstan, the actualization of these opportunities will depend on effective implementation by policymakers. The implementation of such a strategy will involve careful considerations of questions such as: who should be responsible for implementation? How can regional governments and national governments use their respective strengths to support industry development? How will coordination be planned across export promotion agencies and investment promotion agencies? Given the bandwidth of institutions implementing the strategy, how can you measure its success in the short and long terms?

[41] For a study of the dynamics between FDI and sectoral development, see: J.H. Shen, H Wang, and S.C. Wang, "Productivity Gap and Inward FDI Spillovers: Theory and Evidence from China," *China & World Economy* 29, no. 2 (n.d.): 24–48.

Appendix A: Peer Selection for the National-Level Analysis

Section 3 utilizes a set of peer countries in order to benchmark Kazakhstan's economic trajectory and set relevant development expectations given the country's historical context, endowment, and related challenges.

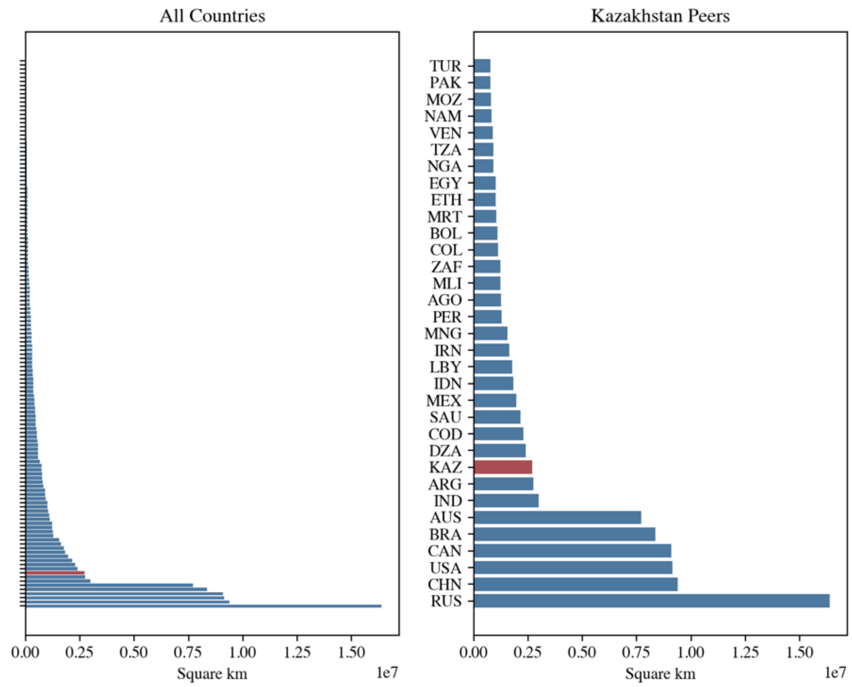
The main dataset used for peer selection was the World Bank's World Development Indicators (WDI). The World Bank WDI contains a wealth of cross-country data on growth, macroeconomic aggregates, demographics, labor, natural resources and more. In order to narrow down prospective comparators from the WDI's list of over 200 countries and territories, we selected countries among the subset of 133 that meet the criteria to also be included in the Economic Complexity Index. These criteria are: (i) a minimum population of 1 million inhabitants in the year selected; (ii) exports over a 3-year period that total at least $1.2 bn; and (iii) defined standards of data quality and reliability. Filtering first by these criteria additionally allows us to compare the evolution of the complexity metrics in these countries with respect to Kazakhstan.

Both time-invariant and performance dimensions can be relevant for peer selection. Comparator countries can be selected based on time-invariant dimensions, including their historical context, geography, or resource endowments. This presents the advantage of narrowing down to peer countries that are structurally relevant for the country being considered. For instance, suggesting Singapore as a comparator to Russia would have little sense, as the countries differ in many structural aspects including size, population, population density, political heritage, and natural resources endowment. Performance dimensions such as income level, recent growth in GDP or in exports, are also relevant to an extent. This is mainly because a country with significantly lower economic and social performance compared to the country of interest is unlikely to yield insights regarding what constraints are binding or the best way forward.

There are geographic features that are essential to consider for Kazakhstan's development process and hence for peer selection. As the ninth largest country by land area and the largest landlocked country in the world, Kazakhstan faces challenges to land connectivity. This pertains to both the separation *between* Kazakhstani cities and other major population centers and to the dispersion of population *within* the country. As such, we use land area and population density as a combined condition to consider countries that are facing similar challenges. We set a threshold for land area above 750,000 square kilometers (33 countries; Kazakhstan = 2,699,700) and for population density below 30 people per square kilometer (30 countries; Kazakhstan = 7). The countries that most closely match Kazakhstan by this criterion are Libya, Australia, Mongolia, Canada, and Russia (Fig. A1).

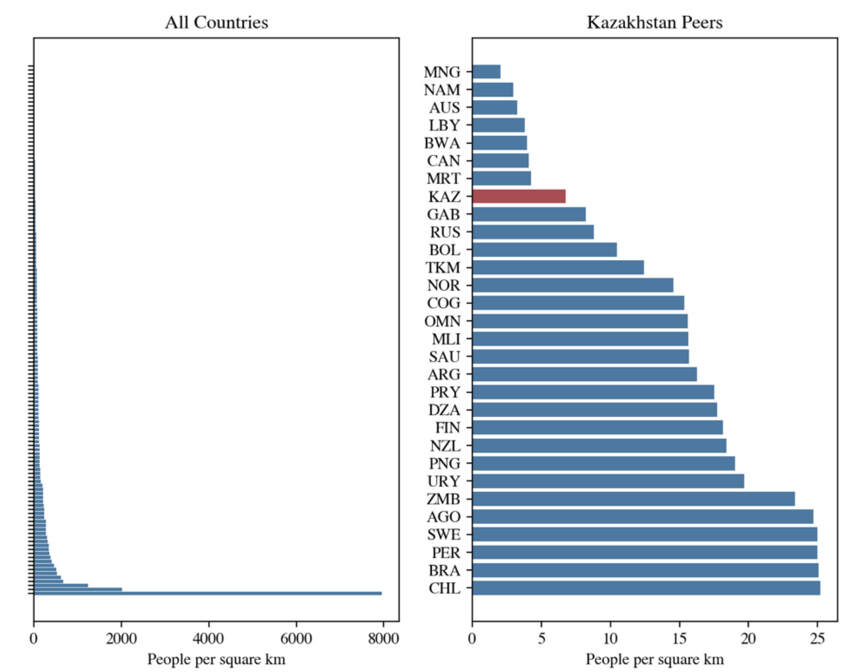
Figure A1

Peer Dimension: Land Area, 2018



Note: Kazakhstan Peers widely defined as countries with population > 1 million and land area > 750,000 square kilometers in 2018.
Source: World Bank WDI

Peer Dimension: Population Density, 2018



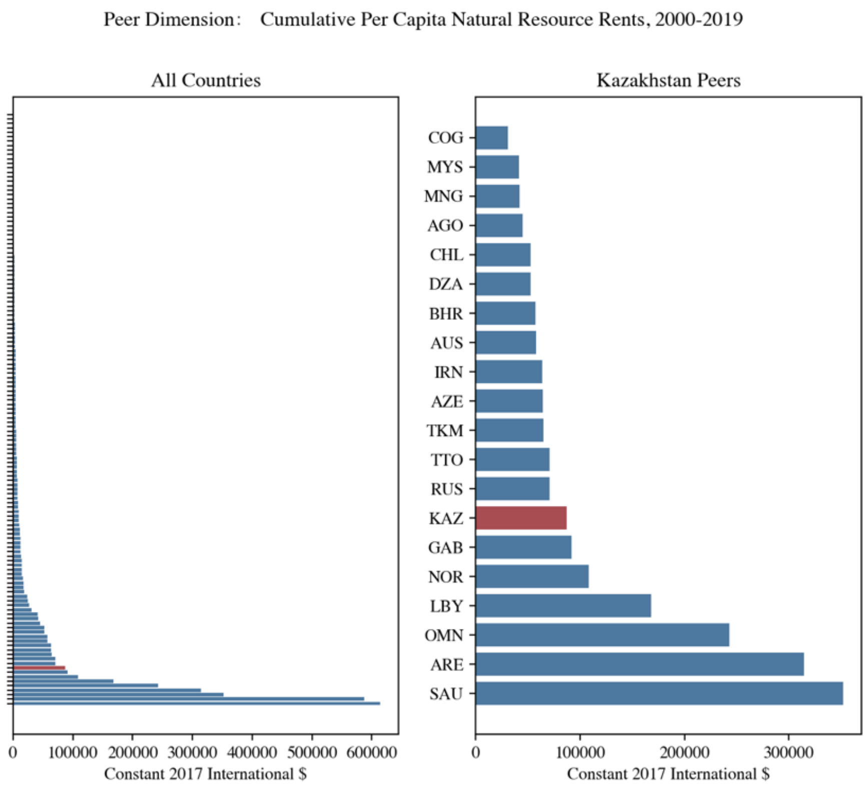
Note: Kazakhstan Peers widely defined as countries with population > 1 million and density < 30 people per square kilometer in 2018.
Source: World Bank WDI

A second criterion we consider as important for Kazakhstan's growth trajectory is the country's history as a former member of the Soviet Union. Several studies establish path dependencies associated with trade, development and innovation in post-Soviet states that can be attributed to the persistence of Soviet institutions and infrastructure after the Soviet Union's dissolution in 1991.[42] To proxy for potential path dependencies that may exist due to formerly belonging to the union, we consider a binary variable that allocates a unit value for the countries Armenia, Azerbaijan, Belarus, Estonia, Georgia, Kyrgyzstan, Kazakhstan, Latvia, Lithuania, Moldova, Russia, Tajikistan, Turkmenistan, Ukraine, and Uzbekistan. In a wider definition of the criterion, we also consider countries in Europe and Central Asia that had communist regimes at one time but were never a part of the Soviet Union (Albania, Bosnia and Herzegovina, Bulgaria, Croatia, Czechia, Hungary, North Macedonia, Poland, Romania, Serbia, Slovakia, and Slovenia).

The third and final criterion we consider is a country's cumulative per capita natural resource rents over the previous two decades. An extensive body of literature — and much of the *Growth Diagnostic of Kazakhstan* — focuses on the implications of high natural resource endowments on fiscal policy, diversification, and growth. Thus, appropriate comparators are those countries that have similar natural resource endowment trajectories as Kazakhstan. We set the threshold for cumulative per capita natural resource rents from 2000-2019 between $30,000-400,000 in constant 2017 international $ (20 countries; Kazakhstan = $87,076). The countries that most closely match Kazakhstan by this criterion are Gabon, Russia, Trinidad & Tobago, Turkmenistan, Azerbaijan and Iran (Fig. A2).

Figure A2

[42] A.B. Krylov, "Post-Soviet States: Challenges of Development."; Assel Mussagulova, "Newly independent, path dependent: The impact of the Soviet past on innovation in post-Soviet states." ; Arman Mazhikeyev & T. Edwards, "Post-colonial trade between Russia and former Soviet republics: back to big brother?"



Note: Kazakhstan Peers widely defined as countries with cum. per capita NR rents (constant 2017 intl $) from 2000-2019 between $30000-400000 (Kazakhstan = $87076).
Source: World Bank WDI

The final peer set was the result of systematic filtering followed by qualitative selection to avoid redundancies. Each of the three criteria are successively filtered with their broad range around Kazakhstan's value, and then the resulting lists are simultaneously compared with contextual considerations to arrive at a narrower set of peers. Priority was given to countries meeting the largest number of narrow criteria, while highly similar countries were streamlined to shorten the final list of peers. For better interpretation, we divided the remaining peers into three subsets: regional, global, and aspirational (OECD) peers. These peers are used for analyses throughout Section 3.[43] The final sets of peers are listed below (Fig. A3 and Fig. A4).

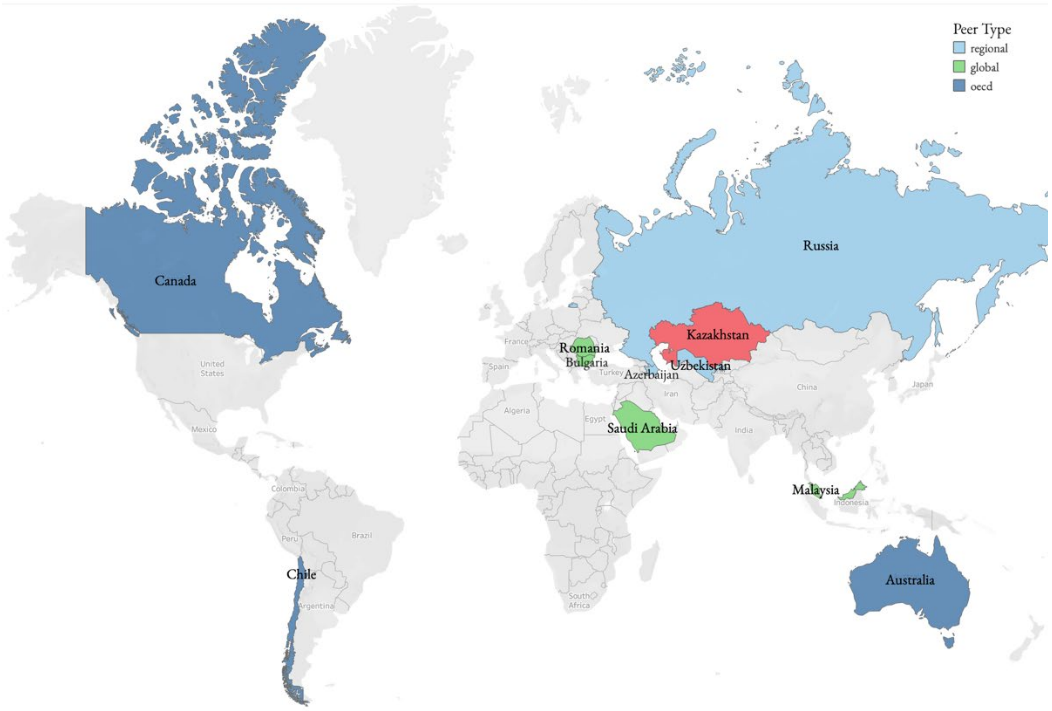
Figure A3

Peer Type	Countries
Regional	Azerbaijan, Russia, Uzbekistan
Global	Bulgaria, Malaysia, Romania, Saudi Arabia
Aspirational (OECD)	Australia, Canada, Chile

[43] While the retained peers are used throughout Section 3, some specific analyses do warrant a different set of comparators. For instance, certain economic diversification results may be best assessed considering oil-rich countries only.

