Diversification vs. specialization as alternative strategies for economic development:

Can we settle a debate by looking at the empirical evidence?

UNITED NATIONS
INDUSTRIAL DEVELOPMENT ORGANIZATION

Diversification vs. specialization as alternative strategies for economic development:
Can we settle a debate by looking at the empirical evidence?

Florian Kaulich
Department of Economics
Vienna University of Economics and Business (WU Wien)
UNIDO Consultant

Table of contents

List of figures

List of tables

Abstract

Few ideas have been more vocally debated in the economic literature than that of diversification versus specialization as drivers of economic growth and development. Scores of scholars have written numerous articles debating the issue and, more importantly, policymakers across the globe have pioneered profoundly different policy approaches on the basis of such disagreements. As usual in the history of economic ideas, the most promising way to address a discussion with such strong ideological connotations is to assess the empirical evidence.

Motivated by the seminal work of Imbs and Wacziarg (2003), this paper attempts to synthesize the vast literature on the pros and cons of diversification vs. specialization, as well as the policy positions that emerge from such literature. It then provides the reader with an introduction of how such characteristics can be measured empirically and on the available evidence in terms of industrial production and exports (with the latter under absolute and relative specialization).

The empirical analysis identifies a positive relation between the diversification of an economy and its income at low levels of income per capita. At the same time, the available evidence is inconclusive about the occurrence of a negative relation between the two at higher levels of income per capita. This study provides further insights into policy measures aimed at facilitating the diversification of economies, especially in low-income countries. The study concludes with a few recommendations that warrant further research, which may provide very important qualifications to the above conclusion.

1. Introduction

For centuries, economists have debated the role of economic specialization and diversification in economic development. On the one hand, starting with the theory of comparative advantage in the early 19th century, the case has been made about the benefits of specializing on "what one does best". On the other hand, it has been argued that the diversification of production and exports can make a country less prone to negative economic shocks. Policymakers, in particular in low-income countries, are thus faced with contradicting theories about the best path to sustainable economic growth.

This protracted discussion on specialization, diversification and economic development has gained new impetus from recent empirical findings (Imbs and Wacziarg, 2003). The authors show that the economy of low-income countries is typically specialized in a narrow range of products. As GDP per capita rises, the structure of the production of goods diversifies through the launch of new products and through diversification within those goods that are already being produced or exported. At higher levels of GDP per capita, this diversification trend slows down and eventually veers towards re-specialization. The pattern of the relation between specialization and GDP per capita can therefore be described as a "U-curve".

This empirical evidence indicates that different theories apply at different stages of the economic growth process. The implication for low-income countries, in particular, is that they can overcome their economic marginalization through the acquisition of skills and knowledge necessary to diversify their economic portfolio rather than by focusing on "what they do best", while high-income countries seem to only benefit from specialization.

Research findings in this area are still preliminary. Some studies confirm the existence of a U-curve for the structure of production (Kalemli-Ozcan, Sorensen and Yosha, 2003; Koren and Tenreyro, 2004), while other studies have identified a U-curve in data on exports as well (Carrere, Strauss-Kahn and Cadot, 2006; Klinger and Ledermann 2004, 2006). Some authors reject the existence of a re-specialization and instead argue that the diversification process continues (De Benedictis, Gallegati and Tamberi, 2007). However, different and sometimes conflicting definitions and measurements of diversification/specialization have been used, together with different datasets. If a robust diversification trend can, at least, be confirmed for developing countries, policy advice will have to take it into account.

The purpose of this study is to contribute new empirical findings to the recent discussion on the stages of economic diversification and specialization. As a basis for the empirical analysis, the dispute over the existence and relevance of a specific trend of diversification, specialization or both is reviewed in the light of recent empirical findings. The conclusions from this debate are then considered by empirically analysing the main issues. The central question is: Does a country's economic structure initially diversify and then re-specialize as income grows? Specifically, what does "diversification" mean in this context, and what conclusions on the structure and dynamics of sectoral economic diversification can be drawn from the literature and from available data?

The structure of this study is as follows. Section 2 outlines the different strands of discussion on diversification and specialization. Section 3 briefly explains the statistical methods and datasets used in the subsequent econometric analysis. Section 4 presents the empirical results of this analysis in detail. Section 5 concludes this study.

2. The debate on specialization and diversification

2.1 Arguments concerning specialization

Traditional trade models suggest that – on an aggregate level – countries benefit from opening to trade and specializing in the production of goods in which they have a comparative advantage. By becoming more specialized, the allocation of resources becomes more efficient, allowing for mutual welfare increases (Krugman and Obstfeld, 2006). This idea goes back to David Ricardo, who pointed out – in his famous example of Portuguese wine and British cloth[1] – that although Portugal requires less labour to produce a unit[2] of either good compared to the United Kingdom, opening up for trade would benefit both countries, because they would be specializing in the good that has lower opportunity cost[3] (Ricardo, 1971:153-154). Hence, poor countries may be able to trade with rich ones and may gain from this trade (Ruffin, 2002:741f).

