TRAGEDIES of Regional Conflicts & PROMISES of Peacebuilding

Responding to Disruptors and Enablers of MENA Development Pathway

ERF 30th Annual Conference

www.erf.org.eg

Stages of Diversification Redux

Hoda Assem,
Roberta Gatti
and Daniel Ledermau

ECONOMIC RESEARCH FORUM

منتدى البحوث الاقتصادية

2024

Stages of Diversification Redux

Hoda Assem, Roberta Gatti and Daniel Lederman[1]

This draft: October 3rd, 2023

Abstract

The existing literature on development and economic diversification finds an inverted-U function between these two variables, whereby economies diversify as they grow up to a point, after which they start specializing. This paper contributes to this literature by investigating the stages of diversification over the course of development during the past 57 years. The paper emphasizes the trajectories of resource-rich and resource-poor countries, an issue that has not been covered by the extant literature. In addition, the paper studies the stages of diversification across three dimensions, namely employment, value-added, and exports. Additionally, it examines the relationship for services. Non-parametric estimations suggest a U-shaped curve between measures of economic concentration and per capita income levels, which is in line with existing evidence. However, these patterns are mainly driven by between-country rather than within-country variation, a finding that had been ignored in the existing literature. Diversification patterns also differ across resource-rich and resource-poor countries: Employment and value added in resource-rich countries are on average more concentrated at low levels of development while in resource poor countries, they are more concentrated at high levels of development. In contrast, at all levels of development, exports are more concentrated in resource-rich countries.

JEL classification: F1, F43, O11, O40

[1] Chief Economist Office, Middle East and North Africa region, World Bank. We are grateful to Jean Imbs, Anne-Celia Disdier, and participants to the 2023 annual International Association of Energy Economics conference for useful comments. All errors are our own.

1. Introduction

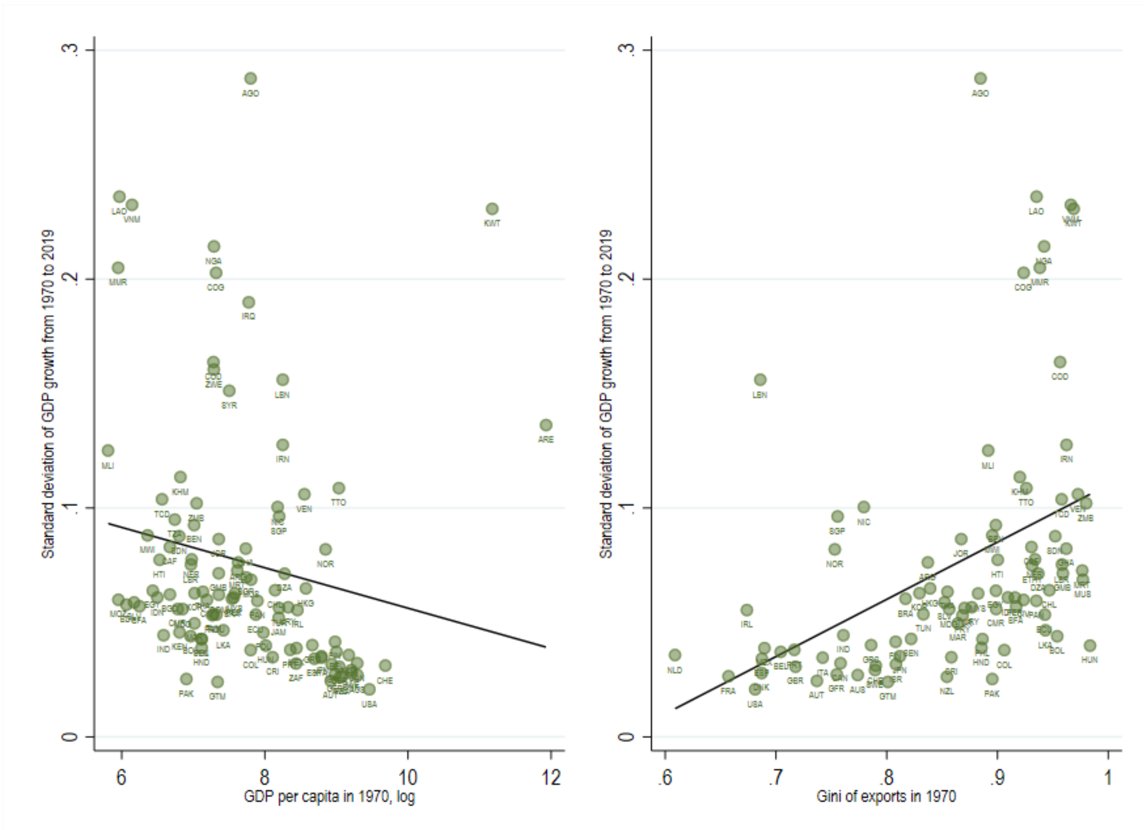
Economic diversification often features prominently in policymakers' development objectives (e.g., Saudi 2030) due to various concerns, including a concern over macroeconomic volatility (Lederman et al. 2021; Lederman and Maloney 2012). That is, it is likely that economic diversification tends to reduce macroeconomic volatility. Figure 1 shows that, in fact, growth volatility (proxied by the standard deviation of GDP per capita growth rates) during 1963-2019 declined with GDP per capita and rose with the concentration of exports across products. These bivariate correlations thus lend credence to the literature on diversification and macroeconomic volatility.

Perhaps more importantly, the academic literature on diversification has focused on the relationship with economic development. Theoretical models describing the relationship between diversification and economic growth can be classified into two sets of models. The first set of models, inspired by either neoclassical Ricardian or factor proportions (Heckscher-Ohlin) trade theory, predicts a negative relationship between diversification and growth. Open economies specialize in sectors or products for which they have a comparative advantage, suggesting that the relationship between diversification and economic development would be negative.

Another set of models predicts a positive relationship between diversification and development. At low levels of development, in the presence of fixed costs and imperfect capital markets, capital-poor countries cannot pay for the fixed cost of investing in new sectors, and therefore have limited diversification and are vulnerable to shocks. As countries accumulate capital and grow, they can afford the fixed costs of entering into new sectors and diversify, thus offering protection against volatility caused by sector-specific shocks (Acemoglu and Zilibotti, 1997). Subsequent literature has linked economic concentration with macroeconomic volatility which can have adverse effects on economic growth and investment levels (Koren and Tenreyro 2007; Van der Ploeg and Poelhekke, 2010; Lederman et al. 2021).

In a panel of 66 countries over 34 years, Imbs and Wacziarg (2003) show that the relationship between GDP per capita and measures of concentration in employment and value added is U-shaped. More specifically, at low level of income, economic activity diversifies as countries grow, but then, at around 9,000 (constant 1985$) GDP per capita, economic activity concentrates as countries grow richer. Klinger and Lederman (2004; 2006), Hesse (2008) and Cadot et al. (2011) also find a U-shaped pattern between measures of diversification of exports and income per capita.

Figure 1: Growth Variability, Development, and Economic Concentration



Source: Authors' calculations based on data from Penn World Table and UN COMTRADE.

Bringing both sets of literature together, this paper studies the patterns of economic diversification over the past 57 years. The contribution of this paper to the literature is fourfold. First, by taking advantage of longer and expanded data series, this paper updates our understanding of diversification patterns across countries and over time to 151 countries over 57 years. Second, it investigates whether the observed patterns are more likely driven by cross-country characteristics or by within-countries trajectories. Thirdly, we expand the study of the relationship between diversification and economic development to services, which has not been covered by the existing literature, to the best of our knowledge. Lastly, it explores whether diversification patterns differ across resource-rich and resource-poor countries.

Consistent with the literature, we measure economic concentration using Gini and Herfindahl indexes. We use employment and value-added data from the UNIDO INDSTAT database as well as data on exports from the UN COMTRADE to calculate measures of concentration for the manufacturing sector. We use data from WTO statistical portal[2] for trade in services. GDP per capita, PPP in constant 1985$ is obtained from Penn World Tables to proxy for the level of economic development.

To describe the relationship between economic diversification and development, we use non-parametric estimation because it imposes little structure on the functional form, allowing us to better identify the relationship between measures of concentration and GDP per capita. More specifically, we employ the *lowess* estimator which essentially carries out a locally weighted OLS regression of the dependent variable on the independent variable and plots the predicted values. We also estimate a quadratic relationship between measures of concentration and income per capita. Finally, we also apply the semi-parametric strategy developed by Imbs and Wacziarg (2003) to our data as a robustness check. Their strategy implements a local fixed-effects regression of the dependent variable on the independent variable. Including country-fixed effects allows the estimation of the intercept and slope coefficients and therefore distinguish between within and between country dimensions.

Estimating a *lowess* curve for different measures of concentration against GDP per capita, we find a U-shaped curve which is in line with the literature on economic diversification and development (Imbs and Wacziarg, 2003; Klinger and Ledermann, 2004 and 2006; Cadot et al., 2011). Countries diversify until they reach a turning point of GDP per capita of 20,000 (constant 1985$) then they start re-specializing. These results are robust across different data sources including employment, value-added and exports, different levels of disaggregation, as well as different sectors (manufacturing versus services).

[2] https://stats.wto.org/

Using the full sample, we break down diversification patterns into between and within-country components. Diversification patterns are mainly driven by between rather than within-country variation,

We distinguish between resource-poor and resource-rich countries and find that these patterns largely hold with some qualifications. The level of economic concentration is on average higher for resource-rich countries, but re-specialization is steeper for resource-poor countries. Additionally, diversification in resource-rich countries reaches a plateau earlier than in resource-poor countries (at GDP per capita of 10,000$) and re-specialization starts later, at a higher income per capita than resource-poor countries.

The rest of the paper is organized as follows. The next section summarizes the existing literature. Section 3 presents the data and explains how measures of concentration are calculated. Section 4 outlines the empirical strategy we use and presents the results. Section 5 concludes.

2. Related Literature

Ricardian trade theory predicts that open economies specialize in their comparative advantages as they grow. Specialization can also be due to agglomeration effects. Decreasing transport costs reduces the number of domestically produced products, thus promoting specialization (Dornbusch et al., 1977). Additionally, demand externalities make it profitable for producers to cluster, so this might also lead to sectoral concentration, but not necessarily to specialization across products.

In contrast, Acemoglu and Zilibotti (1997) develop a model that predicts that countries at early stages of development cannot diversify due to capital scarcity and fixed costs of initiating production in new industries or products. As a result, economic agents only invest in a limited number of sectors. As countries grow, they accumulate capital and diversify risks by investing in more productive projects, leading to more stable economic growth. Their paper also highlights a known correlation between early stages of development and economic volatility. In the same vein, Koren and Tenreyro (2007) build a model to study volatility at different stages of development. They show that sectoral diversification is an essential factor in growth stability. Countries in the early stages of development are more volatile because they specialize in fewer and more volatile sectors and experience more frequent and more severe aggregate shocks. Countries specialize in less volatile sectors as they develop, and country-specific aggregate volatility also falls due to more political stability and sounder macroeconomic policies.

Starting with Imbs and Wacziarg (2003) seminal work, many papers investigated the empirical relationship between concentration and income per capita using concentration measures based on domestic production and labor data. Imbs and Wacziarg (2003) find a U-shaped pattern whereby countries in their early development stages diversify and re-specialize later at higher income levels. In their specification, this pattern is robust across different measures of concentration, levels of disaggregation and datasets. The turning point from domestic diversification to re-specialization is robust at around US$9,000 per capita, which is relatively high, signifying that most developing countries are in the diversifying stage throughout their development path. Their evidence may be consistent with countries diversifying because consumers demand a wider variety of goods as their income rises. Additionally, producers invest in a wide range of sectors to reduce risk, which leads to diversification (Acemoglu & Zilibotti, 1997).