Figure A4

Kazakhstan Country Peers



Appendix B: The Industry Targeting Dashboard

The *Industry Targeting Dashboard* is a companion tool for this report that grants stakeholders additional flexibility in specifying the criteria to identify export opportunities and provides insight into the factor data for over 700 products. The dashboard is built primarily on export data from the Atlas of Economic Complexity and Development Bank of Kazakhstan (KDB), and thus does not include results of the service-based analysis detailed in Section 7. A breakdown of each factor described in this report, including the data and methods to calculate it, is included in the methodology note here.

The dashboard begins with selection of one of the macro regions designated in this report, such as the Industrial Belt. The first section of the tool presents the composite attractiveness and feasibility scores of products for the selected macro region, with the highest-scoring ones ultimately selected for inclusion in this report (Fig. B1). There are separate scatterplots for products that are new to the region ("New and Emerging Products") and those already intensively present ("Intensively Present Products"). On the right-hand side of the charts is a list of the factors and their current weights, defaulted to weigh each equally (as was done to obtain the results of this report). The user can alter the weights of each factor between 0 and 1 based on their priorities and observe how the top-performing products change. The tables below these scatterplots are linked dynamically to this weighting scheme and update with the new attractiveness and feasibility scores (Table B1).

Figure B1

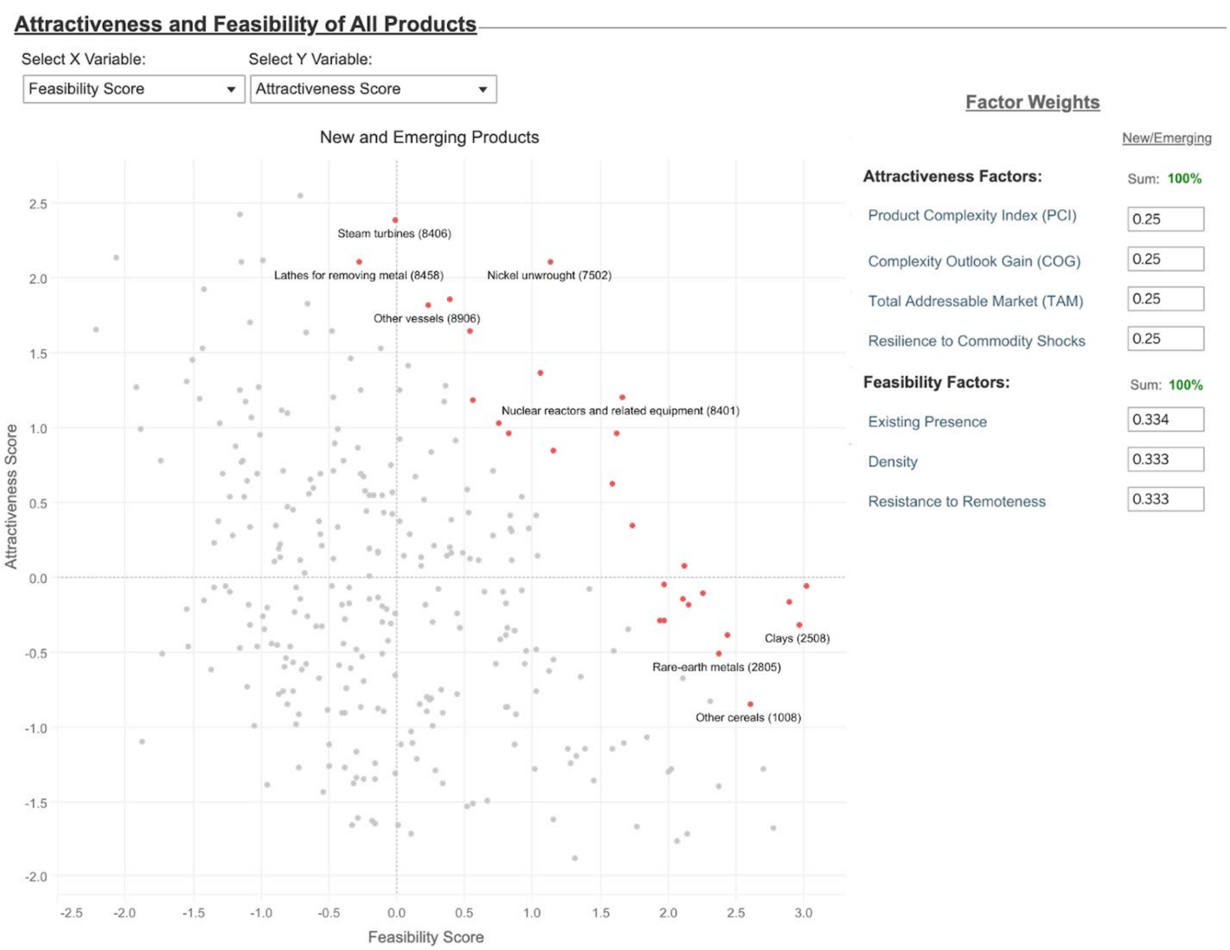


Table B1

Factor Scores: New and Nascent Products

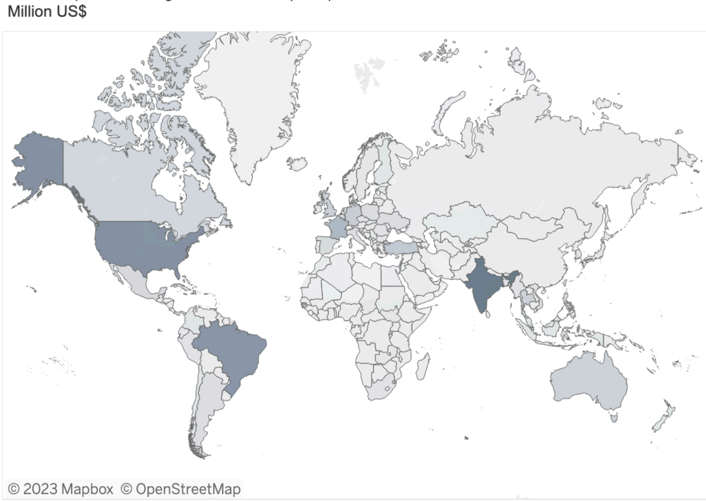
Product	Sector	Rec. in Report	Existing Presence	Density	Resistance to Remoteness	PCI	COG	TAM	Resilience to Commodity Shocks	Attractiveness Score	Feasibility Score	Overall Score
Nickel unwrought (7502)	Metals	1	-0.485	1.896	0.498	0.330	1.311	1.082	1.347	2.101	1.133	3.234
Other tubes and pipes, diameter > 406.4 mm, of iron or steel (7305)	Metals	1	2.000	3.107	-0.023	0.125	-1.308	0.672	0.393	-0.061	3.024	2.963
Nuclear reactors and related equipment (8401)	Machinery	1	-0.437	1.964	1.272	1.981	-0.576	0.080	0.828	1.194	1.663	2.856
Railway cars, not self-propelled (8606)	Vehicles	1	1.430	3.651	-0.222	0.119	-0.740	-0.585	0.877	-0.170	2.890	2.720
Clays (2508)	Minerals	1	3.138	2.018	-0.174	-0.501	1.018	-0.383	-0.766	-0.326	2.965	2.639
Binoculars and telescopes (9005)	Machinery	1	-0.481	2.069	1.135	1.733	-0.352	-0.545	1.019	0.958	1.617	2.575
Tanks and other armored fighting vehicles (8710)	Vehicles	1	-0.485	0.199	2.071	0.760	-1.127	0.399	2.602	1.359	1.060	2.419
Steam turbines (8406)	Machinery	1	-0.456	-0.593	1.034	1.567	0.871	0.393	1.775	2.378	-0.010	2.368
Gas turbines (8411)	Machinery	1	-0.150	-0.661	1.462	1.449	-0.186	2.707	-0.384	1.851	0.387	2.237
Railway coaches, not self-propelled (8605)	Vehicles	1	-0.485	1.325	1.837	0.673	-1.219	-0.155	1.910	0.624	1.590	2.214
Floating structures for scrapping (8908)	Vehicles	1	-0.485	1.070	2.992	-0.490	-0.909	-0.182	1.721	0.072	2.124	2.197
Self-propelled railway coaches (8603)	Vehicles	1	-0.485	-0.088	1.483	2.254	-1.005	0.437	1.492	1.641	0.540	2.181

The second section of the tool allows the user to search for a specific product to view its factor data and additional insights on global trade in the product. Heatmaps of the global imports, addressable market, and current exports from Kazakhstan detail the trade dynamics for the product for each year from 2012-2019 (2012 being the earliest year for which KDB data was available). Fig. B2 and B3 illustrates the example of Nitrogenous fertilizers, an emerging product in the Industrial Belt, detailing the product's factor scores, largest global importers, accessible markets, and destination of Kazakhstan's exports in the product.

Figure B2

Figure B3.

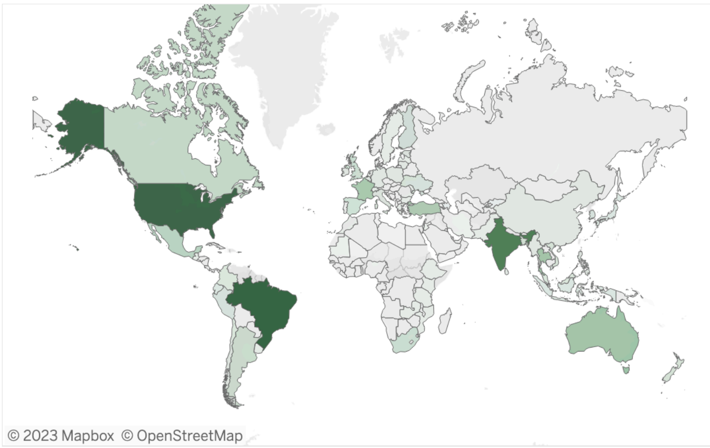
Global Imports of Nitrogenous fertilizers (3102), 2019
Million US$



Global Imports of Nitrogenous fertilizers (3102), 2019 (Million US$)

Country	
India	2,642
United States	2,067
Brazil	1,938
France	1,108
Turkey	749
Germany	639
Australia	600
United Kingdom	554
Thailand	529
Canada	472
Mexico	429
Belgium	407
Spain	388
Argentina	332
Ukraine	323
Italy	319
Poland	303

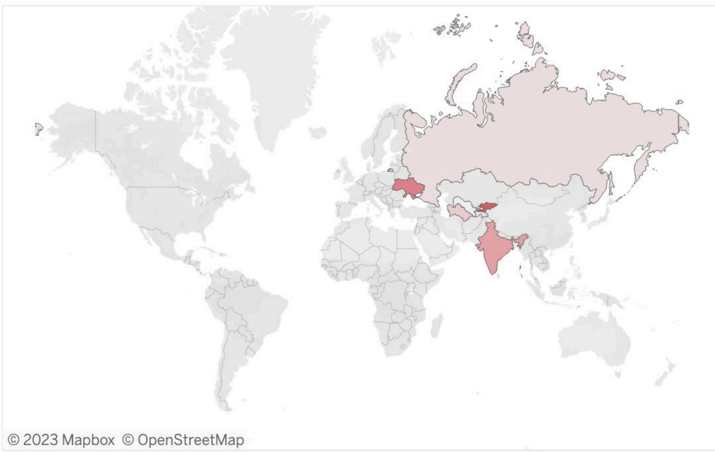
Addressable Markets for Nitrogenous fertilizers (3102), 2019
Million US$



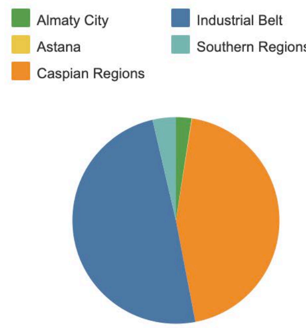
Addressable Markets for Nitrogenous fertilizers (3102), 2019 (Million US$)

Country	
Brazil	1,561
United States	1,523
India	1,258
Australia	481
Thailand	478
France	427
Turkey	401
Mexico	304
Canada	249
Argentina	206
United Kingdom	196
South Africa	193
Bangladesh	186
Finland	164
Spain	156
New Zealand	148

Kazakhstan's Exports of Nitrogenous fertilizers (3102), 2019
Million US$



Kazakhstan's Exports of Nitrogenous fertilizers (3102), 2019



Note: Export data is compiled from the Atlas of Economic Complexity (UN Comtrade) and Kazakhstan Development Bank. Differences in recording and rounding may lead to inconsistencies between regional and national totals across the datasets.

All visualizations and the tables listing the updated factor scores are available to download by clicking the Download tab on the top right of the tool.

Appendix C: Technical Appendix – Complexity Calculations and Industry Prioritization

Mathematical Definitions of Complexity Metrics

Revealed Comparative Advantage is the ratio of a country's exports that come from a particular product, divided by the share of total world exports that come from that product:

$$RCA_{c,p} = \frac{X_{c,p} / \sum_c X_{c,p}}{\sum_p X_{c,p} \,/\, \sum_{c,p} X_{c,p}}$$

An MCP refers to a situation where, in the Matrix of Countries and Products, RCA is greater than or equal to one.

Diversity is equal to the sum of a country's MCPs across all its products:

$$Diversity_c = \sum_p MCP_{c,p}$$

Ubiquity is equal to the sum of a product's MCPs across all countries:

$$Ubiquity = \sum_c MCP_{c,p}$$

The Economic Complexity Index (ECI) and Product Complexity Index (PCI) are the eigenvectors corresponding to the second largest eigenvalues of the matrix of MCPs, for which rows are for countries and columns are for products.