The major contribution of Ricardian theory is to provide a rationale of the *actual* pattern of trade and not necessarily how it *should* be, as the theory simply states that "[a] country exports that commodity in which it has comparative labour-productivity advantage." (Bhagwati and

---

[1] Ricardian theory can also be applied to a higher number of countries and/or goods (Dornbusch, Fischer and Samuelson, 1977; Bhagwati and Srinivasan, 1983:35-51; Becker, 1952).

[2] Maneschi (2004) points out that Ricardo did not mean unit labour coefficients, but that labour had to produce the amount of wine and cloth actually being traded. However, this discussion is not relevant for the present analysis.

[3] "Opportunity cost" in this context denotes how much the production of another good has to be dispensed to produce a good.

Srinivasan, 1983:29) Despite the simplicity of Ricardo's theory, it has been interpreted in several different ways, and continues to have a tremendous impact on how specialization and economic development are (mis)understood, as Deardorff (2005:3) denotes: "Comparative advantage is certainly one of the most basic ideas in economics, underlying much of our understanding of why countries trade the way they do and why they benefit from doing so. But it is also a difficult concept for many people to understand, and seemingly even more difficult for them to believe once they do understand it (and especially if they don't)." In the same vein, Nobel laureate Paul Samuelson has described comparative advantage as the best example of "an economic principle that is undeniably true yet not obvious to intelligent people." (Krugman and Obstfeld, 2006:24).

In fact, the Ricardian theory of comparative advantage states that specialization according to comparative advantage is an important factor in the production of more goods compared to economic autarky. Still, while trade according to comparative advantage is necessary to realize gains from trade, it does not suffice. For example, world prices could equal autarky prices for some countries, resulting in zero gains from trade for them. The simplicity of Ricardian theory of static comparative advantage generally results in extreme predictions about trade patterns, but if one allows for a small amount of more realistic – thus more complex – assumptions, the predictions and political implications become less clear (Deardorff, 2005:5-13).

Thus, the drawing of conclusions for economic policy from the Ricardian model, in particular policy recommendations based on the potential advantages of specialization, should be done with caution. Nevertheless, the necessity of specialization according to comparative advantage for economic development continues to be an integral part of policy advice, as Rodrik (2007:103) affirms: "Those who associate under-development with inadequate exposure to international markets generally imply – although this is often left unstated – that specialization according to comparative advantage is an essential ingredient of development." The World Trade Organization proclaims the general applicability of this concept on its website: "Simply put, the principle of 'comparative advantage' says that countries prosper first by taking advantage of their assets in order to concentrate on what they can produce best, and then by trading these products for products that other countries produce best." (WTO, 2009) When focusing on the industrial sector, Hausmann and Rodrik (2003:23) state that "for all economies except possibly the most sophisticated, industrial success entails concentration in a relatively narrow range of high-productivity activities".

2.2   Arguments concerning diversification

The relevance of economic diversification has been advocated by famous economists such as Nobel laureate Simon Kuznets, who states in his Nobel Prize lecture that "[a] country's economic growth may be defined as a long-term rise in capacity to supply increasingly diverse economic goods to its population […]." (Kuznets, 1971). Grossman and Helpman (1992:334) make an even stronger statement by asserting that "[g]rowing economies produce an ever-increasing quantity, quality and variety of goods and services."

The most straightforward argument for the importance of diversification is that diversified economies are less vulnerable to economic shocks than specialized economies: "[…] [A]lthough there are good theoretical arguments for specialization according to comparative advantage", Osakwe (2007:1) argues, "in practice policymakers in developing countries are interested in diversifying their production and export structure to reduce vulnerability to external shocks." Moreover, more diversified economies are less volatile in terms of outputs, and lower output volatility is associated with higher economic growth (Ramey and Ramey, 1995).

An early concept that highlights the particular problem of specializing in agriculture is the so-called "Graham paradox", which incorporates non-constant unit costs – hence productivity – between different sectors into Ricardian theory (cf. Graham, 1923). More specifically, productivity in the manufacturing sector rises with production as unit costs fall with increasing output due to the benefits of mass production, while the unit costs of agricultural products increase with production. For a country with a comparative advantage in agriculture, specialization according to comparative advantage decreases productivity in both the agricultural and the manufacturing sectors, hence, the country's total output declines. Even global production can decline if the increase in production of countries specializing in manufacturing is not large enough (Raffer, 2004:112-117).