This pattern of diversification and re-specialization also holds for exports, as shown by Klinger and Lederman (2004; 2006), Hesse (2008) and Cadot et al. (2011), but for exports the turning point is at a higher GDP per capita level which means that only very advanced economies re-concentrate their exports. For example, Cadot et al. (2011) find a turning point of 25,000$ per capita at PPP. They also decompose their measure of concentration (the Theil index) into an "intensive" and "extensive" margin, finding that the extensive margin drives the concentration patterns. The hump-shaped pattern is consistent with the assumption that countries travel across diversification cones, as discussed in Schott (2003, 2004), thus linking diversification to neo-classical trade theory.

This paper is also more broadly related to the literature on structural transformation. Structural transformation signifies the reallocation of economic activity across the three main sectors of an economy: agriculture, manufacturing, and services. The literature on structural transformation has shown that as countries grow richer, employment and value-added shares in the agriculture sector have declined, increased in the services sector, and followed an inverted u shape in the manufacturing sector (Herrendorf et al., 2014). Our paper looks at these patterns at higher level of disaggregation.[3]

3. Data description and measurement

We rely on a variety of sources to measure the dimensions of economic diversification.

3.1 Employment and value-added

For employment and value added, we use UNIDO's Industrial Statistics Database (INDSTAT) which provides disaggregated data on the manufacturing sector at two levels of disaggregation (2-digit and 4-digit). INDSTAT 2 is measured at the 2-digit level of ISIC (International Standard Industrial Classification) Rev.3 and comprises 23 manufacturing sectors, such as "food and beverages", "textiles" etc. INDSTAT 4 provides a finer taxonomy across 137 manufacturing sectors. For example, the category of Textiles includes such sub-categories as "preparation and spinning of textile fibers", "Weaving of textiles", "Knitted and Crochet Fabrics" etc.[4] Note that a previous dataset, INDSTAT3, provided disaggregated data on the manufacturing sector at the 3-digit level of ISIC Rev.2 which comprised 28 manufacturing sectors.[5] INDSTAT 3 was used in Imbs and Wacziarg (2003). INDSTAT 2 was initially derived from INDSTAT 3.[6]

While INDSTAT datasets provide several indicators of industrial statistics, we will focus on the number of employees (employment) and value-added in our calculations of measures of concentration. Employment measures the total number of people working in or for the establishment during the reference year. Value-added measures the value of output minus the value of input in current prices during the reference year.

Coverage of INDSTAT datasets varies considerably across countries in terms of years and sectors depending on data availability. INDSTAT2 includes 172 countries for the period

[3] At the most aggregated level (1-digit ISIC Rev 3.1), we find an r-shaped relationship between concentration of exports and GDP per capita across all sectors. This r-shaped relationship could be evidence of structural transformation as countries grow. At more disaggregated levels, we find a u-shaped relationship where countries diversify as they grow up to a given level of GDP then start specializing. Results are available upon request.

[4] See Table A1 and Table A2 in appendix for a list of sectors in INDSTAT 2 and INDSTAT 4 respectively.

[5] See Table A3 in appendix for a list of sectors in INDSTAT 3.

[6] INDSTAT 3 was discontinued in 2007. Table A4 provides a crude conversion between INDSTAT2 and INDSTAT3.

1963-2019 (57 years) and 23 manufacturing sectors. The average number of years per country is 34 years with a minimum of 1 year and a maximum of 57 years.[7] INDSTAT4 includes 143 countries for the period 1990-2018 (29 years) and 127 manufacturing sectors. The average number of years per country is 12 years with a minimum of 1 year and a maximum of 27 years.[8] As a first step, we harmonize the number of sectors for all countries and years by adding the missing sectors and assigning them zero values. The dataset remains unbalanced in terms of countries and years.[9]

There are several advantages to using UNIDO. First, UNIDO is the longest-running dataset with data on cross-country production for the manufacturing sectors, allowing comparability over time and across countries. It also offers different measures of industrial production including employment and value-added. However, the data also suffer from several limitations. UNIDO covers only the manufacturing sector, which means no information on agriculture or services is provided. There is considerable variation in data coverage across countries, years, and sectors. UNIDO doesn't include data on the informal sector which tends to be more important in low-income countries as well as in labor-intensive sectors. Data on value-added is measured in current prices which may pose measurement challenges.

3.2 Exports

Merchandise export data are obtained from the UN COMTRADE database at SITC (Standard international Trade Classification) Rev.1., which provides the longest-running data in UN COMTRADE starting from 1962. The dataset covers 215 countries for 58 years (from 1962 to 2019) over 625 4-digit sectors and 61 2-digit sectors. The average number of years per country is 43 years with a minimum of 1 year and a maximum of 58 years. The dataset remains unbalanced in terms of countries and years.

---

[7] Figure A1 in the appendix shows, for each country, the number of years for which employment data at the 2-digit level (INDSTAT 2) is available. Figure A2 in the appendix shows, for each country, the number of years for which value-added data at the 2-digit level is available.

[8] Figure A3 shows, for each country, the number of years for which employment data at the 4-digit level is available. Figure A4 shows, for each country, the number of years for which value-added data at the 4-digit level is available.

[9] A regression of the number of observations on log GDP per capita indicates that higher income countries have more observations in the dataset, this could be because they have better capabilities in data measuring and data reporting. Controlling for income, resource endowment is not correlated with data coverage. Figure A5 in appendix shows the correlation between income per capita, resource endowment, and data availability in UNIDO.

Services export data are obtained from WTO statistical portal[10]. Two datasets are available, the first provides data on services for the period 1980 to 2013. The dataset covers 22 sectors using EBOPS 2002 classification[11]. The second data covers 94 sectors using EBOPS 2010 classification for the period 2005 to 2021. The two classifications are concorded using Wettstein et al. (2019) and Brochert et al. (2020). The resulting dataset includes 158 countries and 13 sectors for a period of 40 years (1980 – 2019).

3.3 GDP per Capita

The key correlate for the analysis is GDP per capita, which is obtained from Penn World Table. GDP per capita is initially expressed in 2017 PPP dollars and converted to 1985 PPP dollars using US implicit GDP deflator for comparability with Imbs and Wacziarg (2003).[12]

3.4 Measures of Diversification

We can measure diversification using employment, value-added or exports. Each of these dimensions relate to potentially different benefits of diversification. For example, diversification in employment can be important to create productive employment for labor-abundant countries (Gelb, 2010). Additionally, a country can appear diversified in one dimension but not the other. Labor allocation can look very different from value-added or export allocation. Alternatively, a very diversified economy in terms of employment, value-added or exports can be very concentrated in fiscal revenues which can put pressure on the government budget if the sector responsible for the bulk of fiscal revenues experiences an economic shock.

To measure the level of diversification/concentration of an economy, we use the Gini and the Herfindahl (HHI) indices.

[10] https://stats.wto.org/

[11] EBOPS (Extended Balance of Payments Services Classification) is a nomenclature system for measuring service transactions between residents and non-residents created in 1993. The classification currently used (EBOPS 2010) is based on the *6th edition of the Balance of Payments and International Investment Position Manual*.

[12] We convert GDP per capita in 2017$ to GDP per capita in 1985$ by dividing GDP per capita in 2017$ by [US implicit GDP deflator in 2017/US implicit GDP deflator in 1985].

The Gini index is measured using the following formula:

$$\mathrm{Gini} = 1 - \frac{1}{\mathrm{N}}\sum_{\mathrm{k=1}}^{\mathrm{n}}(\mathrm{X_k} + \mathrm{X_{k-1}})$$

Where $X_k = \sum_{l=1}^{k} s_l$represents the cumulative sector shares and $s_k = \frac{x_k}{\sum_{k=1}^{N} x_k}$ is the share of sector k in total employment/value-added and N is the total number of sectors (omitting country and time subscripts).

The Gini coefficient ranges between 0 and 1. The closer the Gini coefficient is to 1, the more unequal the distribution is. For example, a Gini of zero points to a homogeneously diversified economy where all sectors have an equal share of total employment/value-added. On the other hand, a Gini of 1 means the economy is fully concentrated and all employment/value-added is generated by a single sector. The Gini index is a measure of inequality of a distribution. It measures the difference between the actual distribution of unemployment and the equal distribution of employment across sectors. s. It is therefore affected by adding sectors with zero values and it increases as the number of sectors increases.

The Herfindahl index is measured as follows:

$$\mathrm{HHI} = \frac{\sum_{\mathrm{k=1}}^{\mathrm{n}} \mathrm{s_k^2} \; - \; \frac{1}{\mathrm{N}}}{1 - \; \frac{1}{\mathrm{N}}}$$

Where $s_k = \frac{x_k}{\sum_{k=1}^{n} x_k}$ is the share of sector k in total employment/value-added and N is the total number of sectors (omitting country and time subscripts).

The Herfindahl index is normalized to range between 0 and 1. The closer the index is to 1, the more unequal the distribution. HHI is often employed as a measure of market concentration or market power. It is unaffected by the number of sectors with zero values. HHI decreases as the number of non-zero sectors increases (less market concentration/power) and increases with dispersion in size between sectors (more market concentration/power). HHI is more affected by the sizes of sectors than by the number of sectors because squaring the shares gives larger weight to larger sectors. This means that while the Gini index would be the same if a market has 5 sectors with equal shares or 100 sectors with equal shares as it focuses on equality of a distribution, HHI would be higher for a market with 5 sectors rather than 100 sectors because in the first case there is more market concentration/power.

Other measures of concentration include coefficient of variation, max min spread, log variance of sector shares and the Theil index. Table A5 in appendix shows summary statistics on Gini and HHI.

3.5 Defining Resource Abundance

An important distinction is how to define resource abundance. First, it is necessary to define what are natural resources. Here we employ a narrow definition that includes fuel (which comprises SITC section 3: mineral fuels, lubricants, and related materials) and ores and metals (which comprise SITC sections 27: crude fertilizer, minerals n.e.s (not elsewhere specified); 28: metalliferous ores, scrap; and 68: non-ferrous metals). A broader definition would include natural resource-intensive commodities as defined by Leamer (1995) which comprise SITC sections 0-9, 11, 12, 21-29, 32-35, 41-43, 63, 64, 68, 94. The broad definition includes agriculture, forestry, animal products in addition to fuel, ores, and metals.

Consistent with Lederman and Maloney (2003), we define a country to be resource abundant if it's, on average, a net exporter of natural resources. For a more a nuanced measure of resource abundance we use the ratio between natural resources net exports and working-age population to capture a country's comparative advantage in natural resources vis a vis labor in the spirit of Heckscher-Ohlin. On this basis, resource-rich countries can be divided into two groups: highly resource-rich are countries that are above median net exporters of natural resources per worker and moderately resource-rich are countries that are below median net exporters of natural resources per worker.

Resource-rich countries can also be divided based on the type of natural resource endowments into fuel-rich countries which are net exporters of fuel (which comprise SITC section 3: mineral fuels, lubricants, and related materials) and metal-rich countries which are net exporters of metals and ores (which comprise SITC sections 27: crude fertilizer, minerals nes; 28: metalliferous ores, scrap; and 68: non-ferrous metals).