Density is conventionally defined as:

$$Density_{c,p=j} = \frac{\sum_{p \neq j} (1 - M_{c,p}) Proximity_{p,j}}{\sum_{p \neq j} Proximity_{p,j}}$$

Where Proximity is the minimum of the conditional probabilities of co-incidence in the MCP matrix for two products. I.e., if the probability of having an MCP in Product B given having an MCP in Product A is 50% and the vice versa is only 40%, then the proximity between these two products is defined as 40%. Mathematically, this is as follows:

$$Proximity_{i,j} = Minimum\{Prob(MCP_i | MCP_j), Prob(MCP_j | MCP_i)\}$$

Complexity Outlook Gain is conventionally defined as:

$$COG_{c,p=j} = \sum_{p \neq j} \frac{Proximity_{p,j}}{\sum_{p' \neq p} Proximity_{p',p}} (1 - MCP_{c,p})\, PCI_p$$

Random Forest Algorithm for Density

The Random Forest algorithm to calculate density answers the following prediction problem: given knowledge of a particular country's MCPs in every product except product *p*, does that country have an MCP in product *p*? We evaluated its performance against the standard measure of density and two alternative approaches by training each method on data covering the same set of 99 countries, and then comparing the predictive accuracy of each approach for out-of-sample data covering another 34 countries.

We specifically use the F1 score as our metric of out-of-sample predictive accuracy. This is a metric applied to binary outcomes (i.e., 0 or 1) which is suitable for sparse datasets where there are many more 0s than 1s. The F1 score is defined as follows:

$$F1 = \frac{\#\,True\ Positives}{\#\,True\ Positives + \frac{1}{2}(\#False\ Positives + \#False\ Negatives)}$$

There are three major steps in our Random Forest procedure: data sampling, variable selection, and model training. After these steps are complete, we evaluate performance.

We begin by creating a blocked random sample of countries, according to diversity, to divide into our training and validation data. This is to ensure that samples are reasonably representative of the world and are not randomly skewed towards or away from particular kinds of countries.

It is critical to divide the training and validation data by country, and not naïvely by observations at the country-year-product level, to avoid information leakage from the training to validation data. The latter naïve procedure could easily result in a situation where, for example, French exports of wine in 2000 are allocated to the training data and French exports of wine in 2001 are allocated to the validation data. However, year-on-year changes in the export profiles of countries are typically small. As such, nearly all the information from the training dataset would be leaked into the validation dataset. This would result in validation performance results that are artificially and incorrectly high. Instead, strictly choosing entire countries to be either exclusively in the training or the validation dataset prevents this kind of information leakage.

To execute our sampling procedure, we calculate each country's diversity in each year, and then calculate each country's average diversity across all years. We then order countries from highest to lowest average diversity and divide the countries in this ordered list into blocks of four. For example, the four countries at the top of this list will be in Block 1; the next four in Block 2; and so on. Within each block we then randomly select one country for the validation dataset and use the remaining three for training.

We additionally randomly assign the three training countries within each block IDs of 1, 2, or 3 to create "training folds" —representative random subsamples of the training data—for subsequent use in our cross-validated variable selection procedure. The goal of this variable selection procedure is to determine which products are not relevant for predicting any product *p*, so that they can be safely ignored by the algorithm when predicting a particular product *p* to reduce the possibility of overfitting. This cross-validated variable procedure is as follows:

1. Choose a product *p* to make predictions for.
2. Choose one training fold as a validation dataset for cross-validation
3. Combine the other two training folds
4. Fit a Random Forest on the two combined training folds
5. Extract the Random Forest's Feature Importance, scoring how important each product was for making predictions about product *p*
6. Check the Random Forest's prediction accuracy for the cross-validation fold, and record this value
7. Repeat steps 2-6, using the other training folds for cross-validation instead
8. Record the average prediction accuracy for all three possible choices of training folds as cross-validation folds
9. Select the top *n* products in each possible combination of training folds according to the previously extracted Feature Importance Scores
10. Repeat steps 2-8 for every desired *n*
11. Identify the *n* number of products that are associated with the highest average cross-validated prediction accuracy, from step 8
12. Fit a Random Forest on all the training data together
13. Extract the top *n* products according to Feature Importance scores, at the level of *n* determined in step 11. Record these products as the selected products which are relevant for predicting product *p*.
14. Repeat steps 1-13 for a different product *p*.

Having determined which products are relevant for predicting each product *p*, we now fit Random Forests on the full set of training data. We fit the algorithm separately for each product *p*, in each instance only considering the relevant predictor products previously selected. The Random Forest procedure is as follows:

1. Choose a product *p* to make predictions for.
2. Among the training data, up sample the portion of observations for which the MCP for product *p* is 1 so that there are an equal number of observations where MCP = 0 and MCP = 1. This helps to force the algorithm to not simply guess that the MCP should almost always be 0, and instead to rely on the structural information present in the products a country exports.
3. Duplicate the entirety of this up sampled training data three times and append it to the up sampled training data.
4. Separate the output training data, covering product *p*, from the input training data, covering every other relevant product. Among the input data, randomly set 15% of MCP entries to zero. This helps to prevent overfitting by ensuring the algorithm cannot rely on an exact combination of products to make predictions, and instead must make predictions based on approximate combinations of products.
5. Fit the Random Forest algorithm on this modified training data.
6. Record the Random Forest algorithm's predictions for the validation data.
7. Repeat steps 1-6 for every product *p*.

8. Take all validation predictions together and use them to compute an overall validation F1 score.

We compare the out-of-sample prediction accuracy of the Random Forest approach to that for three ways of computing density that do not involve machine learning:

- The standard approach to density
- A K-Nearest Neighbors (KNN) approach to density, where only the top 35 products with the highest proximity to each product *p* are considered when making predictions for product *p*
- A Ratio of Nearest Proximities approach to density, where the proximities of the top 10 most proximate products a country has an MCP in vis a vis a product *p* are summed; and then this sum is divided by the sum of proximities for the top 10 most proximate products overall for product *p*

In each of these three approaches to density that do not involve machine learning, we choose a decision threshold for whether predict MCP = 0 or MCP = 1 by selecting the decision threshold that maximizes the F1 score within the training data.

We find that for all 34 validation countries, the F1 score for prediction accuracy is 41.6% for standard density; 49.5% for KNN density; 49.8% for the Ratio of Nearest Proximities density; and 51.7% for the Radom Forest approach. If, however, one only considers the 10 validation countries which in any year had at least 20% of GDP accounted for by natural resources (which includes Kazakhstan), the F1 score is 18.2% for standard density; 28.6% for KNN density; 28.5% for the Ratio of Nearest Proximities density; and 39.0% for the Random Forest approach. As such, the Random Forest approach outperforms alternatives moderately when considering all countries and by a wide margin when considering resource-intensive countries.

Applying Random Forest Probabilities to Calculate Complexity Outlook Gain (COG)

The conventional definition of Complexity Outlook Gain is a weighted sum of a target product's proximities to other non-MCP products (as a share of all proximities to the other product), weighted by PCI. The underlying logic is that the target product's proximities to other non-MCP products represent the increased probability of acquiring that other product if the target product is acquired.
In a Random Forest setting we can thus substitute the proximity measure for the calculated increased probability of acquiring another product if the target product were acquired. In other words:

$$RF\ COG_{c,p=j} = \sum_{p \neq j} \left(Density_{c,p,MCP_j=1} - Density_{c,p,MCP_j=0} \right) \left(1 - MCP_{c,p}\right) PCI_p$$

UMAP Algorithm for Clustering Regions of Kazakhstan

To cluster Kazakh regions together in terms of export similarity, we take the MCPs of each world country and Kazakh region in 2019 and run them through a UMAP (Uniform Manifold Approximation and Projection) algorithm. This dimension reduction algorithm transforms complex non-linear relationships in the MCP space to continuous measures along two dimensions, which allows for easy visualization of how countries and Kazakh regions relate to each other. We select settings in the algorithm to encourage granular localized relationships between economies; specifically, we set the minimum number of neighbors to two and the minimum distance between each point to 0.1.

Gravity Model of Trade to Determine Distance Sensitivity of Products

We construct a gravity model of trade for each product as follows:

$$Export_{o,d,p} = C \cdot E_{o,p}^{\beta_1} \cdot I_{d,p}^{\beta_2} \cdot D_{o,d}^{\beta_3} \cdot F_{o,d}^{\beta_k} + \epsilon_{o,d,p}$$

Where o is for origin country, d is for destination country, p is for product, E is total exports of origin country o in product p, I is total imports of destination country d I product p, D is the distance between origin country o and destination country d (centroid to centroid), and F is a vector of controls representing additional frictions or lubricants between origin country o and destination country d (such as sharing a common border or language). Epsilon is the error term.

We log transform this equation to:

$$Log(Export_{o,d,p}) = \beta_0 + \beta_1 Log(E_{o,p}) + \beta_2 Log(I_{d,p}) + \beta_3 Log(D_{o,d}) + \beta_k F_{o,d} + \epsilon_{o,d,p}$$

Then estimate it in an OLS regression framework using global trade data covering all bilateral trade relationships between different countries.

This yields β_3, a coefficient that indicates how sensitive to distance a product is when traded. The more negative the coefficient, the more sensitive the product is to being traded across large distances.

Total Addressable Market Adjusted for Distance Sensitivity

In contexts without remoteness as an economic constraint, the Total Addressable Market for an export is usually considered to be the global value of all trade in that export. Given that Kazakhstan is constrained by its geographic remoteness, however, it is arguably important to adjust trade levels downwards to account for Kazakhstan's distance to import markets around the world.

To execute this calculation, we take the product-specific coefficients for the distance decay of trade from the Gravity Model of Trade and look at every bilateral trade relationship for every possible combination of countries and products in the world. In each case, we calculate how much trade would decay—if at all—if the exporting country's centroid was changed to that of Kazakhstan, i.e., if its location was changed to that of Kazakhstan. We sum up these counterfactual trade flows for each product to determine what Kazakhstan's Total Addressable Market is in each respective product.

This counterfactual calculation presents two possibilities for each pair of countries: one in which the exporting country becomes closer to the importing country, and one in which it becomes further. In the former case we treat counterfactual export levels as the same since exports are not made any more difficult by distance. In the latter case, however, it is necessary to calculate how trade flows would diminish relative to the baseline. One can derive the formula for the appropriate adjustment as follows:

$$\text{let } Export_{o,d,p} = C \cdot E_{o,p}^{\beta_1} \cdot I_{d,p}^{\beta_2} \cdot D_{o,d}^{\beta_3} \cdot F_{o,d}^{\beta_k}$$

Now consider how Country C's exports to Country B would change if it moved to the position of Country A.

$$\frac{Export_{C,B,p}}{Export_{C\to A,B,p}} = \frac{C \cdot E_{C,p}^{\beta_1} \cdot I_{B,p}^{\beta_2} \cdot D_{C,B}^{\beta_3} \cdot F_{C,B}^{\beta_k}}{C \cdot E_{C,p}^{\beta_1} \cdot I_{B,p}^{\beta_2} \cdot D_{A,B}^{\beta_3} \cdot F_{C,B}^{\beta_k}} = \frac{D_{C,B}^{\beta_3}}{D_{A,B}^{\beta_3}}$$

$$\therefore Export_{C\to A,B,p} = \frac{D_{A,B}^{\beta_3}}{D_{C,B}^{\beta_3}} Export_{C,B,p}$$

In other words, the ratio by which to adjust the export flow is equal to the ratio of distance-imposed frictions in counterfactual versus original location of the country.

Remoteness Tolerance for Industries of Employment

To quantify how remoteness-tolerant a particular industry is, we first quantify how remote a US Commuting Zone (CZ) is and then see how the presence of an industry depends on the remoteness of the places it is located in.

The remoteness of each CZ is calculated based on three ideas that capture the economic remoteness of a place: first, the driving time to the nearest moderately large population center (of at least 200,000 people); second, a gravity-type calculation of proximity to population centers, where a location counts as less remote if it is closer to more and larger population centers; and third, the driving time to the nearest port. Collectively, these variables aim to capture how proximate or remote a location is with respect to important economic inputs.

To begin, we calculate remoteness indices for each CZ as follows:
1. Obtain the geographic centroids of each CZ
2. Obtain the locations of all ports in the US
3. Calculate the driving times from each CZ, to each other CZ with a population of at least 200,000 (i.e. population centers)
4. Take the logarithm of the driving time from each CZ to the nearest population center, then normalize this range from 0 to 1. This represents the first relevant remoteness factor.
5. Calculate a gravity-type interpretation of driving time to population centers as follows:

a. Create a matrix of driving times from each CZ to each population center
b. Subtract each driving time from the overall maximum driving time in the matrix
c. Raise each value in the matrix to the power of 10 (this gives exponentially more weight to places that are closer rather than further)
d. Divide each value in the matrix by the largest value in the matrix
e. Take the dot product of the transformed matrix of driving times and a vector of each population center's population (this effectively assesses how close each CZ is to each population center and how big the population center is, assigning a higher weight if the CZ is closer and if the population center is larger)
f. Divide each value in the resultant vector by the maximum value in the vector. This yields the second relevant remoteness factor.