Prebisch (1950) and Singer (1950) have further elaborated the argument of the importance of diversification for economic growth. Their most influential contribution may have been the formulation of the so-called Prebisch-Singer hypothesis (PSH), which asserts that economic growth cannot be based on resource-based products, because world prices for primary exports relative to manufactured exports decline over time. Consequently, the ratio of export prices to import prices – the terms of trade – for developing countries, which are mostly heavily dependent on exports of commodity products, is declining as well. The proposed explanations are: (1) Strong labour unions in industrialized countries cause wages in the manufacturing sector

of each business cycle to rise at a much higher extent than wages in developing countries;[4] (2) Monopoly power in manufactures prevents technological increase and results in lower prices; (3) Demand for primary commodities shows a relatively lower income elasticity, which means that income growth tends to reduce the relative demand for, and hence price of, primary commodities; and (4) Raw-material-saving technical progress in manufacturing causes a relatively slow-growing demand for primary products (Cuddington, Ludema and Jayasuriya, 2002:5). Eventually, the PSH was used as a theoretical justification for economic diversification through Import Substitution Industrialization,[5] which is labelled a "great historical mistake" by Sachs and Warner (1995:4), because it was based on prolonged trade barriers rather than on export promotion.

The PSH has been widely discussed in the literature, with conclusions being drawn both for and against its validity. Lutz (1999) builds on this mixed evidence and confirms the validity of the PSH, as do Ocampo and Parra (2004), and Raffer (2004:119) concludes that the PSH has been widely accepted since the 1990s. Cuddington, Ludema and Jayasuriya (2002), however, demonstrate that the terms of trade of primary products have experienced a few abrupt shifts – or structural breaks – downwards, but do not follow any particular trend.

Overall, the Graham paradox and the PSH do not provide arguments in favour of diversification per se, but explain the disadvantage of being specialized in the "wrong" sector, namely, agriculture, as opposed to being specialized in manufacturing. In principle, these arguments can therefore serve as a rationale for changing the respective sector in which a country specializes or as justification for overall economic diversification.

The literature on endogenous growth theory also highlights the importance of the nature of the sector in which a country specializes, as the returns to scale depend on the sector itself. Once increasing returns to scale are assumed in the manufacturing sector, and constant returns to scale are assumed in the agricultural sector, it obviously follows that when a country "initially has a comparative advantage in manufacturing (agriculture), its manufacturing productivity will grow faster (slower) than the rest of the world and accelerate (slow down) over time." (Matsuyama 1991:11) Structural models of economic development show that countries should develop their

---

[4] During economic booms, strong labour unions can negotiate for wage increases, while during recessions unions can prevent wages from falling. In the absence of labour unions, wage increases during booms are lower, while recessions can cause decreasing wages.

[5] Import Substitution Industrialization is a strategy to replace imports with domestic products by diversifying the domestic production structure. This strategy was used in Latin American countries for the first time during the Great Depression (Nuscheler, 2004:627).

export structure from primary exports into manufactured exports in order to achieve sustained economic growth.

Collier (2002) lists three additional severe problems developing countries face, stemming from their heavy dependence on exports of primary commodities: (1) As commodity prices are highly volatile, countries have to cope with large external shocks. (2) Rents generated by primary commodities are usually associated with poor governance. (3) Dependence on a narrow range of natural commodities increases the risk of civil war as natural resources might generate income for rebel groups. Generally, the negative impact of natural resource abundance and economic growth has been coined the "curse of natural resources" (cf. Sachs and Warner 1995, 2001) or – in the context of a single booming sector that negatively influences the industrial sector – the "Dutch Disease" (cf. Corden and Neary, 1982).

The portfolio effect – as in the finance literature – might apply to the export structure as well: A specialized export structure, especially when a country depends on commodity products with volatile market prices, discourages necessary investments by risk-averse firms. Diversification of exports therefore helps to stabilize export earnings in the long run (Hesse, 2008). However, Bebczuk and Berretoni (2006:8) warn against explaining export diversification from an aggregate viewpoint only, since the decision to diversify the export portfolio is taken by individual firms (assuming that the government has no direct influence on the export structure). Extending the insights about incentives to diversify financial assets to export diversification might be misleading, as flexible financial markets differ from the inflexible production decisions of firms, which are more irreversible and depend on a much broader set of conditions.

The stimulating effect of export diversification on the creation of new industries can also take place through forward and backward linkages (Hirschman, 1958:98-119). Export diversification does not always mean climbing the ladder of value added, however. For example, the case of Chile's export diversification since the 1970s has seen neither the emergence of heavy industry through industrial policy, nor the imitation of high-technology products, but instead the emergence of new agricultural products (De Pineres and Ferrantino, 1997:389).