As a result, our sample includes 151 countries, with 100 resource-poor countries and 51 resource-rich countries (24 highly resource-rich and 26 moderately resource-rich countries). Highly resource-rich countries are richer with an average GDP per capita of 14,471.19 (constant 1985$) while GDP per capita of resource-poor is 6,838 (constant 1985$) and moderately resource-rich countries are poorest with an average GDP per capita of 2,600.98 (constant 1985$).[13] Moderately resource-rich countries are much poorer than highly resource-rich countries and they are even poorer than resource-poor countries.

4 Results

4.1 Non-Parametric "Lowess" Estimation

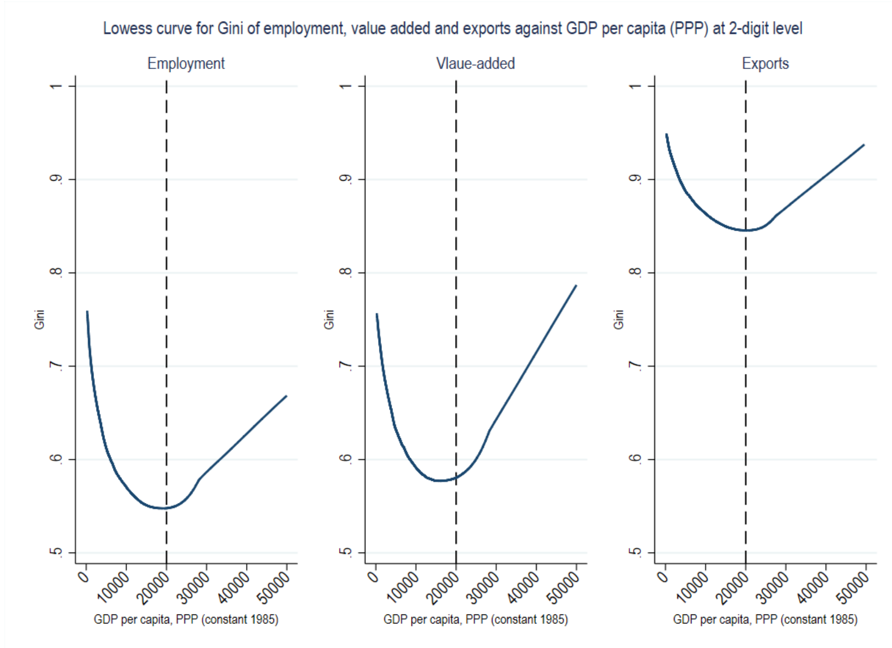
We estimate non-parametric *lowess* curves for different measures of concentration using employment, value-added, and exports concentration measures. *Lowess* refers to locally weighted scatter plot smoothing where the smoothed values are obtained by running a weighted OLS regression of the dependent variable (here the Gini) on the independent variable (here GDP per capita) for each observation and a few of the data points around it.

The first two panels in Figure 22 display the *lowess* curves for employment and value-added. We use the UNIDO INDSTAT2 database which includes 23 manufacturing sectors based on ISIC Rev.3 classification over the period (1962-2019). The third panel displays the lowess curve for exports. We use the UN COMTRADE database for the manufacturing sector which includes 55 sectors based on SITC Rev. 1 classification over the period (1962-2019).[14]

[13] See Figure A6 in appendix for a list of resource-rich and resource-poor countries. See Table A6 for summary statistics on income differences between resource-rich and resource poor countries.

[14] We use concordance tables to convert SITC Rev.1 to ISIC Rec 3.1 for better comparison with UNIDO data.

Figure 2: Lowess Curve for the Gini index at 2-digit level



Source: Authors' calculations

The Gini index follows a U-shaped pattern: Countries first diversify then they start specializing at around PPP GDP per capita of 20,000 (constant 1985 US dollars). This result is in line with the literature on diversification and economic development (Imbs and Wacziarg, 2003; Klinger and Lederman, 2004 and 2006; and Cadot et al.,2011). Imbs and Wacziarg (2003) find a turning point at PPP GDP per capita 8,675 (constant 1985 US dollars) and they note that their turning point occurs around the level of income for Ireland in 1992. Since we use data on PPP GDP per capita in constant 2017 US$, the value of the turning point in our estimation is not directly comparable. We thus convert PPP GDP per capita (constant 2017 US dollars) into constant 1985 dollars to compare with Imbs and Wacziarg (2003) using the U.S GDP implicit price deflator provided by the U.S Bureau of Economic Analysis who use GDP per capita in constant 1985$. Our turning point corresponds roughly to Ireland's income level in 2000. Countries to the left of the turning point include Ghana, Malaysia, Egypt, Nigeria. Countries to the right of the turning point include: Saudi Arabia (1974-1981, 2008-2019), Qatar (1978-1979, 1999-2004), Kuwait (1974-1984, 1993-2019), UAE, Japan (2005-2019), UK (2006-2019), New Zealand (2017-2019), Australia (2001-2019), Ireland (2000-2019), France (2007-2019), Canada (1999-2019), USA (1992-2019).

A significant difference with Imbs and Wacziarg (2003) is that they find that the curve is not symmetric, and the upward sloping part of the curve does not go back to the initial level of concentration. Instead, we find that the upward sloping part of the curve is steep and can even exceed the initial level of concentration.

Measures of concentration using exports data are on average higher than those using employment and value-added data.[15] [16]

Using three different data sources (employment, value-added and exports) and 2 different measures of economic concentration (Gini and HHI) [17] gives similar U-shaped patterns with very close turning points. This reinforces the results in the literature that there are 2 stages of diversification in manufacturing.

We observe that the granularity of sectoral disaggregation does not affect the *shape* of the curve, but it affects the *level* of the curve, especially for the Gini index. The higher the level of disaggregation, the higher the Gini index is because there are more sectors with zero values and therefore more inequality in distribution. For example, exports are classified across 625 sectors, yet the mean (median) country exports only in 354 (378) sectors. HHI is also affected by the higher level of disaggregation but to a smaller extent. The higher the level of disaggregation, the lower the HHI because there is less market power/market concentration.[18]

We also investigate patterns of diversification for trade in services. Figure 3 shows the Gini index for services exports. The lowess curve has a U-shaped curve which is in line with the literature. Countries first diversify then they start specializing at around PPP GDP per

[15] This could be because a country's exports are more concentrated than production because it only exports in sectors where it has a comparative advantage. Another reason could be because exports classification has more sectors than production which would inflate the value of the Gini index.

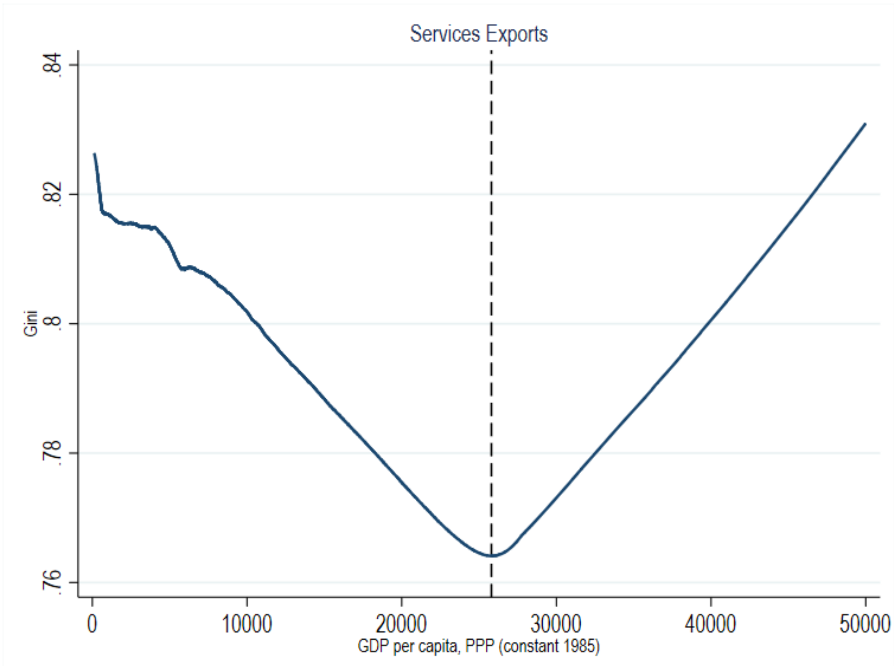
[16] Lowess curve for exports using all sectors in COMTRADE data instead of only the manufacturing sector is qualitatively similar to panel 3 of Figure 2. Results are available upon request.

[17] Figure A7 in the appendix uses HHI instead of Gini and finds similar results.

[18] Figure A8 and Figure A9 in the appendix present the lowess curves for different levels of disaggregation.

capita of 25,000 (constant 1985 US dollars). [19] Re-specialization in export services occurs at a higher level of GDP per capita compared with merchandise exports. However, the data is classified into only 13 sectors which is more aggregated than the other datasets used.

Figure 3: Lowess Curve for the Gini index for services exports



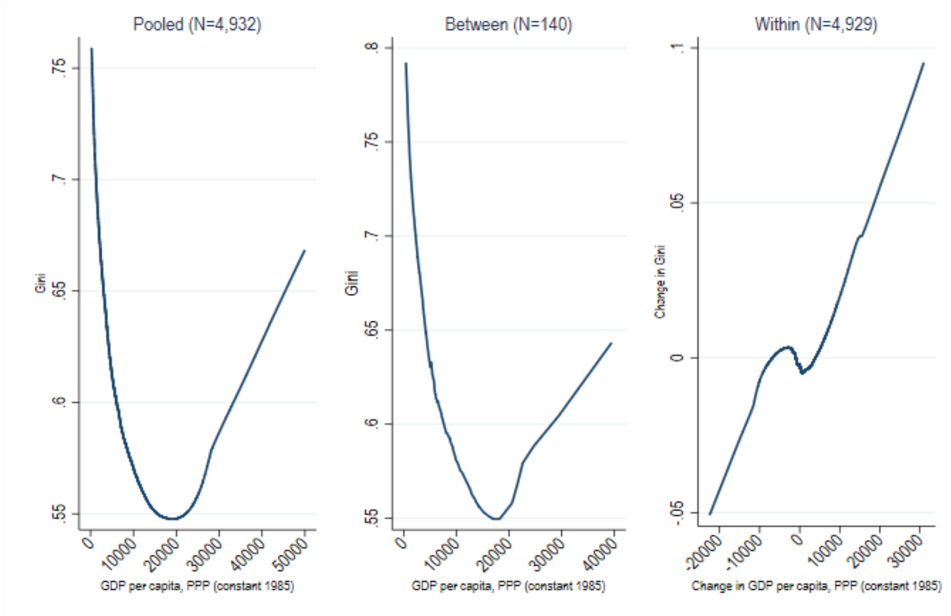
Source: Authors' calculations

4.2 Breaking-down sources of variation for pooled, between and within relationships using employment, value-added and exports

In this section, we try to understand what is driving the U-shaped pattern of the *lowess* curve and whether it is due to differences between-country or within-country. Figure 4 shows non-parametric *lowess* curves for the pooled, between and within-relationship between employment and GDP per capita.

[19] The lowess curve has a U-shaped curve for all economies. Looking at resource-rich and resource-poor countries separately, it seems that the downward section of the U shaped is driven by diversification resource-poor countries and the upward section is driven by concentration in resource-rich countries. Results are available upon request.

Figure 4: Fixed- and within-effects estimation



Source: Authors' calculations

The first panel of Figure 4 displays the same *lowess* curve as Figure 2: the non-parametric *lowess* curve for the Gini index on GDP per capita, on pooled data. The smoothed values of the Gini index are obtained from local weighted regressions of the Gini index on GDP per capita according to the following equation: $y_{it} = \alpha + \ \beta x_{it} + \epsilon_{it}$. The *lowess* curve has a U-shaped pattern showing two stages of diversification.