6. Calculate the logarithm of the driving times from each CZ to the nearest port, then normalize this range from 0 to 1. This represents the third relevant remoteness factor
7. Take the average score of the three remoteness factors to create an overall remoteness index for a location

Next, we relate each industry to its production patterns in different places based on how remote they are. This is executed as follows:

1. We transform RCA into a range from -1 to 1 as follows:

$$RCA' = \frac{RCA - 1}{RCA + 1}$$

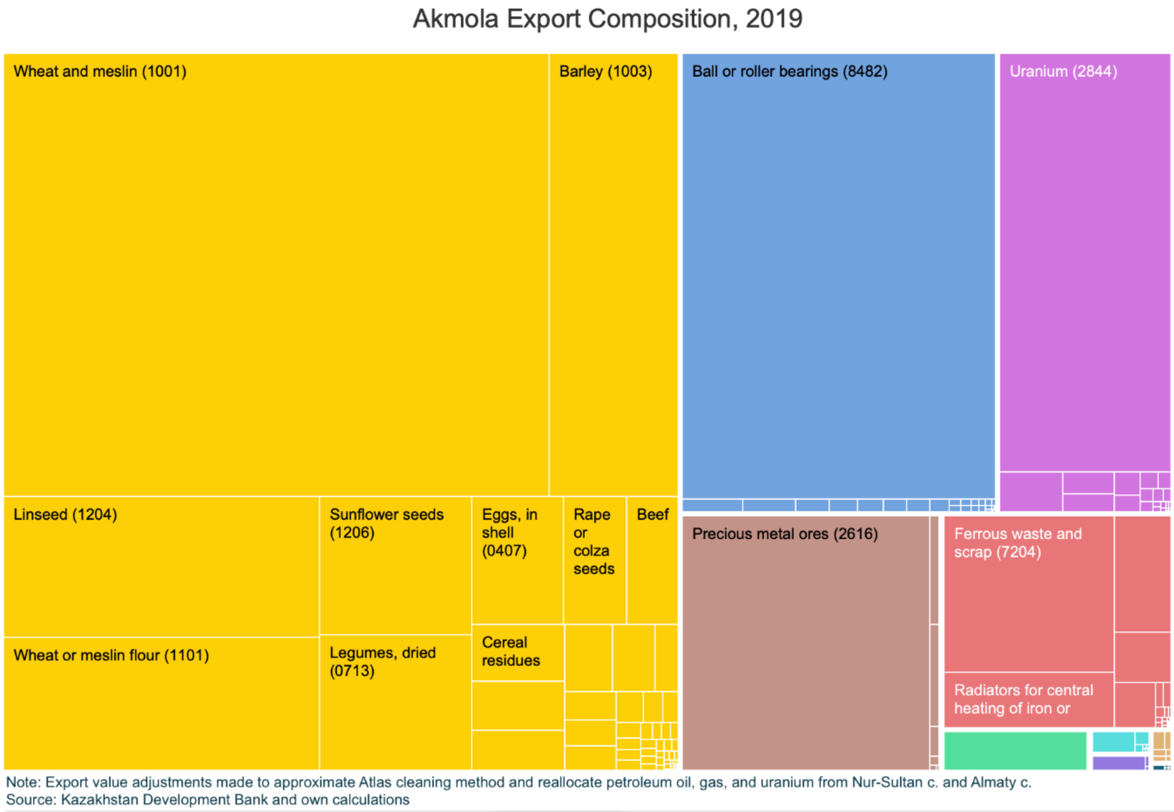
2. We run the following series of regressions where *i* is for industry and *cz* is for Commuting Zone; a separate regression is run for each industry:

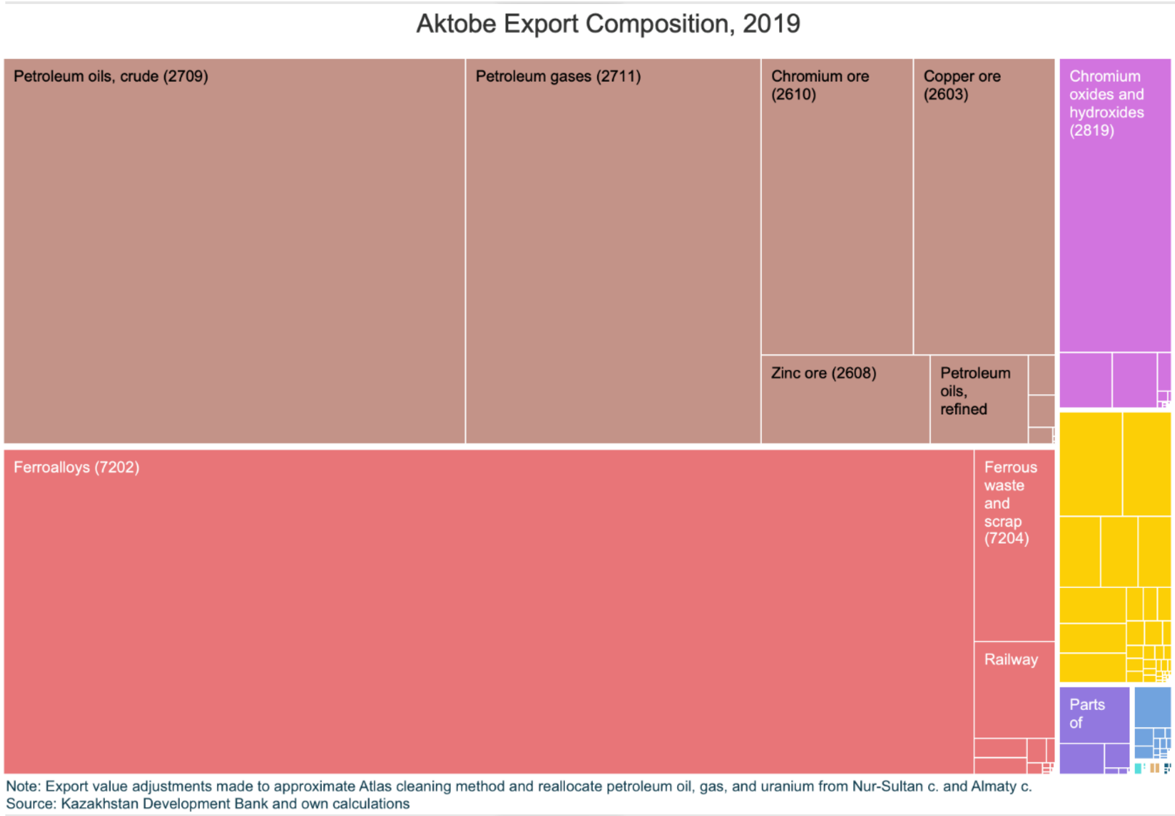
$$RCA'_{i,cz} = \beta_0 + \beta_1 Remoteness_{cz} + \varepsilon_{i,cz}$$

3. We extract the coefficient β_1 for each industry to quantify how remoteness-tolerant it is

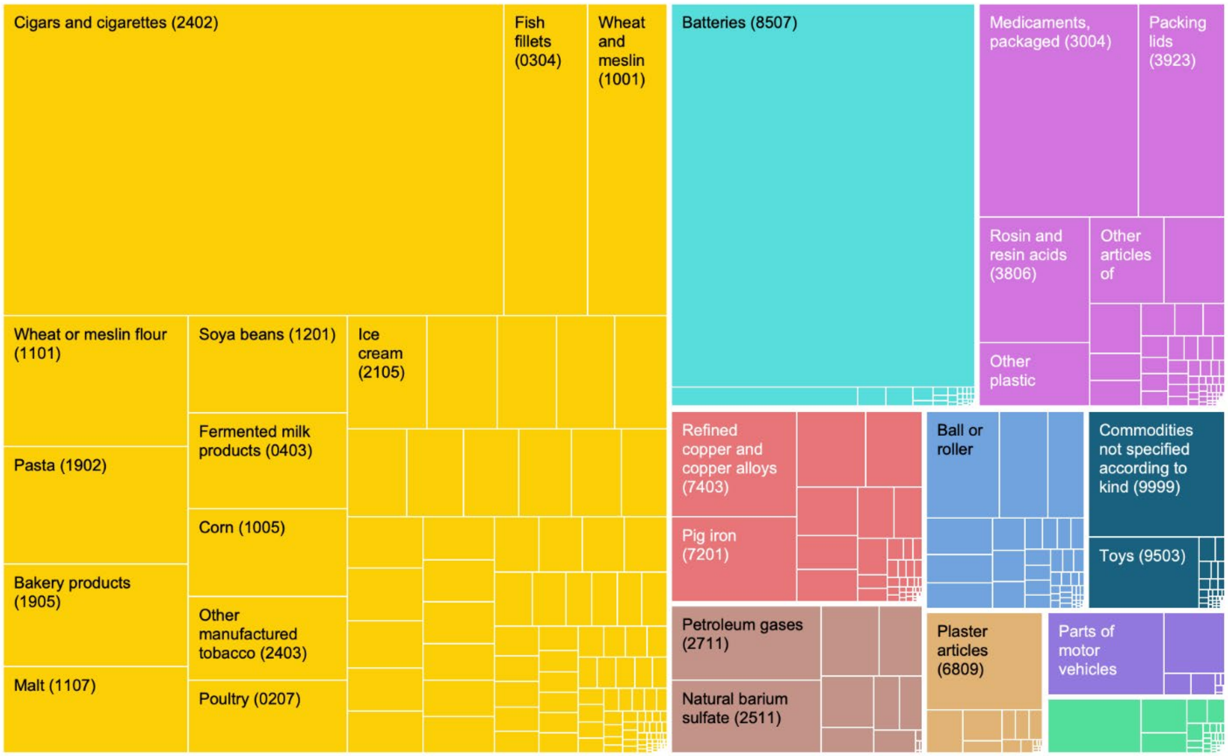
Appendix D: Additional Figures

Figure D1. Export Compositions of the Regions of Kazakhstan



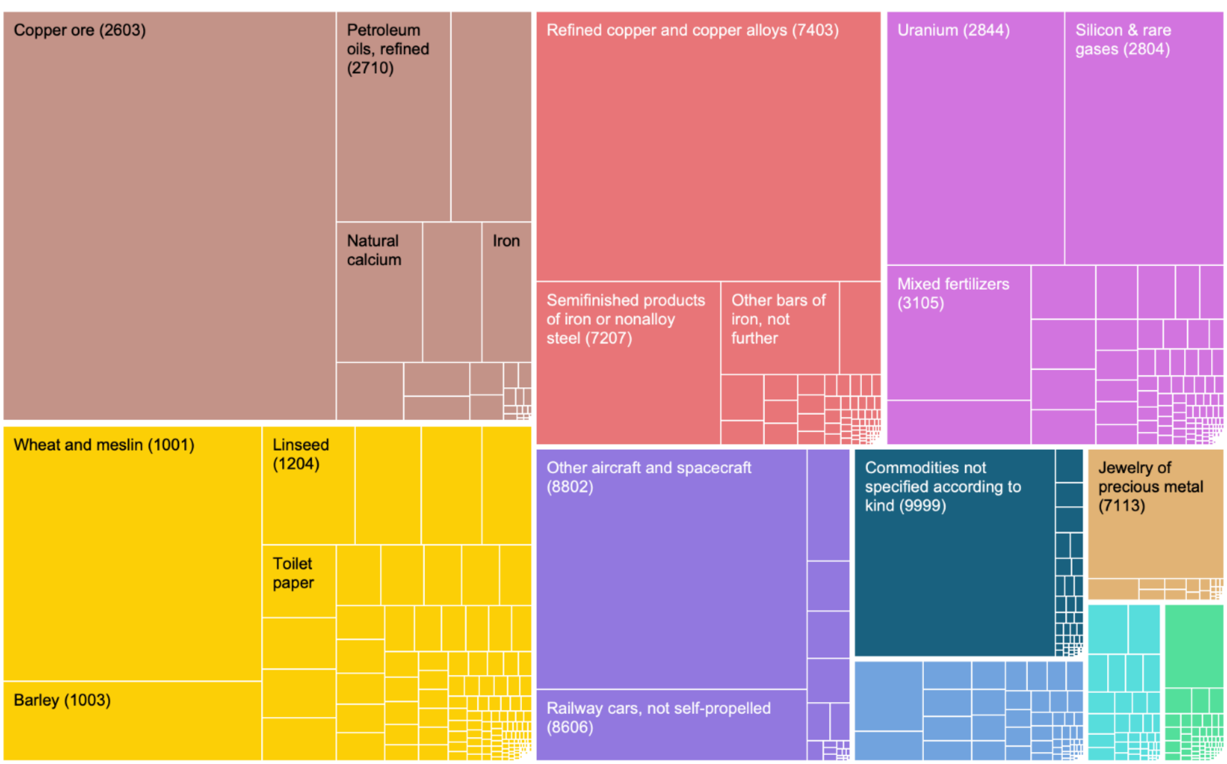


Almaty Export Composition, 2019



Note: Export value adjustments made to approximate Atlas cleaning method and reallocate petroleum oil, gas, and uranium from Nur-Sultan c. and Almaty c.
Source: Kazakhstan Development Bank and own calculations

Almaty c. Export Composition, 2019

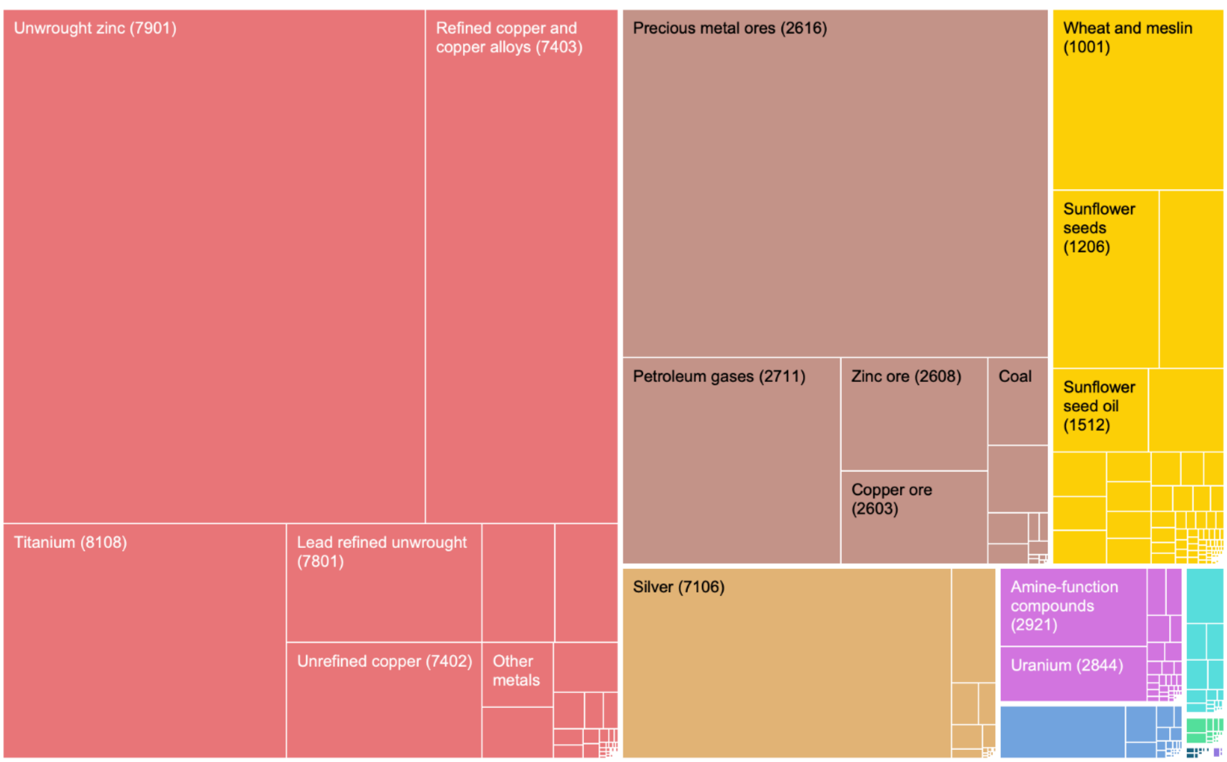


Note: Export value adjustments made to approximate Atlas cleaning method and reallocate petroleum oil, gas, and uranium from Nur-Sultan c. and Almaty c.
Source: Kazakhstan Development Bank and own calculations