The second panel displays the non-parametric lowess curve of the average Gini index on average GDP per capita. The smoothed values of the Gini index are obtained from local weighted regressions of the country-average Gini index on the country-average GDP per capita according to the following equation: $\overline{y}_i = \alpha^{BE} + \ \beta^{BE}\, \overline{x_i} + \overline{\varepsilon_i}$. The *lowess* curve displays again a U-shaped pattern, this characterizing the between-country relationship between Gini and GDP per capita.

The third panel displays the non-parametric *lowess* curve of the de-meaned Gini index on the de-meaned GDP per capita. The smoothed values of the Gini index are obtained from local weighted regressions of deviation from the country-average of the Gini index on deviation from the country-average of the GDP per capita according to the following equation: $(y_{it} - \overline{y}_{it}) = \beta^{FE} \ \ x_{it} - \overline{x_i} \ + \ \varepsilon_{it} - \overline{\varepsilon}$. The *lowess* curve does not show a U-shaped pattern. Instead, we see an upward linear relationship between the Gini index and GDP per capita. Countries concentrate as they grow richer. The same results largely hold for resource rich and resource poor countries. See figures Figure A 12 and Figure A 13 in appendix.

Taken together, these results suggest that the change in the patterns of economic diversification (U-shaped pattern) across the development path is mainly driven by between-country rather than within-country variation. Results are similar if we use a different data source (value-added or exports) or different measures of concentration (Gini vs HHI).[20]

4.3 Quadratic Estimation

In this section, we compare within country and between country variation. Table 1 shows a set of regressions of the Gini index on income and income squared with country and year fixed effects. We find that evidence for a quadratic relationship doesn't hold. Results hold for resource-poor countries but the magnitude of the coefficients for income and income squared are very small. Results don't hold for resource rich countries. Table 2 presents a set of regressions of average Gini on average income and income squared to capture cross-country effects. The quadratic relationship holds in a cross-country setting.

Table 1: Within Regressions for pooled, Resource-Rich, and Resource-Poor

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	**All Linear**	**All Quadratic**	**Resource-Poor Linear**	**Resource-Poor Quadratic**	**Resource-Rich Linear**	**Resource-Rich Quadratic**
GDP per capita	0.00071***	0.00054***	0.00122***	-0.00584***	0.00037**	-0.00017
	(0.00013)	(0.00021)	(0.00020)	(0.00052)	(0.00017)	(0.00036)
GDP per capita square		0.00000		0.00022***		0.00000*

[20] See Figure A9 and Figure A10 in appendix for results for value-added and exports. Figures using HHI display similar results and can be available upon request.

		(0.00000)		(0.00002)		(0.00000)
Constant	0.60639***	0.60740***	0.59386***	0.62422***	0.62756***	0.63086***
	(0.00125)	(0.00160)	(0.00176)	(0.00270)	(0.00195)	(0.00277)
Observations	4,970	4,970	3,393	3,393	1,577	1,577
R-squared	0.00648	0.00668	0.01165	0.07075	0.00306	0.00489
Number of iso3c1	140	140	93	93	47	47

Table 2: Between Regressions for Pooled, Resource-Rich, and Resource-Poor

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	All Linear	All Quadratic	Resource-Poor Linear	Resource-Poor Quadratic	Resource-Rich Linear	Resource-Rich Quadratic
GDP per capita	-0.00434***	-0.01185***	-0.00670***	-0.01521***	-0.00232	-0.01754
	(0.00115)	(0.00197)	(0.00178)	(0.00499)	(0.00294)	(0.01466)
GDP per capita square		0.00019***		0.00036*		0.00043
		(0.00004)		(0.00020)		(0.00041)
Constant	-3.66375*	-2.77362	-7.74183***	-6.66882**	-0.10016	0.05693
	(1.97675)	(1.79207)	(2.53702)	(2.53674)	(6.26328)	(6.21142)
Observations	4,970	4,970	3,393	3,393	1,577	1,577
R-squared	0.71510	0.77155	0.88303	0.89216	0.88934	0.90677
Number of iso3c1	140	140	93	93	47	47

4.4 Revisiting Imbs and Wacziarg (2003) using Updated Data

This section replicates the semi-parametric method developed by Imbs and Wacziarg (2003) on our updated data including 151 countries and 57 years. To implement it, we partition the data into S smaller subsamples based on GDP per capita of size J and overlap of J-Δ.[21] We use S=796, J=10,000$, Δ=50, N=4,929.[22,23] Our next step is to run a fixed-effects linear regression (xtreg) for each subsample and save the coefficients and standard errors for the intercept and GDP per capita for each subsample. Then, we predict the Gini index at the midpoint of each subsample using the coefficients of the fixed-effects regression where s= 1,..,S and alpha and beta are, respectively, the fixed effects estimates of the intercept and the slope on GDP per capita in a fixed-effects linear regression of the Gini index on GDP per capita for the subsample s. Additionally, we estimate the 95% confidence interval by calculating the standard error of the predicted Gini and plotting the bands around the fitted curve. Finally, we plot the predicted Gini index against the midpoint GDP per capita of each subsample along with the 95% confidence interval.

There are some differences between the lowess, and the semi-parametric technique described above. *Lowess* performs a simple OLS regression of y on x while the semi-parametric technique performs a fixed effects regression of y on x using country fixed effects. *Lowess* is more computationally intensive, performing N regressions for N observations. The semi-parametric technique reduces the number of regressions by using increments of 25$. *Lowess* typically uses 80% of observations for each regression. The semi-parametric technique uses an overlap of 5,000$ (30% of observations on average). *Lowess* performs a weighted regression for each observation so that points further away from the observation receive less

[21] We try different values for S, J and Δ. Our results are not sensitive to these choices. Results are available upon request.

[22] The average number of observations per subsample is n = 649 (ranging from 29 to 3,585 observations). The average value of GDP per capita (constant international 1985 $) is 10,542 $, minimum value is 233 $ and maximum value is 151,826 $. Only the last 36 observations have values higher than 50,000$. Subsamples are calculated up to 50,000$. This significantly reduces the number of subsamples by 40%.

[23] Imbs and Wacziarg (2003) use S=489, J=5,000$, Δ =25$, N=1,556. The average number of observations per subsample is n = 388 (ranging from 108 to 963 observations). The average value in GDP per capita (constant international 1985 $) is 3,935 $, minimum value is 268 $ and maximum value is 33,946 $. Only the last 19 observations have values higher than 17,500 $. Subsamples are calculated until 17,500$.

weight. The semi-parametric technique uses equal weighing of observations in each subsample.

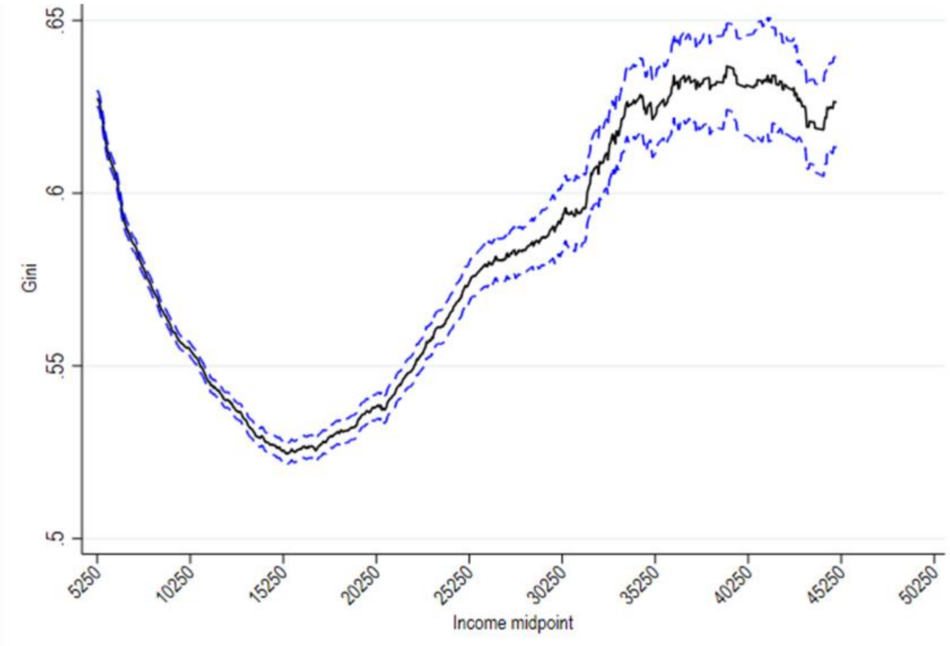
The curve estimated with the semi-parametric method has a turning point of GDP per capita PPP 15,000 (constant 1985$). We include country-fixed effects. Therefore, each country has its own intercept which captures country-specific, time-invariant factors affecting the Gini index. The fitted value of the Gini index is computed using the average value of the intercept plus the slope coefficient multiplied by the midpoint of each subsample. The average value of the intercept is the average of the individual country fixed effect for the countries present in each subsample. Imbs and Wacziarg (2003) explain that changes in the fitted value of the Gini index could be due to changes in the average value of the intercept or the slope coefficient or the GDP per capita midpoint. They also explain that the main source of variation in the estimated curve is due to changes in the value of the intercept across subsamples and the smoothed curve would look very similar if the slope estimate * midpoint was removed from the fitted values.

Figure 5 presents a U-shaped curve which shows that variation in the estimated non-parametric curve is mainly driven by changes in the intercept which measures the average degree of diversification across subsamples rather than changes in the slope coefficient which measures the relationship between economic concentration level of development.[24]

Figure 6 plots the slope coefficients of the fixed-effects regression of each subsample against its midpoint GDP per capita. The downward sloping blue line shows the number of observations per subsample, indicating that, similarly to Imbs and Wacziarg (2003), there are very few observations at the right tail end of the sample even in the 151-country dataset. When the number of observations drops, so does the precision of the estimates, which is highlighted in bold.

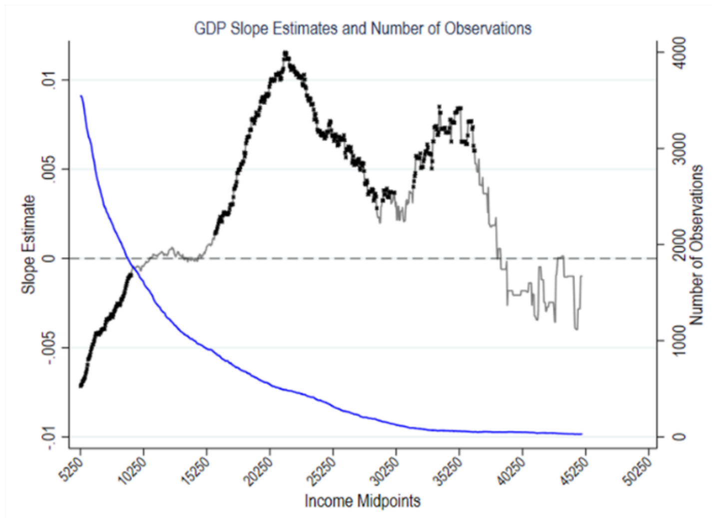
[24] The results for value-added are similar to those for employment. The figures are available upon request.

Figure 5: Estimated Semi-parametric Curve for UNIDO 2-digit Employment Data



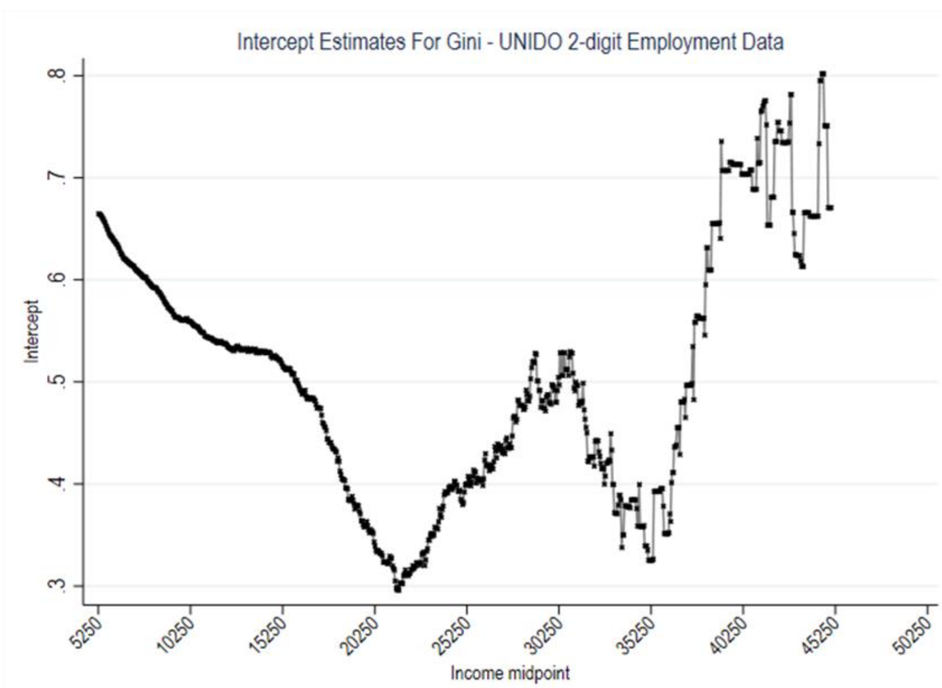
Source: Authors' calculation

Figure 6: Slope coefficients and their significance in the Fixed-effect Regression with the semi-parametric Method and Updated Data



Source: Authors' calculations

Figure 7: Intercept Coefficients and their Significance in the Fixed-effect Regression with the semi-parametric Method and Updated Data



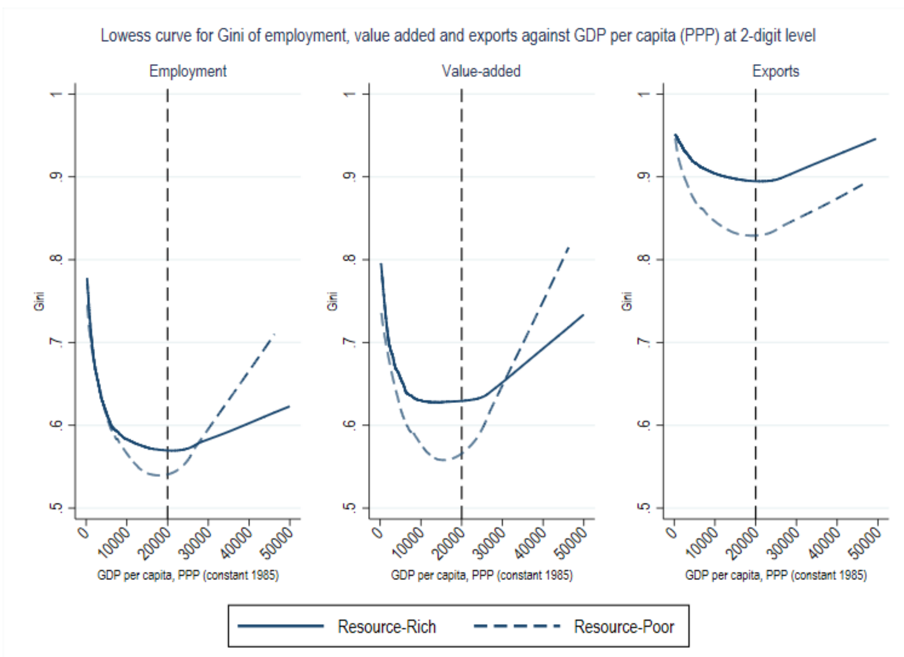
Source: Authors' calculations

Similarly, Figure 7 plots the intercept of the semi-parametric fixed-effects regression of each subsample against its midpoint GDP per capita. Consistent with our earlier evidence that between-effects account for much the U-shaped relationship between diversification and GDP per capita, the estimated intercept curve is U-shaped and the intercept is estimated with consistent precision.

4.5 Patterns of diversification for Resource-Poor, Resource-Rich, and Highly Resource-Rich Countries.

Figure 8 displays the *lowess* curve for the Gini index of employment, value-added and exports for resource-rich and resource-poor countries separately.

Figure 8: Diversification patterns in resource-rich and resource-poor countries



Source: Authors' calculations

In resource-poor countries, the Gini index follows a similar pattern to Figure 2. Countries first diversify then they start specializing at a turning point around PPP GDP per capita of 20,000 (constant 1985 US dollars). Instead, resource-rich countries reach a plateau at around PPP GDP per capita of 10,000 (constant 1985 US dollars) and start specializing again later at a higher-turning point. Resource-rich countries have a higher Gini index on average than resource-poor countries and therefore look more concentrated. This observation is most noticeable in the third panel of Figure 8. After the turning point, re-specialization in resource-poor countries is higher in employment and value-added than in resource-rich countries. While re-concentration is higher for resource-rich countries' exports.[25]

The upward bending part of the lowess curve in the 3 panels is steeper for resource-poor countries than for resource-rich countries suggesting that richer resource-poor countries concentrate more than richer resource-rich countries. On average, resource-rich countries are on average more specialized both in employment and value-added compared to resource-poor countries. The difference between the mean of the two groups is statistically significant.

5. Conclusion

This paper investigates the diversification of economies over the past 60 years with a special focus on trajectories of resource-rich and resource-poor countries across three dimensions: employment, value-added, and exports. We estimate non-parametrically a U-shaped curve between measures of economic concentration and per capita income levels. The evidence suggests that these patterns are driven by between-country rather than within-country variation.

Diversification patterns also differ across resource-rich and resource-poor countries. Employment and value added in resource-rich countries are on average more concentrated at low levels of development than resource-poor ones, but the relationship is opposite at high levels of development, with resource poor countries displaying higher concentration. At all levels of development exports are more concentrated in resource-rich countries.

The aforementioned results come from bi-variate estimations, thus leaving broad room for future research about the drivers of diversification across stages of development. Given that we know that the U-shaped relationship is due to the between country variation, it is worth looking at potential confounding factors with little over-time variation. Thus, this paper has taken a first stab by focusing on the role of natural resource abundance, which tends to

[25] Figures using value-added and exports display similar results and can be available upon request.

vary little over time across countries. But there are other potential relatively time-invariant factors that might be correlated with the distribution of GDP per capita across countries. On the other hand, there are lingering measurement issues that could explain the observed patterns in the data.

First, the reported cross-country differences could be due to the exclusion of primary and tertiary sectors from our analysis because the global datasets from UNIDO primary sectors or services. Likewise, the data on merchandise exports does not include services trade. We have taken a first step by looking at data on services exports. However, the available data is still at a highly aggregated level which make concordance with other datasets on the manufacturing sector at a more disaggregated level difficult. Another potential explanation concerns the role of economic size as a determinant of economic diversification, whereby large economies tend to be more diversified than small economies (Lederman and Maloney 2012; Lederman et al. 2021). Thus, it is plausible that if the sample of rich countries that appear beyond the hump of diversification might be smaller than the sample of countries. These are further avenues of research that we are exploring as part of this project.

References

Acemoglu, A, Ziljbotti, F. (1997). “Was Prometheus Unbound by Chance? Risk, Diversification, and Growth.” *The Journal of Political Economy*, 105(4), 709-751.

Alexeev, M., & Conrad, R. (2009). “The elusive curse of oil”. *The review of Economics and Statistics*, 91(3), 586-598.

Borchert, I., Larch, M., Shikher, S., & Yotov, Y. V. (2021). “The international trade and production database for estimation (ITPD-E).” *International Economics*, 166, 140-166.

Bruno, M., & Sachs, J. (1982). “Energy and resource allocation: a dynamic model of the “Dutch Disease”.” *The Review of Economic Studies*, 49(5), 845-859.

Bulte, E. H., Damania, R., & Deacon, R. T. (2005). “Resource intensity, institutions, and development.” *World development*, 33(7), 1029-1044.

Busse, M., & Gröning, S. (2013). “The resource curse revisited: governance and natural resources.” *Public choice*, 154(1), 1-20.

Cadot, O., Carrère, C., & Strauss-Kahn, V. (2011). “Export diversification: What’s behind the hump?” *Review of Economics and Statistics*, 93(2), 590-605.

Corden, W. M., & Neary, J. P. (1982). “Booming sector and de-industrialisation in a small open economy.” *The economic journal*, 92(368), 825-848.

Dornbusch, R., Fischer, S., & Samuelson, P. A. (1977). “Comparative advantage, trade, and payments in a Ricardian model with a continuum of goods.” *The American Economic Review*, 67(5), 823-839.

Gelb, A. H. (1988). “Oil windfalls: Blessing or curse?” *Oxford university press.*

Ghura, D., & Pattillo, C. (2012). “Macroeconomic policy frameworks for resource-rich developing countries.” *International Monetary Fund*, Washington, DC.

Gylfason, T. (2001). “Natural resources, education, and economic development.” *European economic review,* 45(4-6), 847-859.

Herrendorf, B., Rogerson, R., & Valentinyi, A. (2014). “Growth and structural transformation.” *Handbook of economic growth*, 2, 855-941.

Hesse, H. (2009). "Export diversification and economic growth." Breaking into new markets: Emerging lessons for export diversification, 2009, 55-80.

Imbs, J., & Wacziarg, R. (2003). "Stages of diversification." *American economic review*, 93(1), 63-86.

Klinger, B., & Lederman, D. (2004). "Discovery and development: an empirical exploration of" new" products" (Vol. 3450). World Bank Publications.

Klinger, B., & Lederman, D. (2006). "Diversification, innovation, and imitation inside the global technological frontier." World Bank policy research working paper, (3872).

Ross, M. L. (2017). "What do we know about economic diversification in oil-producing countries?" Available at SSRN 3048585.

Koren, M., & Tenreyro, S. (2007). "Volatility and development." *The Quarterly Journal of Economics*, 122(1), 243-287.

Lashitew, A. A., Ross, M. L., & Werker, E. (2021). "What drives successful economic diversification in resource-rich countries?" *The World Bank Research Observer*, 36(2), 164-196.

Leamer, E. E. (1995). "The Heckscher-Ohlin model in theory and practice."

Lederman, D., & Maloney, W. (2002). "Open questions about the link between natural resources and economic growth: Sachs and Warner revisited (Vol. 141)." Banco Central de Chile.

Lederman, D., & Maloney, W. F. (2007). "Neither curse nor destiny: introduction to natural resources and development." Natural resources: Neither curse nor destiny, 1-14.

Lederman, D., and Maloney. (2008). "In search of the missing resource curse [with comments]". Economía, 9(1), 1-57.

Lederman, D, and Maloney, W.F. (2012). "Does What You Export Matter? In Search of Empirical Guidance for Industrial Policies." World Bank: Washington, DC.

Lederman, D., Pienknagura, S., & Rojas, D. (2021). "Latent trade diversification and its relevance for macroeconomic stability." *The World Bank Economic Review*, 35(1), 58-91.