Atyrau Export Composition, 2019

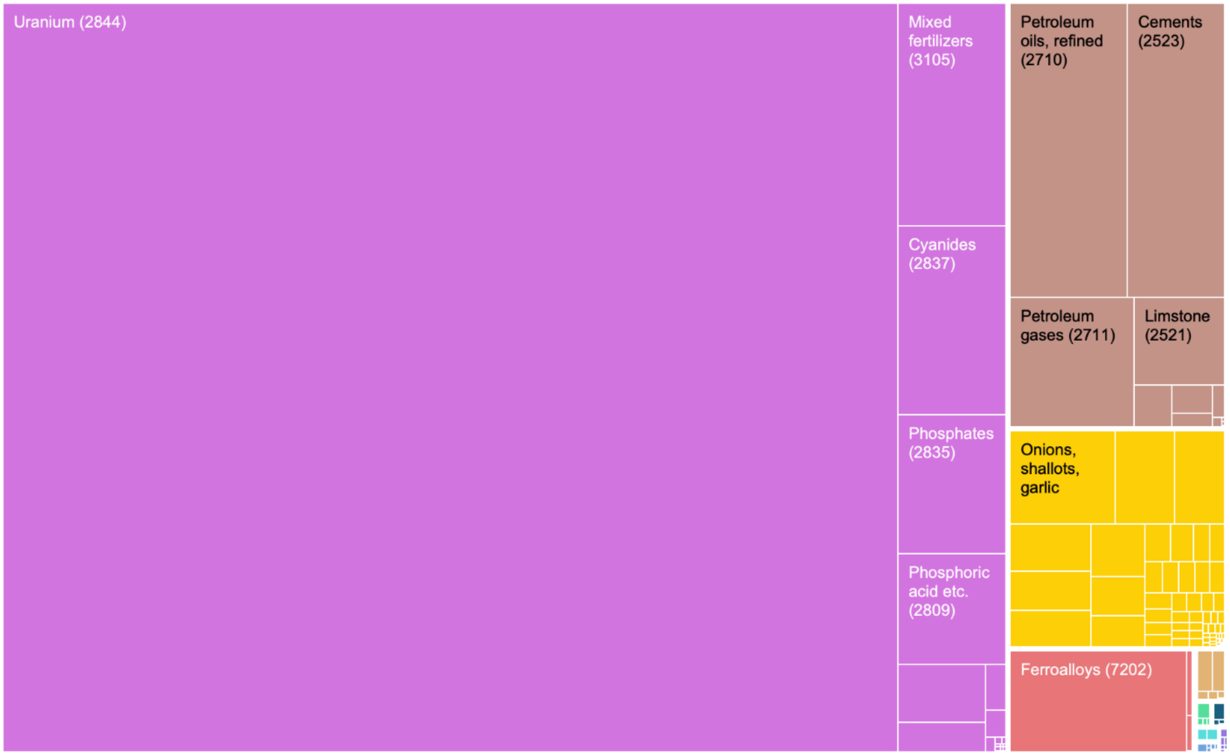
Note: Export value adjustments made to approximate Atlas cleaning method and reallocate petroleum oil, gas, and uranium from Nur-Sultan c. and Almaty c.
Source: Kazakhstan Development Bank and own calculations

East Kazakhstan Export Composition, 2019



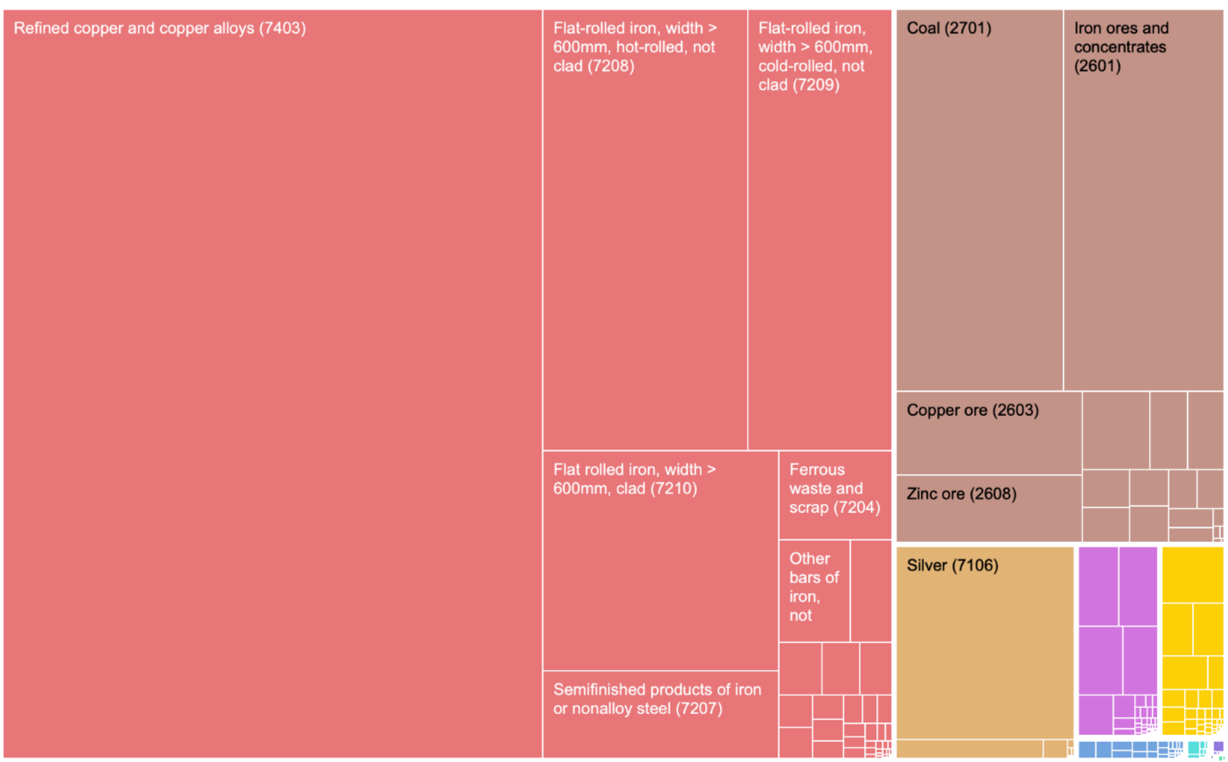
Note: Export value adjustments made to approximate Atlas cleaning method and reallocate petroleum oil, gas, and uranium from Nur-Sultan c. and Almaty c.
Source: Kazakhstan Development Bank and own calculations

Jambyl Export Composition, 2019



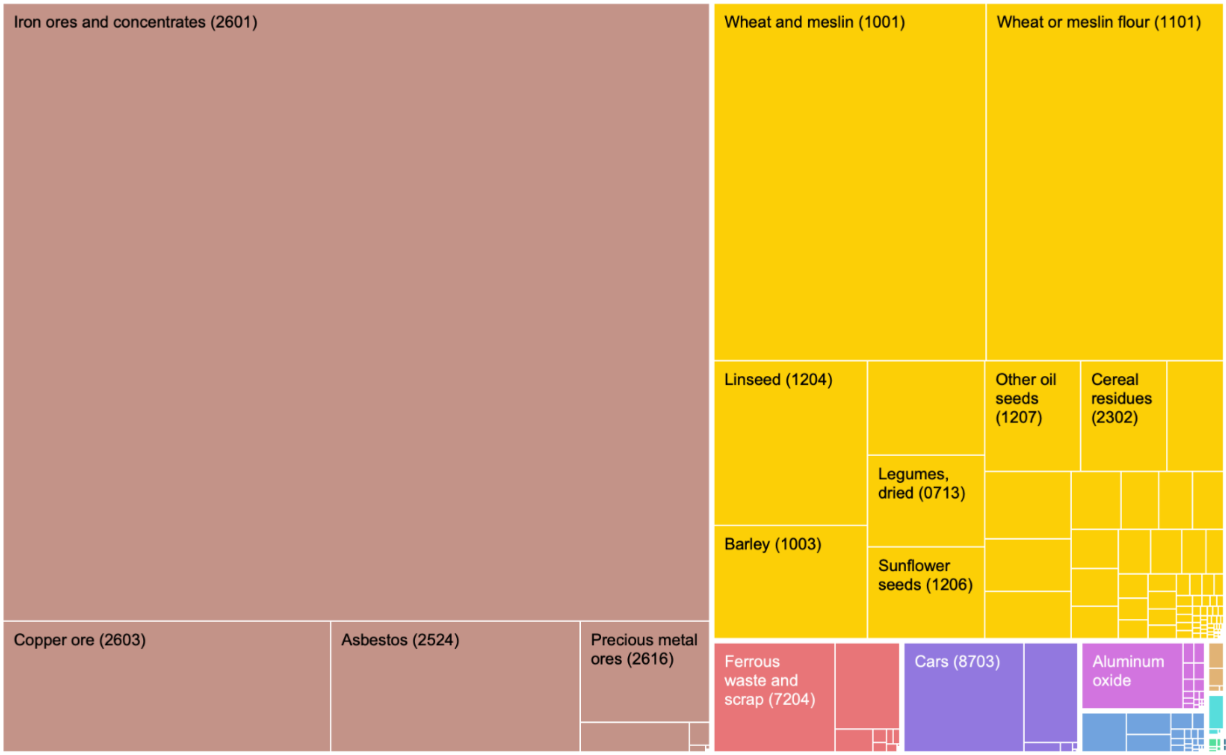
Note: Export value adjustments made to approximate Atlas cleaning method and reallocate petroleum oil, gas, and uranium from Nur-Sultan c. and Almaty c.
Source: Kazakhstan Development Bank and own calculations

Karagandy Export Composition, 2019



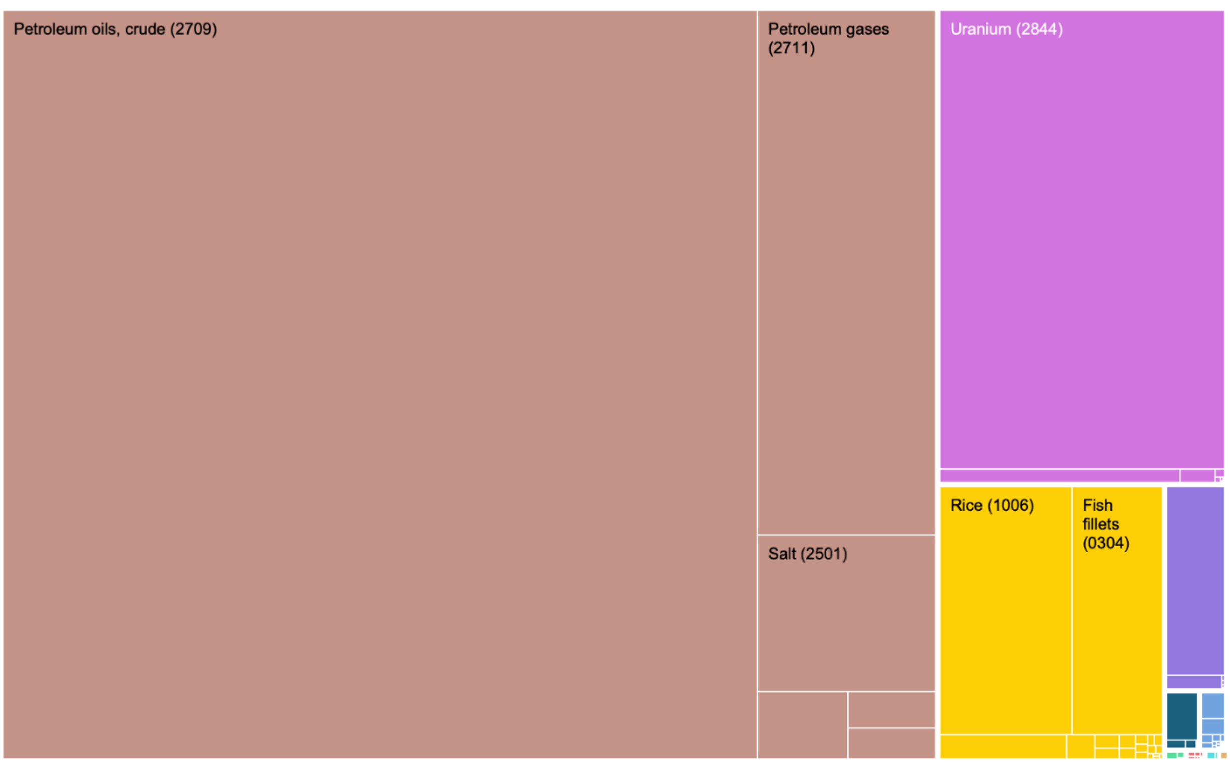
Note: Export value adjustments made to approximate Atlas cleaning method and reallocate petroleum oil, gas, and uranium from Nur-Sultan c. and Almaty c.
Source: Kazakhstan Development Bank and own calculations

Kostanay Export Composition, 2019



Note: Export value adjustments made to approximate Atlas cleaning method and reallocate petroleum oil, gas, and uranium from Nur-Sultan c. and Almaty c.
Source: Kazakhstan Development Bank and own calculations