Malik, A., & Temple, J. R. (2009). "The geography of output volatility." *Journal of Development Economics*, 90(2), 163-178.

Mehlum, H., Moene, K., & Torvik, R. (2006). "Cursed by resources or institutions?" *World Economy*, 29(8), 1117-1131.

Papageorgiou, M. C., & Spatafora, M. N. (2012). "Economic diversification in LICs: Stylized facts and macroeconomic implications." International Monetary Fund.

Robinson, J. A., Torvik, R., & Verdier, T. (2006). "Political foundations of the resource curse." *Journal of development Economics*, 79(2), 447-468.

Ross, M. L. (2001). "Timber booms and institutional breakdown in Southeast Asia." Cambridge University Press.

Ross, M. (2003). "How does mineral wealth affect the poor." UCLA Department of Political Science.

Sachs, J. D., & Warner, A. (1995). "Natural resource abundance and economic growth."

Sachs, J. D., & Warner, A. M. (2001). "The curse of natural resources." *European economic review*, 45(4-6), 827-838.

Sala-i-Martin, X., & Subramanian, A. (2013). "Addressing the natural resource curse: An illustration from Nigeria." *Journal of African Economies*, 22(4), 570-615.

Schott, P. K. (2003). "One size fits all? Heckscher-Ohlin specialization in global production." *American Economic Review,* 93(3), 686-708.

Schott, P. K. (2004). "Across-product versus within-product specialization in international trade." *The Quarterly Journal of Economics*, 119(2), 647-678.

Tornell, A., & Lane, P. R. (1999). "The voracity effect." *American economic review*, 89(1), 22-46.

Van der Ploeg, F., & Poelhekke, S. (2009). "Volatility and the natural resource curse." Oxford economic papers, 61(4), 727-760.

Venables, A. J. (2016). "Using natural resources for development: why has it proven so difficult?" *Journal of Economic Perspectives*, 30(1), 161-84.

Wettstein, S., Liberatore, A., Magdeleine, J., & Maurer, A. (2019). "A global trade in services data set by sector and by mode of supply (TISMOS)."

Appendix A

Table A1: List of Sectors in INDSTAT 2

Rev. 3 Code	Definition
15	Food and beverages
16	Tobacco products
17	Textiles
18	Wearing apparel, fur
19	Leather, leather products and footwear
20	Wood products (excluding furniture)
21	Paper and paper products
22	Printing and publishing
23	Coke, refined petroleum products, nuclear fuel
24	Chemicals and chemical products
25	Rubber and plastics products
26	Non-metallic mineral products
27	Basic metals
28	fabricated metal products
29	Machinery and equipment n.e.c.
30	Office, accounting, and computing machinery
31	Electrical machinery and apparatus
32	Radio, television, and communication equipment
33	Medical, precision, and optical instruments
34	Motor vehicles, trailers, semi-trailers
35	Other transport equipment
36	Furniture; manufacturing n.e.c.
37	Recycling

Table A2: List of Sectors in INDSTAT 4

Rev. 4 Code	**Definition**	**Rev. 4 Code**	**Definition**
1010	Processing/preserving of meat	2432	Casting of non-ferrous metals
1020	Processing/preserving of fish, etc.	2511	Structural metal products
1030	Processing/preserving of fruit, vegetables	2512	Tanks, reservoirs and containers of metal
1040	Vegetable and animal oils and fats	2513	Steam generators, excl. hot water boilers
1050	Dairy products	2520	Weapons and ammunition
1061	Grain mill products	2591	Forging, pressing, stamping, roll-forming of metal
1062	Starches and starch products	2592	Treatment and coating of metals; machining
1071	Bakery products	2593	Cutlery, hand tools and general hardware
1072	Sugar	2599	Other fabricated metal products n.e.c.
1073	Cocoa, chocolate and sugar confectionery	2610	Electronic components and boards
1074	Macaroni, noodles, couscous, etc.	2620	Computers and peripheral equipment
1075	Prepared meals and dishes	2630	Communication equipment
1079	Other food products n.e.c.	2640	Consumer electronics
1080	Prepared animal feeds	2651	Measuring/testing/navigating equipment,etc.
1101	Distilling, rectifying and blending of spirits	2652	Watches and clocks
1102	Wines	2660	Irradiation/electromedical equipment, etc.
1103	Malt liquors and malt	2670	Optical instruments and photographic equipment

1104	Soft drinks, mineral waters, other bottled waters	2680	Magnetic and optical media
1200	Tobacco products	2710	Electric motors, generators, transformers, etc.
1311	Preparation and spinning of textile fibers	2720	Batteries and accumulators
1312	Weaving of textiles	2731	Fibre optic cables
1313	Finishing of textiles	2732	Other electronic and electric wires and cables
1391	Knitted and crocheted fabrics	2733	Wiring devices
1392	Made-up textile articles, except apparel	2740	Electric lighting equipment
1393	Carpets and rugs	2750	Domestic appliances
1394	Cordage, rope, twine and netting	2790	Other electrical equipment
1399	Other textiles n.e.c.	2811	Engines/turbines, excl. aircraft, vehicle engines
1410	Wearing apparel, except fur apparel	2812	Fluid power equipment
1420	Articles of fur	2813	Other pumps, compressors, taps and valves
1430	Knitted and crocheted apparel	2814	Bearings, gears, gearing and driving elements
1511	Tanning/dressing of leather; dressing of fur	2815	Ovens, furnaces and furnace burners
1512	Luggage, handbags, etc.; saddlery/harness	2816	Lifting and handling equipment
1520	Footwear	2817	Office machinery, excl. computers, etc.
1610	Sawmilling and planing of wood	2818	Power-driven hand tools
1621	Veneer sheets and wood-based panels	2819	Other general-purpose machinery

1622	Builders' carpentry and joinery	2821	Agricultural and forestry machinery
1623	Wooden containers	2822	Metal-forming machinery and machine tools
1629	Other wood products; articles of cork, straw	2823	Machinery for metallurgy
1701	Pulp, paper and paperboard	2824	Mining, quarrying and construction machinery
1702	Corrugated paper and paperboard	2825	Food/beverage/tobacco processing machinery
1709	Other articles of paper and paperboard	2826	Textile/apparel/leather production machinery
1811	Printing	2829	Other special-purpose machinery
1812	Service activities related to printing	2910	Motor vehicles
1820	Reproduction of recorded media	2920	Automobile bodies, trailers and semi-trailers
1910	Coke oven products	2930	Parts and accessories for motor vehicles
1920	Refined petroleum products	3011	Building of ships and floating structures
2011	Basic chemicals	3012	Building of pleasure and sporting boats
2012	Fertilizers and nitrogen compounds	3020	Railway locomotives and rolling stock
2013	Plastics and synthetic rubber in primary forms	3030	Air and spacecraft and related machinery
2021	Pesticides and other agrochemical products	3040	Military fighting vehicles

2022	Paints, varnishes; printing ink and mastics	3091	Motorcycles
2023	Soap, cleaning and cosmetic preparations	3092	Bicycles and invalid carriages
2029	Other chemical products n.e.c.	3099	Other transport equipment n.e.c.
2030	Man-made fibres	3100	Furniture
2100	Pharmaceuticals, medicinal chemicals, etc.	3211	Jewellery and related articles
2211	Rubber tires and tubes	3212	Imitation jewellery and related articles
2219	Other rubber products	3220	Musical instruments
2220	Plastics products	3230	Sports goods
2310	Glass and glass products	3240	Games and toys
2391	Refractory products	3250	Medical and dental instruments and supplies
2392	Clay building materials	3290	Other manufacturing n.e.c.
2393	Other porcelain and ceramic products	3311	Repair of fabricated metal products
2394	Cement, lime and plaster	3312	Repair of machinery
2395	Articles of concrete, cement and plaster	3313	Repair of electronic and optical equipment
2396	Cutting, shaping and finishing of stone	3314	Repair of electrical equipment
2399	Other non-metallic mineral products n.e.c.	3315	Repair of transport equip., excl. motor vehicles
2410	Basic iron and steel	3319	Repair of other equipment
2420	Basic precious and other non-ferrous metals	3320	Installation of industrial machinery/equipment
2431	Casting of iron and steel		

Table A3: List of Sectors Included in INDSTAT 3

Rev. 2 Code	Definition
311	Food products
323	Leather products
324	Footwear, except rubber or plastic
331	Wood products, except furniture
332	Furniture, except metal
341	Paper and products
342	Printing and publishing
351	Industrial chemicals
352	Other chemicals
353	Petroleum refineries
354	Miscellaneous petroleum and coal products
355	Rubber products
356	Plastic products
361	Pottery, China, earthenware
362	Glass and products
369	Other non-metallic mineral products
371	Iron and steel
372	Non-ferrous metals
381	Fabricated metal products
382	Machinery except electrical
383	Machinery, electric
384	Transport equipment
385	Professional and scientific equipment
390	Other manufactured products

Table A4: Crude Conversion Between INDSTAT2 and INDSTAT3

Rev. 3 code	**Rev. 2 code**
15	311+312+313
16	314
17	321
18A (18+19)	322+323+324
20	331
21	341
22	342
23A (23+24)	351+352+353+354
25	355+356
26	361+362+369
27	371+372
28	381
29C (29+30)	382
31A (31+32)	383
33	385
34A (34+35)	384
36	332+390

Table A5: Comparison between Values of Gini and HHI And Values in Imbs and Wacziarg (2003)

	Obs.	Mean	Standard Deviation	**Minimum**	**Maximum**	
Employment at 2-digit	5333	1763226	5913544	600	85200000	Authors' calculations based on INDSTAT2
Gini for employment at 2-digit	5333	0.62	0.11	0.35	0.91	
HHI for employment at 2-digit	5333	0.10	0.09	0.02	0.77	
For comparison						
Gini for employment from Imbs and Wacziarg (2003)	1556	**0.57**	**0.10**			Imbs and Wacziarg (2003)
HHI for employment from Imbs and Wacziarg (2003)	1556	**0.12**	**0.10**			Imbs and Wacziarg (2003)
Value added at 2-digit	4906	57400000000	235000000000	3140202	3560000000000	Authors' calculations based on INDSTAT2
Gini for value added at 2-digit	4906	**0.63**	**0.12**	0.36	0.96	
HHI for value added at 2-digit	4906	**0.12**	**0.11**	0.02	1.00	
For comparison						
Gini for value-added from Imbs and Wacziarg (2003)	1493	**0.57**	**0.11**			Imbs and Wacziarg (2003)
HHI for value added from Imbs and Wacziarg (2003)	1493	**0.11**	**0.07**			Imbs and Wacziarg (2003)

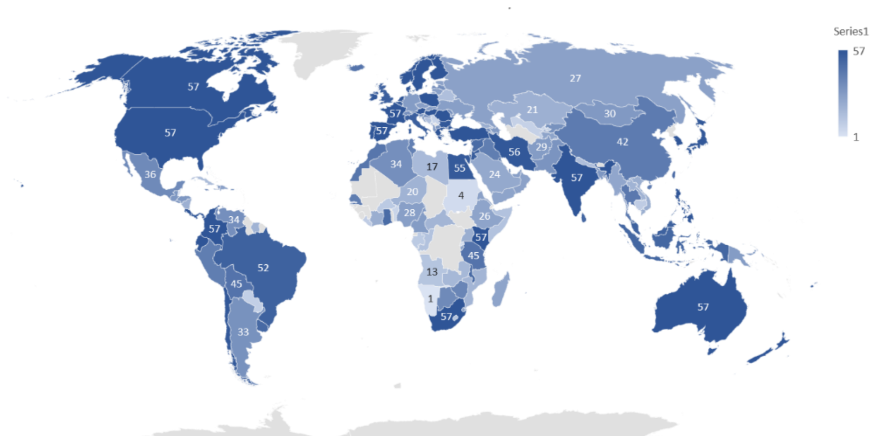
Table A6: Income Distribution of Countries According to their Resource Endowment