Kyzylorda Export Composition, 2019

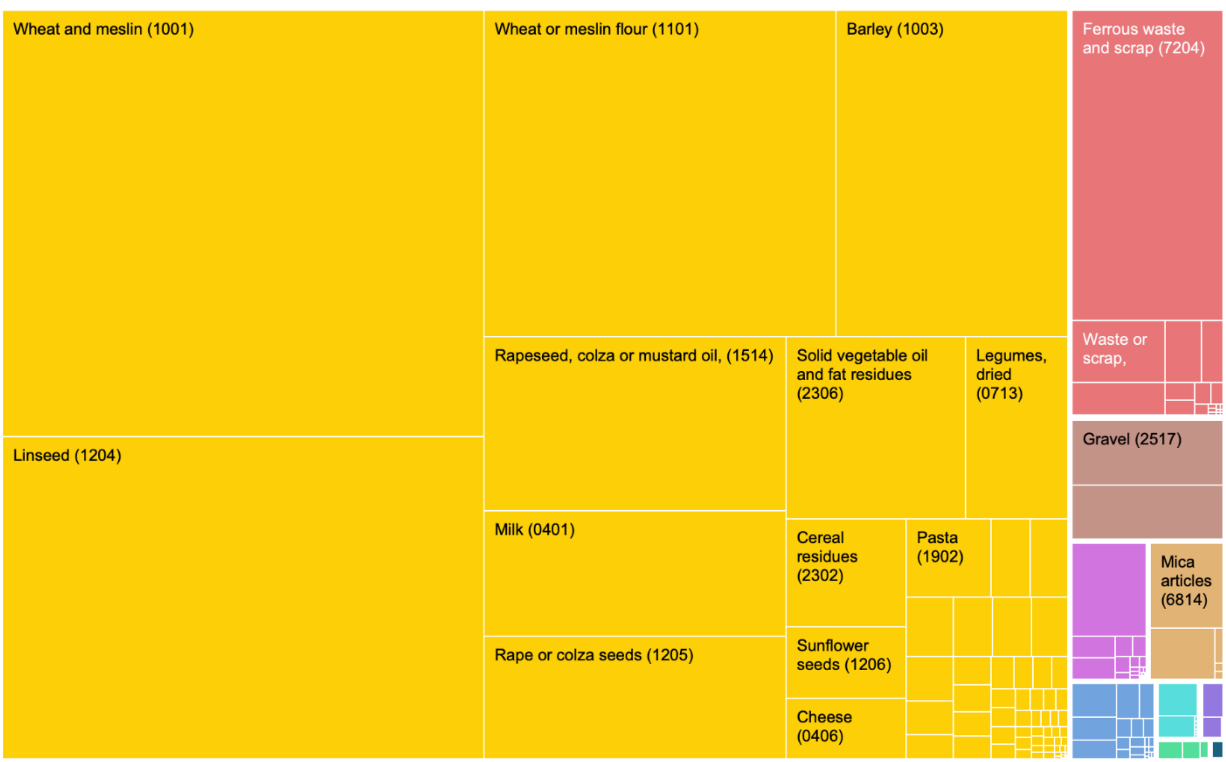


Note: Export value adjustments made to approximate Atlas cleaning method and reallocate petroleum oil, gas, and uranium from Nur-Sultan c. and Almaty c.
Source: Kazakhstan Development Bank and own calculations

Mangystau Export Composition, 2019

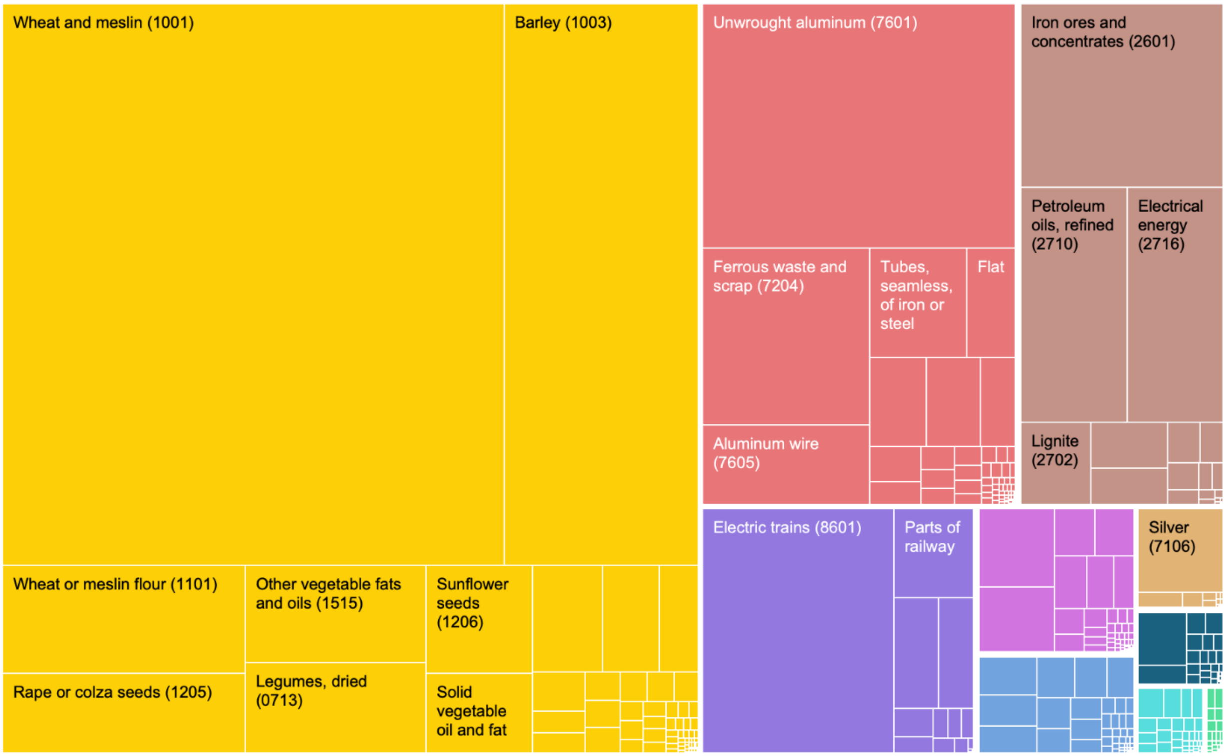
Note: Export value adjustments made to approximate Atlas cleaning method and reallocate petroleum oil, gas, and uranium from Nur-Sultan c. and Almaty c.
Source: Kazakhstan Development Bank and own calculations

North Kazakhstan Export Composition, 2019



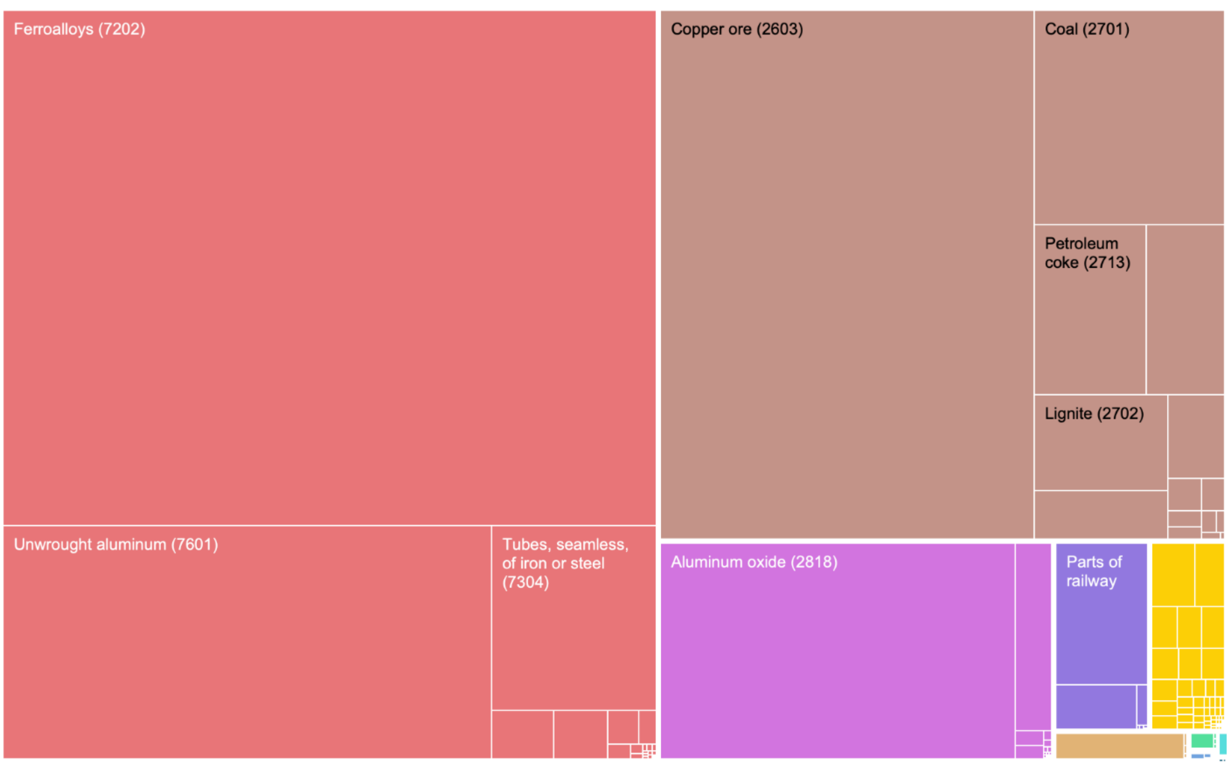
Note: Export value adjustments made to approximate Atlas cleaning method and reallocate petroleum oil, gas, and uranium from Nur-Sultan c. and Almaty c.
Source: Kazakhstan Development Bank and own calculations

Nur-Sultan c. Export Composition, 2019

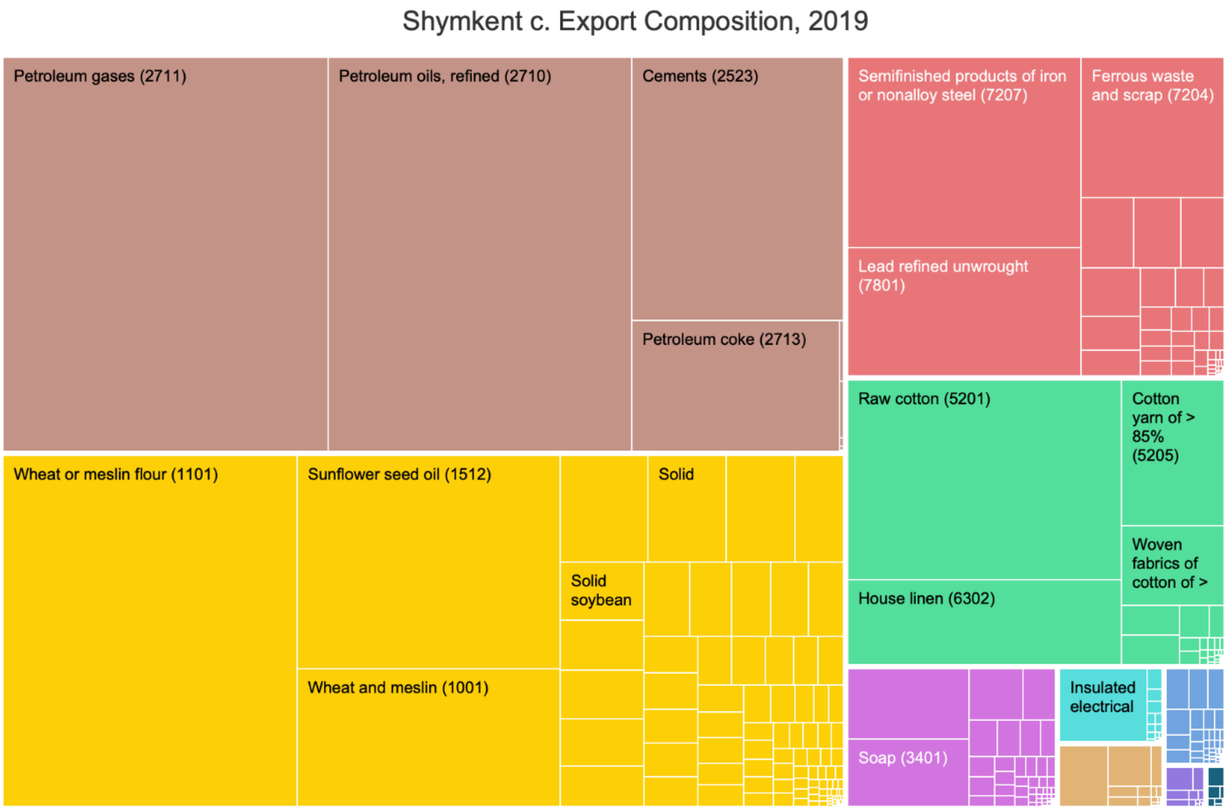


Note: Export value adjustments made to approximate Atlas cleaning method and reallocate petroleum oil, gas, and uranium from Nur-Sultan c. and Almaty c.
Source: Kazakhstan Development Bank and own calculations

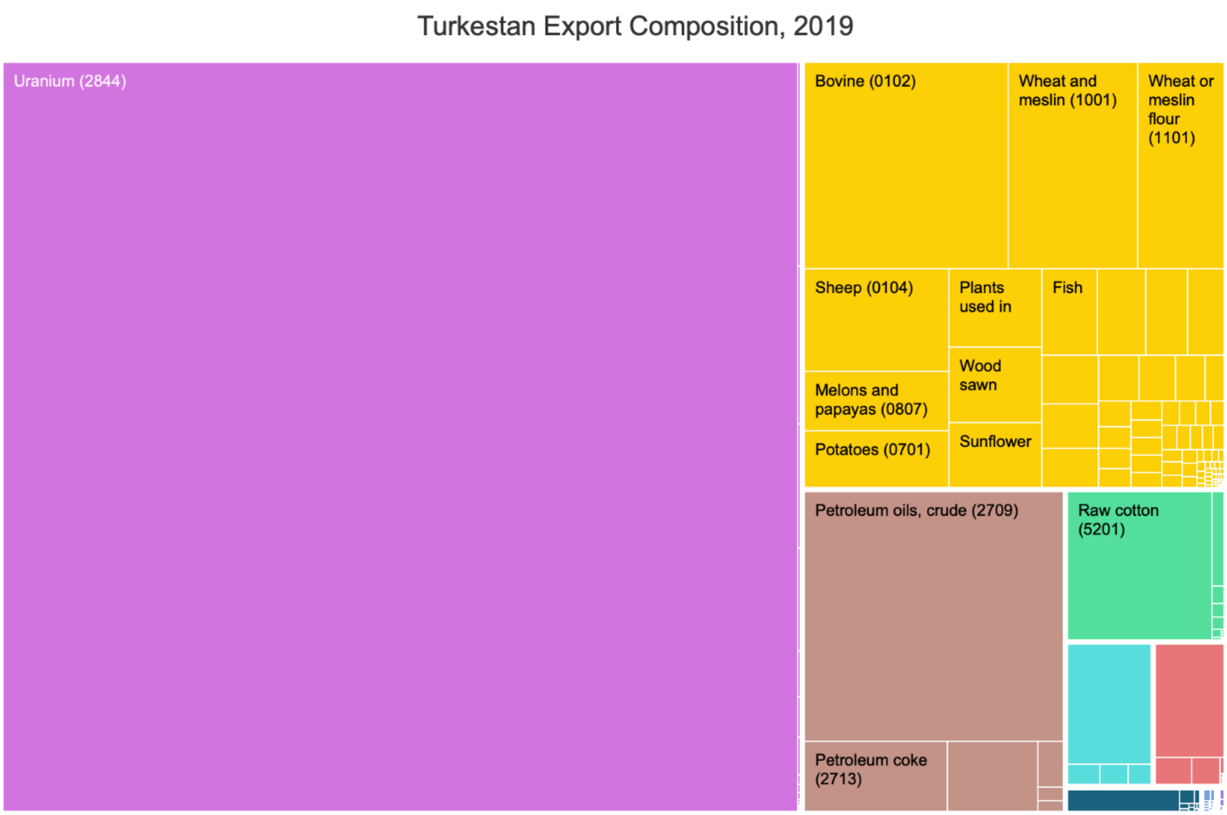
Pavlodar Export Composition, 2019



Note: Export value adjustments made to approximate Atlas cleaning method and reallocate petroleum oil, gas, and uranium from Nur-Sultan c. and Almaty c.
Source: Kazakhstan Development Bank and own calculations