PPP GDP per Capita (constant 1985$)	**Mean**	**Median**	**Minimum**	**Maximum**
Resource-Poor (N=100)	6838.003	6023.79	375.7068	27286.78
Resource-Rich (N=51)	8340.347	12773.2	564.1001	72442.78
Moderately Resource-Rich (N=26)	2699.976	2014.04	564.1001	7592.511
Highly Resource-Rich (N=24)	14471.19	16391.5	1801.431	72442.78

Table A7: Summary Statistics for Resource-Rich and Resource-Poor Countries

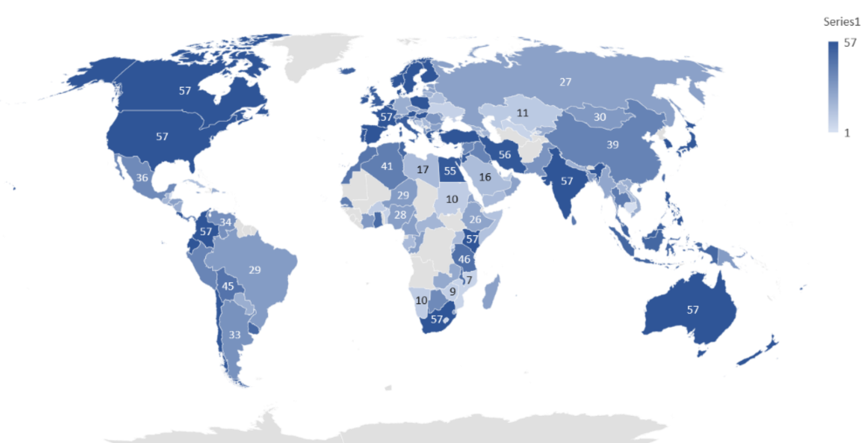
		Mean	**Std. dev.**	**Min**	**Max**
Resource-Poor (N=101)	Number of years per country	37.60	15.76	6	57
	Average Gini of employment	0.61	0.11	0.36	0.91
	Average Gini of value-added	0.61	0.12	0.36	0.96
Resource-Rich (N=50)	Number of years per country	34.30	15.67	5	57
	Average Gini of employment	0.63	0.10	0.35	0.90
	Average Gini of value-added	0.67	0.11	0.41	0.93

Figure A1: Coverage of Employment at 2-digit level (INDISTAT 2) per country for the period (1963-2019)



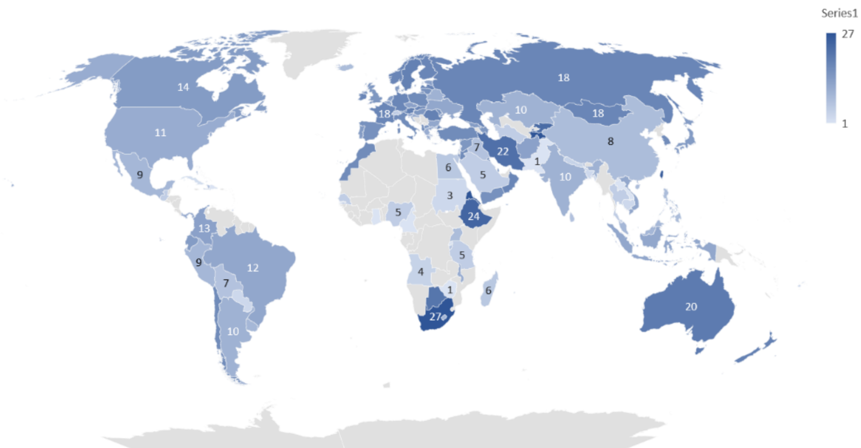
Source: Authors' calculations

Figure A2: Coverage for Value-added at 2-digit level (INDISTAT 2) per country for the period (1963-2019)



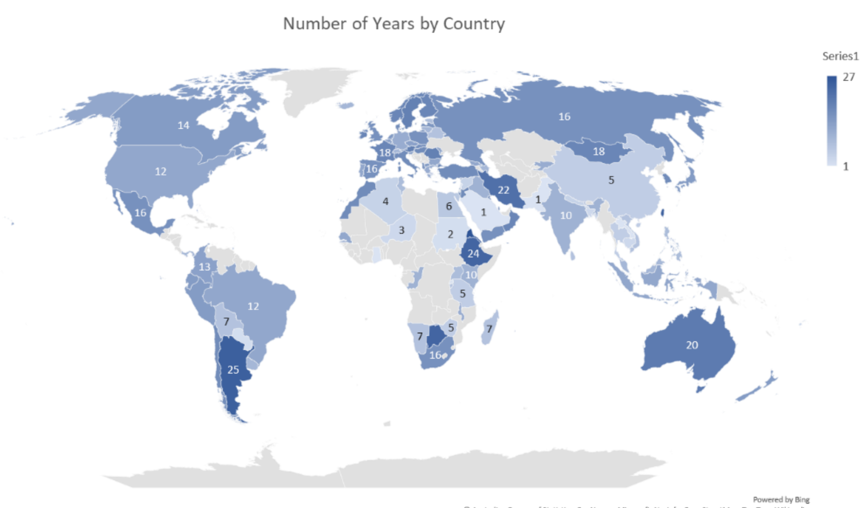
Source: Authors' calculations

Figure A3: Coverage for Employment at 4-digit level per country for the period (1990-2019)



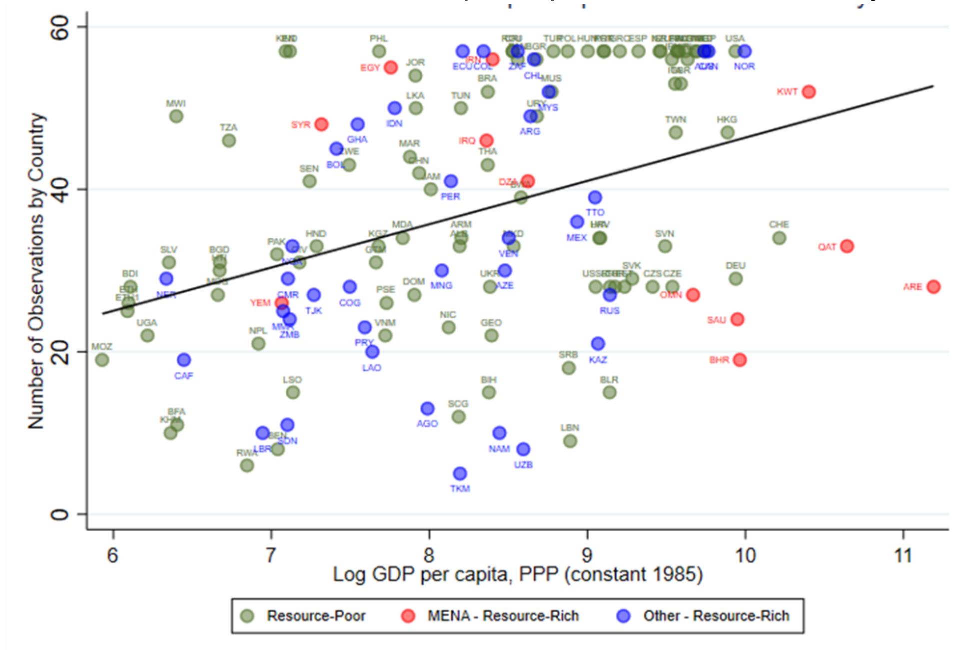
Source: Authors' calculations

Figure A4: Coverage for Value-added at 4-digit level per country for the period (1990-2019)



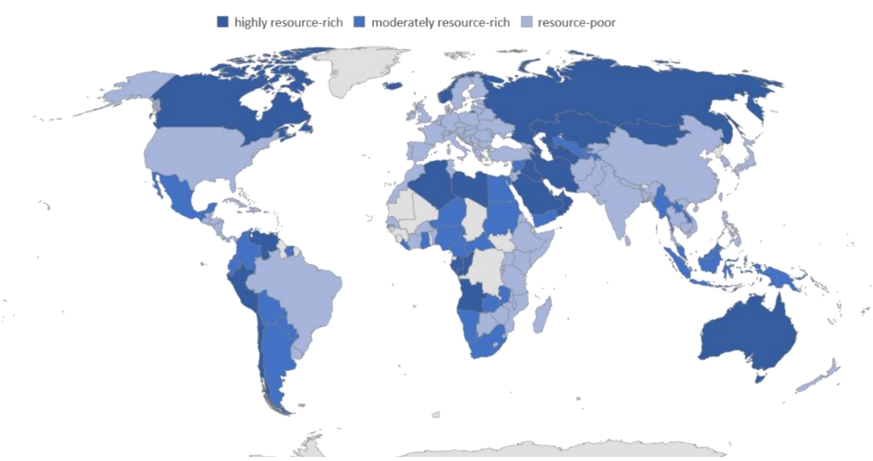
Source: Authors' calculations

Figure A5: Correlation Between Income per Capita and Data Availability in UNIDO.



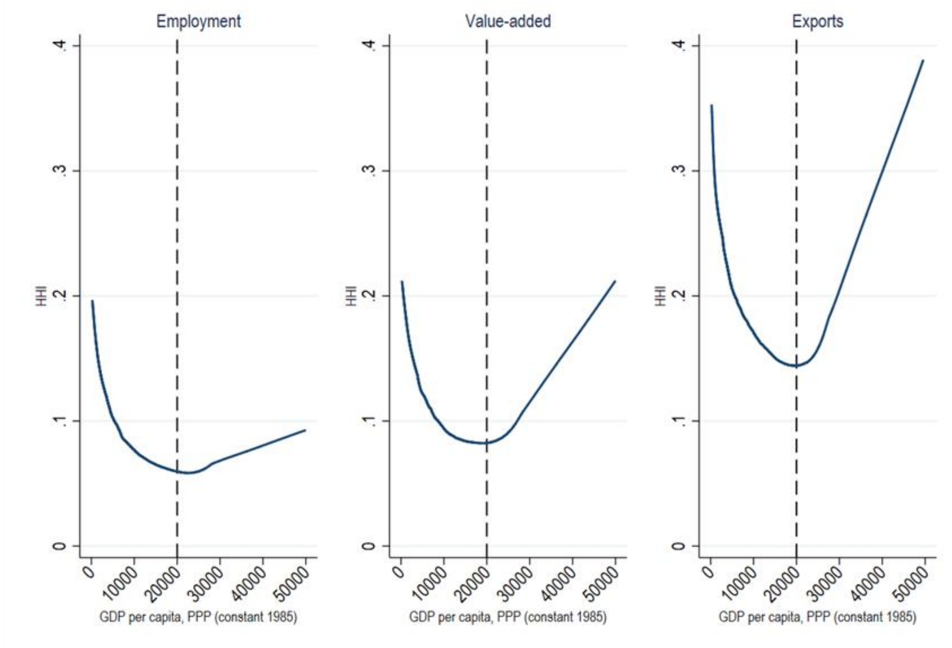
Source: Authors' calculations

Figure A6: Classification of Resource-Poor, Moderately Resource-Rich, and Highly Resource-Rich countries.