Note: Export value adjustments made to approximate Atlas cleaning method and reallocate petroleum oil, gas, and uranium from Nur-Sultan c. and Almaty c.
Source: Kazakhstan Development Bank and own calculations



Note: Export value adjustments made to approximate Atlas cleaning method and reallocate petroleum oil, gas, and uranium from Nur-Sultan c. and Almaty c.
Source: Kazakhstan Development Bank and own calculations

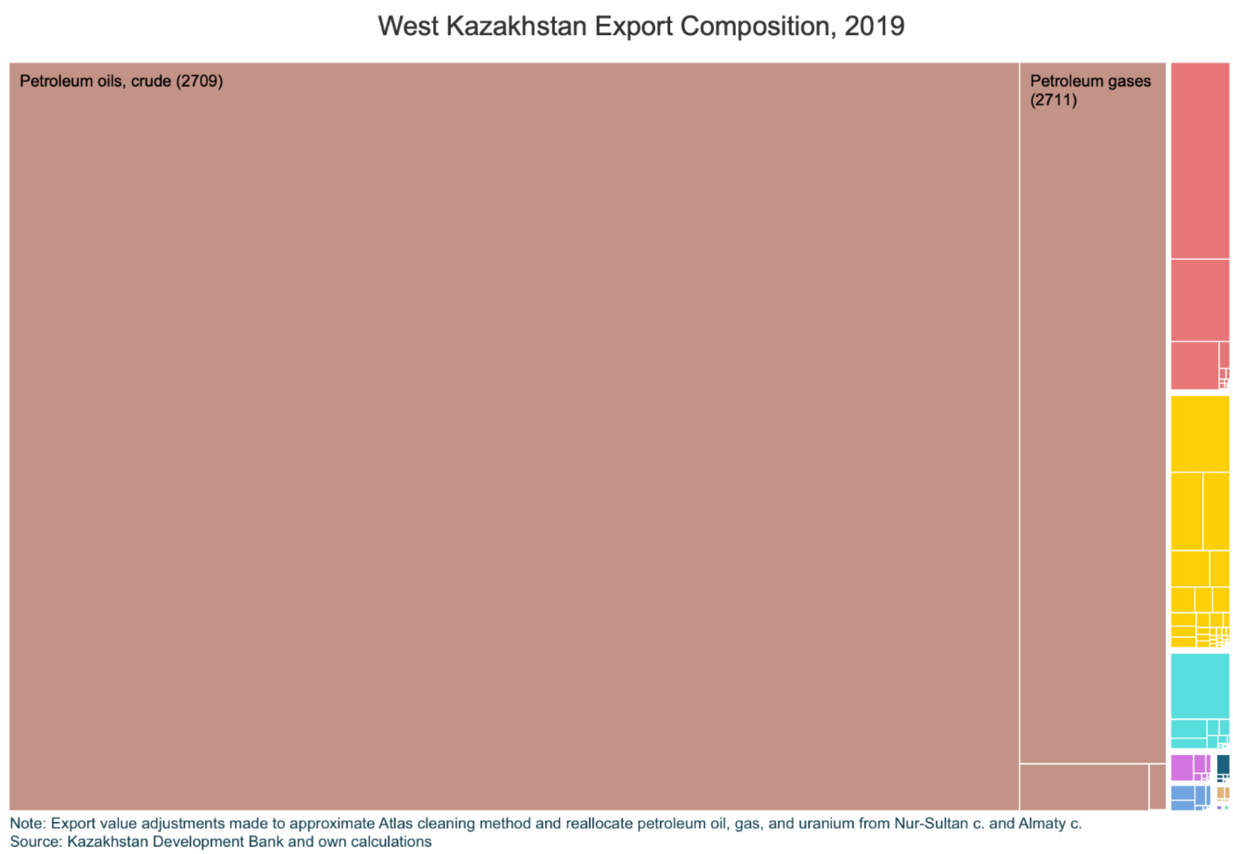
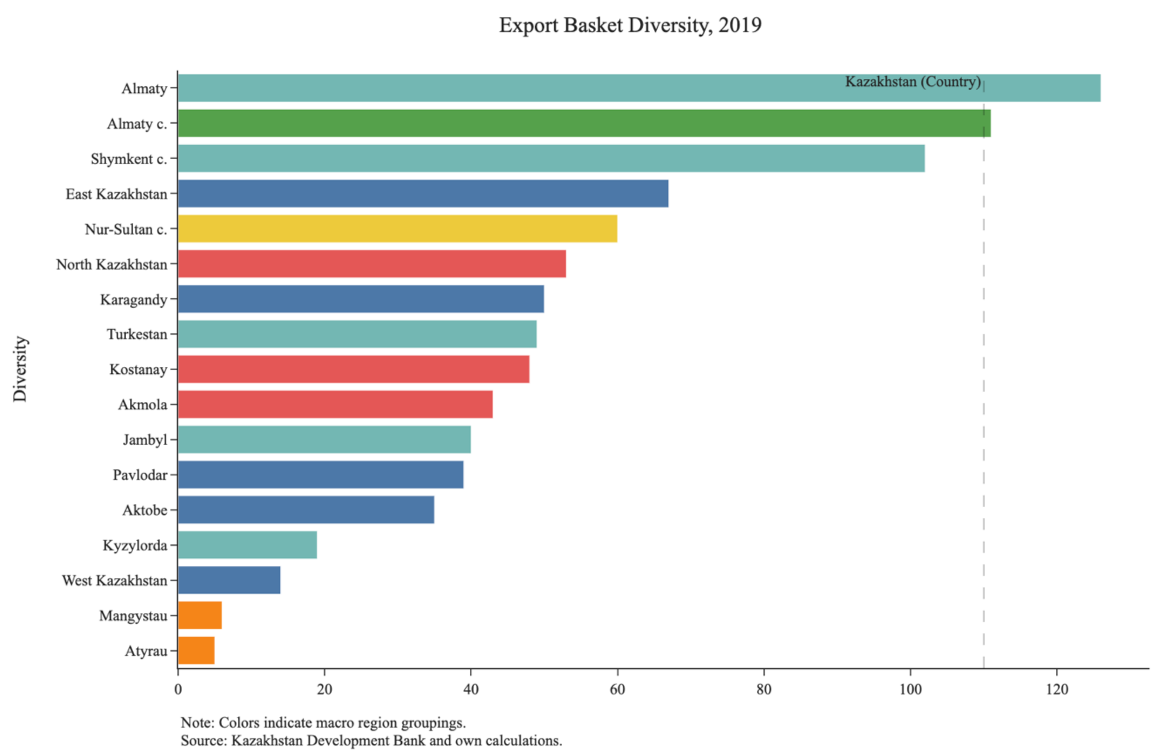
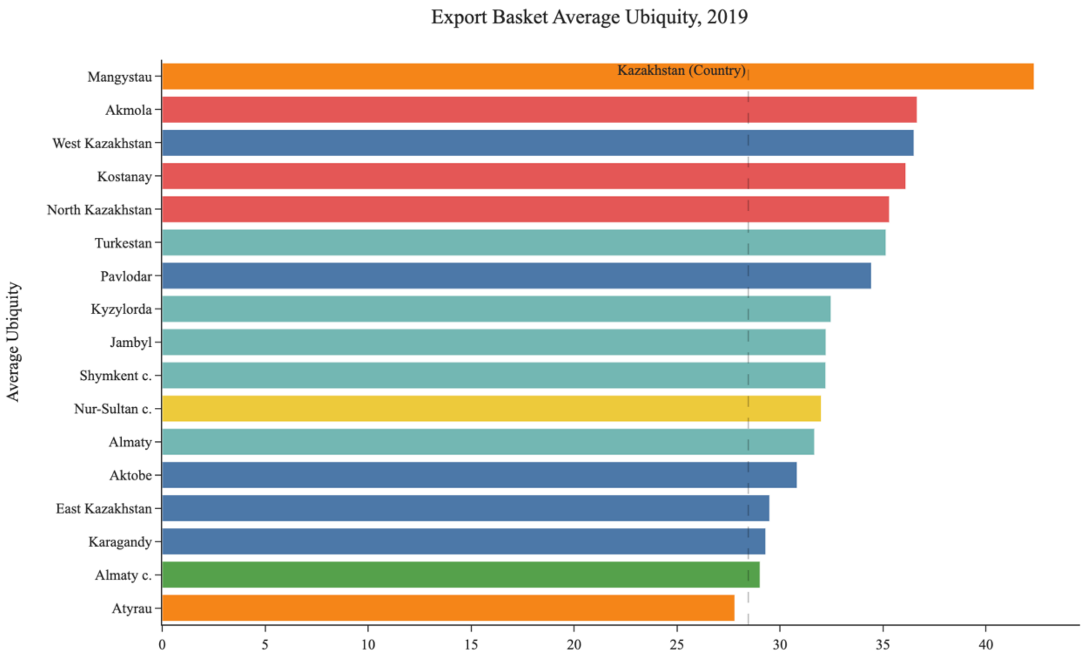
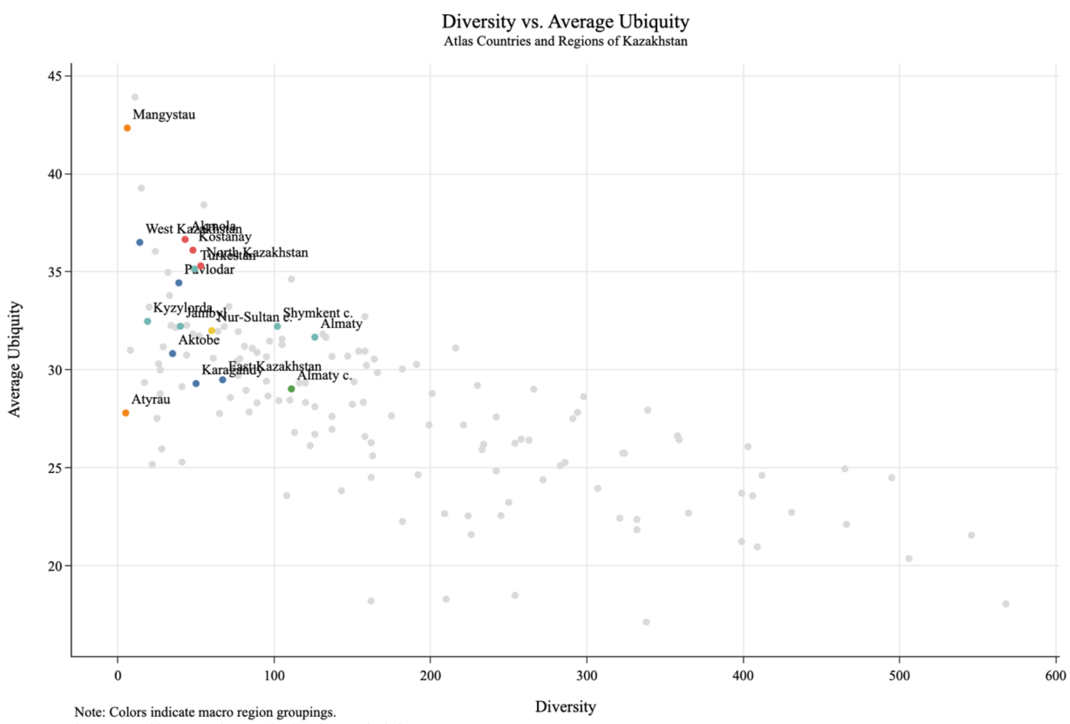