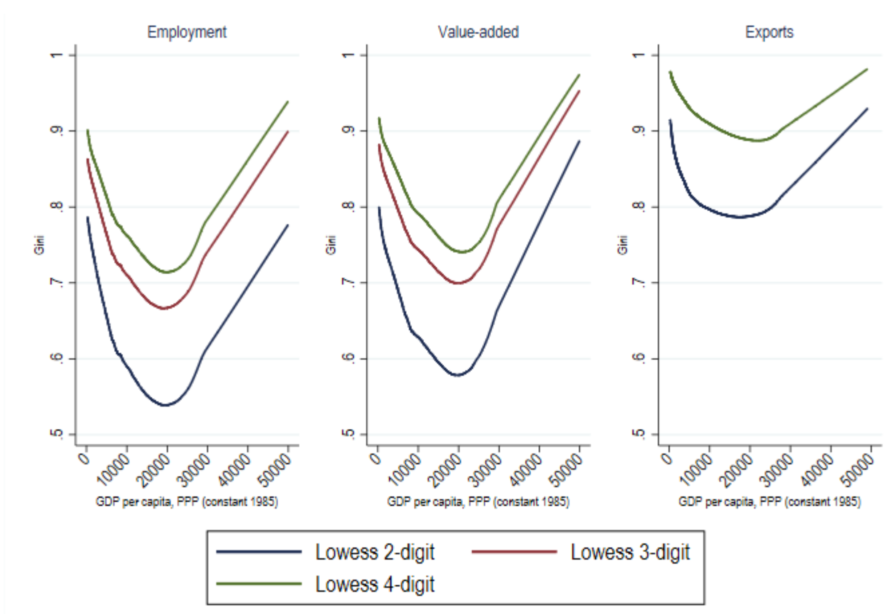
Source: Authors' calculations

Figure A7: Lowess Curve for the HHI index for Employment, Value-added and Exports



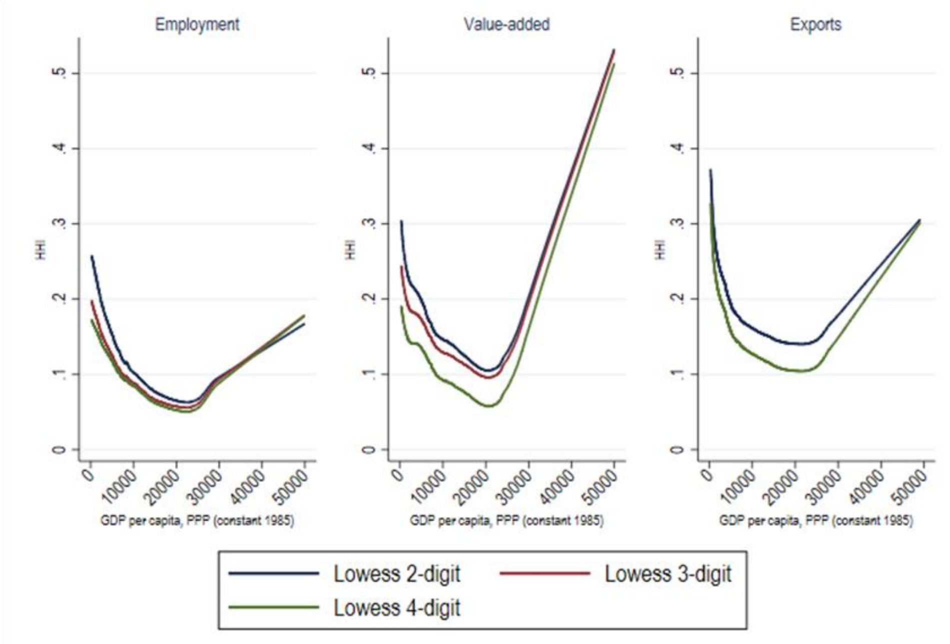
Source: Authors' calculations

Figure A8: Lowess Curve for the Gini of Employment, Value-added and Exports at Different Levels of Disaggregation



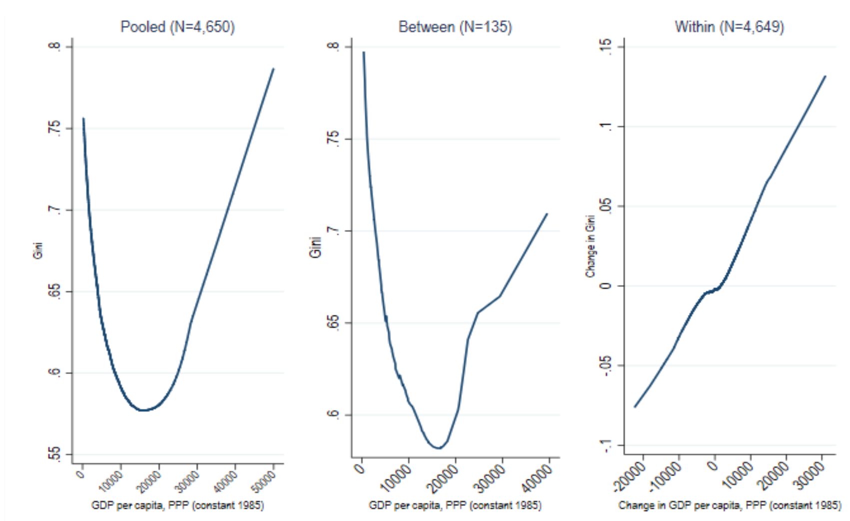
Source: Authors' calculations

Figure A9: Lowess Curve for the HHI of Employment, Value-added and Exports at Different Levels of Disaggregation



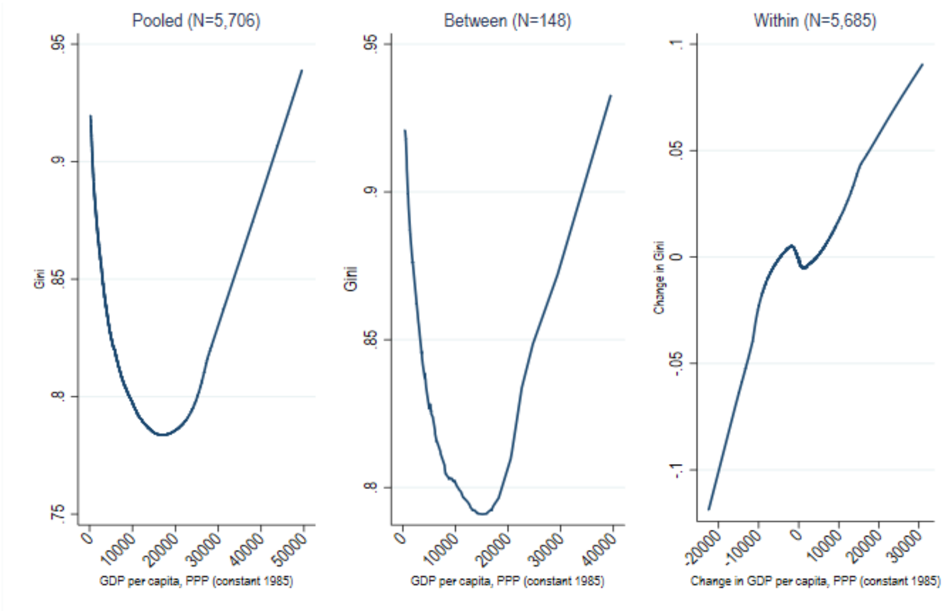
Source: Authors' calculations

Figure A10: Non-parametric Lowess Curves for the Pooled, Between and Within-Relationships Between Value-added and GDP per capita



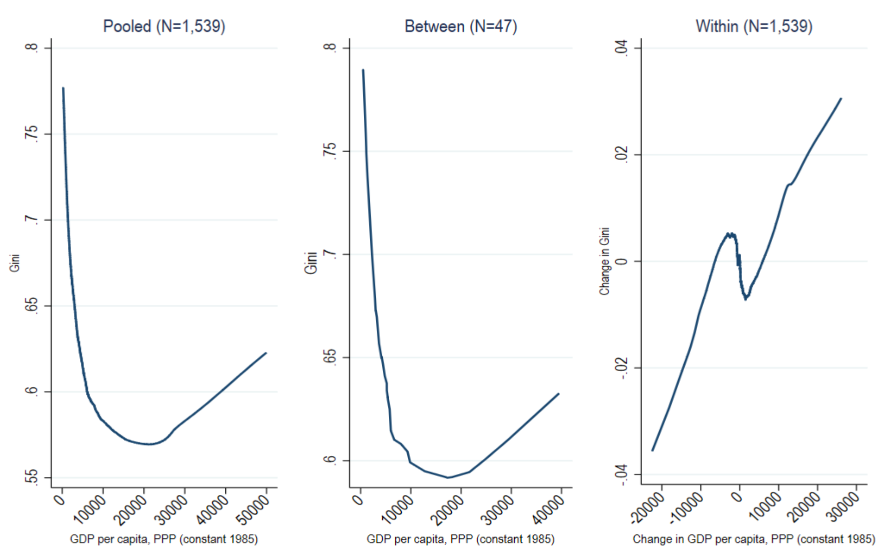
Source: Authors' calculations

Figure A11: Non-parametric Lowess Curves for the Pooled, Between and Within-Relationships between Exports and GDP per Capita



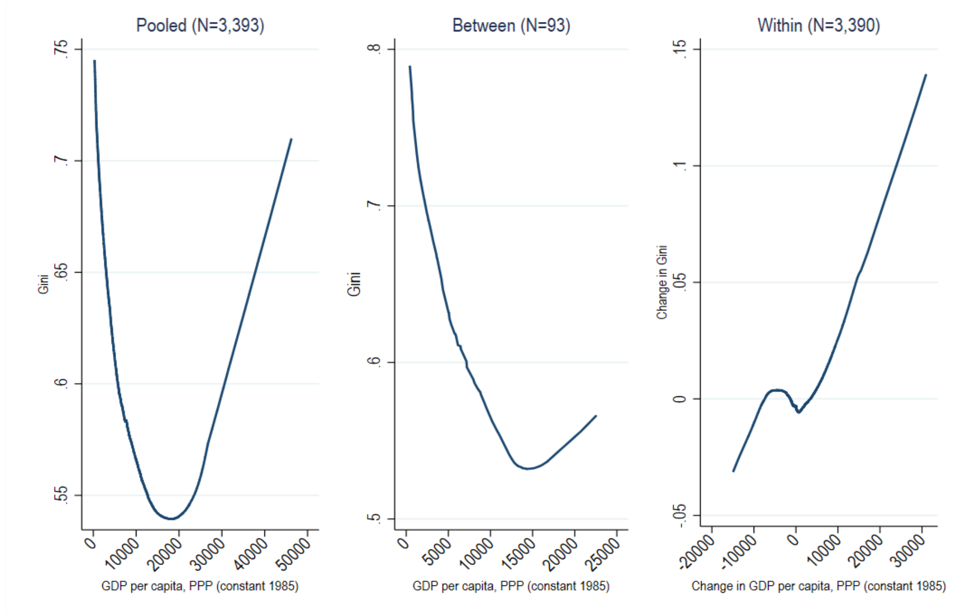
Source: Authors' calculations

Figure A 12: Non-parametric Lowess Curves for the Pooled, Between and Within-Relationships between Employment and GDP per Capita for Resource Rich countries



Source: Authors' calculations

Figure A 13: Non-parametric Lowess Curves for the Pooled, Between and Within-Relationships between Employment and GDP per Capita for Resource Poor countries



Source: Authors' calculations