Figure D2. Complexity Metrics of the Regions of Kazakhstan

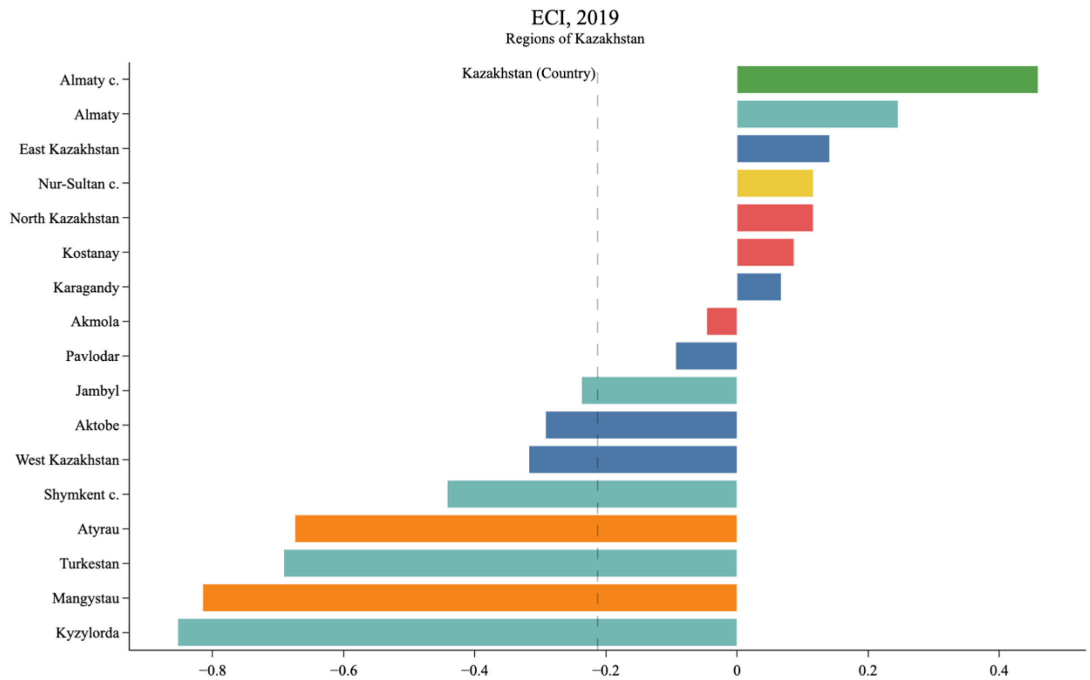




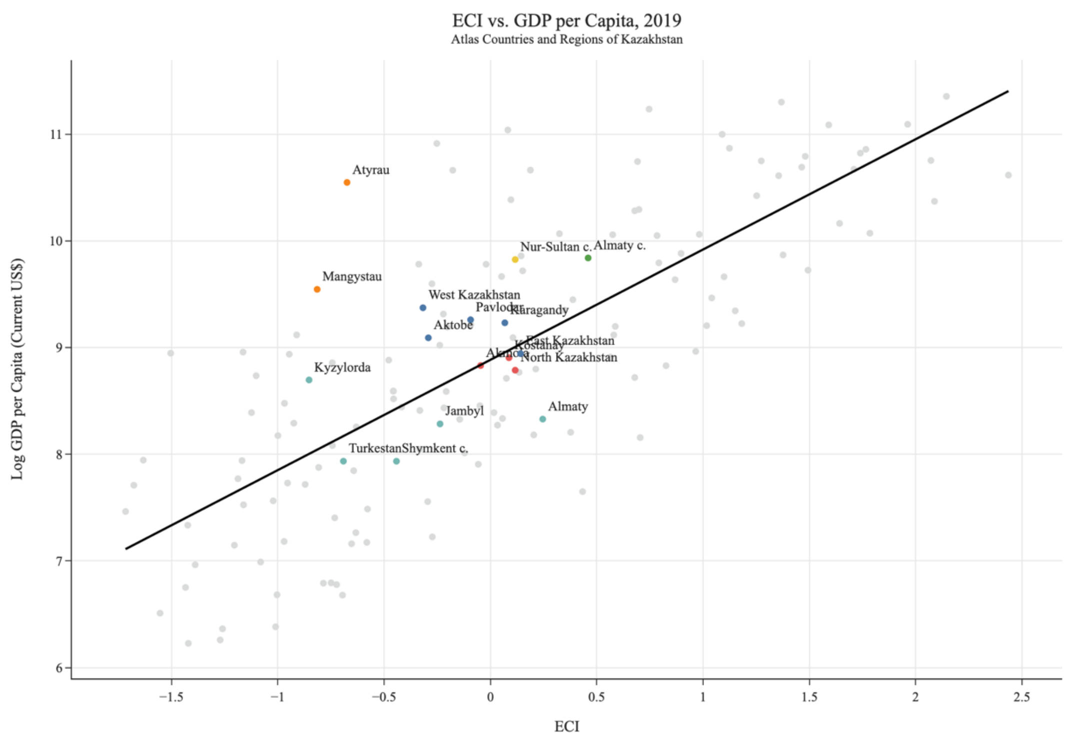
Note: Colors indicate macro region groupings.
Source: Kazakhstan Development Bank and own calculations.



Note: Colors indicate macro region groupings.
Source: Kazakhstan Development Bank and own calculations.



Source: Kazakhstan Development Bank and own calculations



Source: Kazakhstan Development Bank, Bureau of National Statistics, Atlas of Economic Complexity and World Bank WDI

Figure D3. RF Product Space Mapping of Product Recommendations Across All Macro Regions

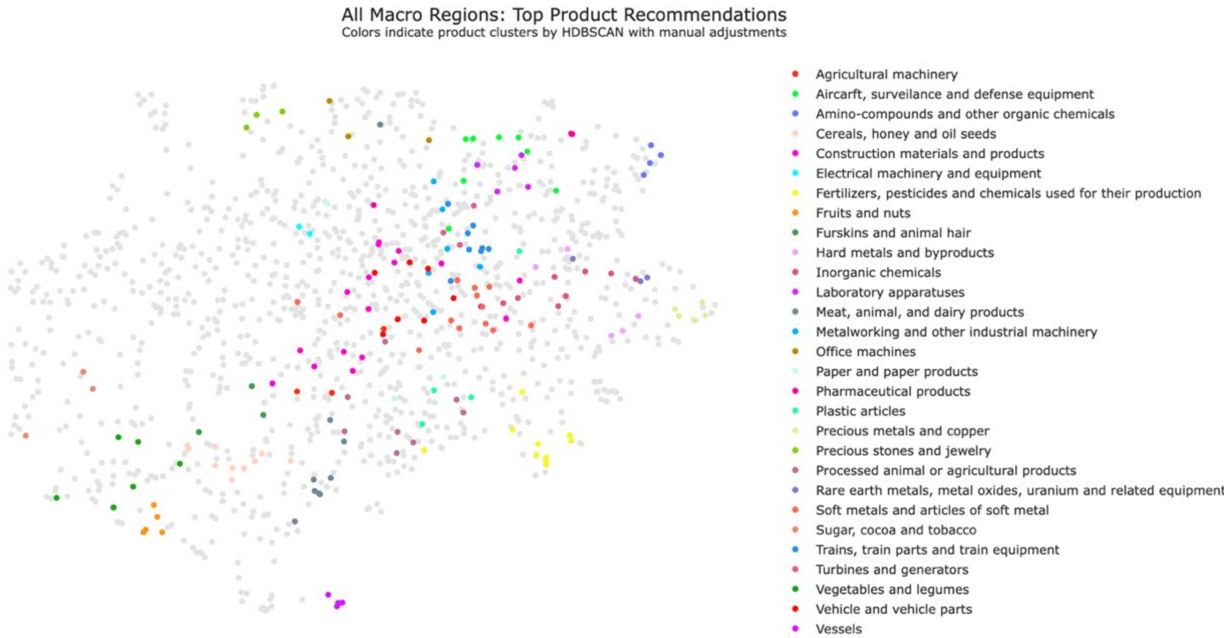
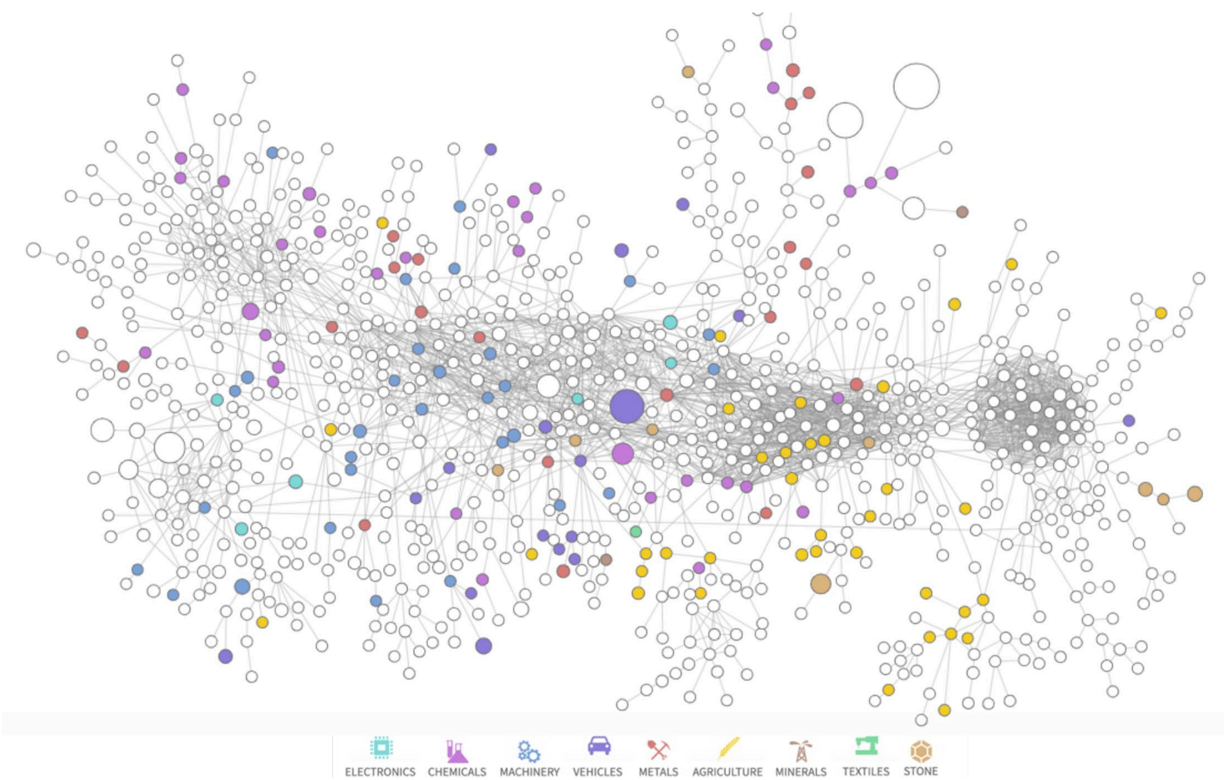


Figure D4. Traditional Product Space Mapping of Product Recommendations Across All Macro Regions



References

Asia Development Bank. "Assessing the Results of Industrial Innovation Development in Kazakhstan," June 11, 2020. https://development.asia/summary/assessing-results-industrial-innovation-development-kazakhstan.

Bekbolaeva, Dinara. "Billions of Wasted Budget Funds Spent on Industrialization Were Revealed by the Accounts Committee." *Baige News*, August 14, 2019. https://baigenews.kz/milliardy-vpustuyu-potrachennyh-na-industrializatsiyu-byudzhetnyh-sredstv-vyyavil-schetnyy-komitet_136825/.

Delgado, Mercedes, Michael E. Porter, and Scott Stern. "Clusters, convergence, and economic performance." *Research policy* 43.10 (2014): 1785-1799.

Development Bank of Kazakhstan. "Export Statistics," 2021. https://kdb.kz/en/.

Growth Lab at Harvard University. (2019). "International Trade Data (HS, 92)."

Hausmann, Ricardo, and Bailey Klinger. "Structural transformation and patterns of comparative advantage in the product space." (2006).

Hausmann, Ricardo, Cesar Hidalgo, Sebastián Bustos, Michele Coscia, Sarah Chung, Juan Jimenez, Alexander Simoes, and Muhammed Yildirim (2011). "The Atlas of Economic Complexity". Puritan Press. Cambridge MA.

Hausmann, Ricardo, et al. "Growth Perspective on Western Australia." (2021).

Hausmann, Ricardo, et al. The atlas of economic complexity: Mapping paths to prosperity. MIT Press, 2014.

Kazakhstan Bureau of National Statistics. "Data - Taldau," 2022.

Kubayeva, Gulaikhan, and Alibek Konkakov. "Progress in Diversification of the Economy in Kazakhstan." The Netherlands, 2016.

Obaya, Martin, and Ernest Stein. "Public-Private Dialogue for the Formulation of Productive Policies: The Experience of Sectoral Roundtables in Argentina." *Inter-American Development Bank*, 2021. https://publications.iadb.org/es/el-dialogo-publico-privado-para-la-formulacion-de-politicas-productivas-la-experiencia-de-las-mesas.

OECD. "OECD Territorial Reviews: Kazakhstan," June 15, 2017. https://doi.org/10.1787/9789264269439-en.

———. "Regional Policies to Support Diversification and Productivity Growth in Kazakhstan." OECD Eurasia Competitiveness Programme, April 2020. https://www.oecd.org/eurasia/competitiveness-programme/central-asia/Regional-Policies-to-Support-Diversification-and-Productivity-Growth-in-Kazakhstan-ENG.pdf.

Office of the President of the Republic of Kazakhstan. "Experts on the Address: Forecast Scheme Will Allow Regions to Use Competitive Advantages and Identify Reserves of Economic Growth," October 20, 2018. https://primeminister.kz/en/news/17323.

———. "State Program for Industrial and Innovative Development 2010-2014," March 19, 2010. https://afmrk.gov.kz/en/activity/strategy-and-program/the-state-program-of-industrial-innovative-develo.html.

———. "State Program for Industrial and Innovative Development 2015-2019," August 1, 2014. https://policy.asiapacificenergy.org/node/309.

———. "State Program for Industrial and Innovative Development 2020-2025," February 15, 2018. https://baiterek.gov.kz/en/programs/gosudarstvennaya-programma-industrialno-innovacionnogo-razvitiya-respubliki-kazakhstan-na-2020-2025.

———. "Technological Upgrade, Investment Inflow, Innovation Boom, or How Kazakhstan Made an Industrial Breakthrough," December 12, 2019.

https://primeminister.kz/en/news/reviews/technological-upgrade-investment-inflow-innovation-boom-or-how-kazakhstan-made-an-industrial-breakthrough.

Office of the Prime Minister of the Republic of Kazakhstan. "Forecast Scheme for Spatial-Territorial Development of Kazakhstan Till 2020," July 21, 2011.

———. "On Industrialization Map," December 31, 2014. https://adilet.zan.kz/eng/docs/P1400001418.

OffshoreTechnology.com. "Karachaganak Conventional Gas Field, Kazakhstan". (2021).

Shen, J.H., H Wang, and S.C. Wang. "Productivity Gap and Inward FDI Spillovers: Theory and Evidence from China." *China & World Economy* 29, no. 2 (n.d.): 24–48.

World Bank. "World Development Indicators," 2022. https://data.worldbank.org/indicator.