While the benefits of a diversified export structure have been well-established in the literature, no unified framework to describe the drivers of export diversification exists. The *causality* between export diversification and growth can potentially run in two directions: On the one hand, the acquisition of new comparative advantages can lead to countries entering new markets

and increasing their income while, on the other, countries with a low GDP income per capita tend to have a comparative advantage in a limited range of goods, as they lack the skills or inputs to apply knowledge that already exists elsewhere. As GDP per capita rises, a country becomes increasingly able to produce a wider range of goods and compete in international markets (Agosin, 2007:4).

Once the issue of uncertainty is incorporated into the Ricardian model, the predicted specialization patterns can oppose those under certainty, as a risk-averse country will shift its production towards another good if price uncertainty in the initial good is too large. The expected gains from trade for a country with absolute risk aversion can become negative, causing it to cease trading altogether (Propositions 7 and 9 in Turnovsky, 1974:211-215). Acemoglu and Zilibotti (1997) developed a theoretical model with uncertainty on the return to investments, where economic development goes hand in hand with better opportunities for diversification and a more productive use of financial funds. At early stages of economic development, the presence of indivisible projects limits the degree of diversification – hence risk spreading – that the economy can achieve. This inability to diversify risks introduces a large amount of uncertainty in the growth process, and the desire to avoid highly risky investments slows down capital accumulation. The chance of conducting profitable projects determines how long countries remain in the stage of initial capital accumulation before they reach a takeoff stage where full diversification of risks can be achieved.

2.3 Arguments concerning both specialization and diversification

The long-lasting discussion described concerning the relevance and trend of economic specialization or diversification has consisted of theories, political arguments or evidence either exclusively *for* or exclusively *against* economic specialization. A non-linear relationship – that is, a relationship that consists of specialization trends in both directions – was not considered until recently. In their seminal econometric analysis of the stages of diversification, Imbs and Wacziarg (2003) were the first to consider and detect a non-linear – namely, a U-curved – relationship between diversification of production and GDP per capita. Their findings consist of the following empirical stylized facts: (1) Low-income countries have a very specialized production structure; (2) As countries' levels of GDP per capita increase, the sectoral distribution of economic activity diversifies; (3) This diversifying trend decreases with rising GDP per capita, and after a turning point – which takes place at a very high level of income – the sectoral distribution exhibits re-specialization.

Although simple in nature, this discovery proved to be a novelty and initiated a new debate on the structure of growth, as Imbs and Wacziarg (2003:63) predicted: "This new finding has potentially important implications for theories of trade and growth. Most existing theories predict a monotonic relationship between income and sectoral concentration." In his weblog, Rodrik (2007b) even labelled this finding "[o]ne of my favourite stylized facts about development […]".

Imbs and Wacziarg (2003) use several datasets to demonstrate the robustness of their results. The first dataset covers employment data from the ILO on nine economic sectors for the period 1969 to 1997 for a large set of countries ranging from low to high income. The second dataset uses OECD data on value added and employment for the period 1960-1993, covering 14 economic sectors. The third dataset from UNIDO consists of employment and value added covering 28 manufacturing sectors. The datasets are edited in a way that the number of sectors available through time for each country is constant, which involves dropping several insufficient observations. To estimate within-country variation, countries for which 27 or more sectors were available from UNIDO data, and 6 or more from ILO data are retained. For between-country variation, the sample was restricted to observations where all sectors were reported. To demonstrate the robustness of their results, Imbs and Wacziarg (2003) employ several measures of sectoral concentration, among them the Gini coefficient and the Herfindahl Index.[6]

To investigate the structure of the data without imposing any specific functional form, Imbs and Wacziarg (1993:67-72) first employ a non-parametric methodology based on locally robust weighted scatter plot smoothing (LOWESS), but they allow for country fixed effects. The results indicate a U-curved relationship between specialization and GDP per capita,[7] where countries diversify over a large range of GDP per capita. The upward-bending part, at the highest levels of GDP per capita, is distinct but does not reach the level of sectoral concentration on the left part of the curve (Figure 1). Although less pronounced than the initial specialization, this upward-sloping part of the curve appears to be statistically significant for all datasets. The turning point at which the diversification trend switches to re-concentration appears to be at quite a high level of income per capita, with re-concentration within the manufacturing sector occurring earlier than across a broader range of sectors.[8]

---

[6] A higher value of the Gini coefficient or Herfindahl Index indicates a higher degree of specialization, while a value of zero means that the distribution of economic activity is equal across sectors.
[7] Here measured in constant international dollars at purchasing power parity (PPP).
[8] The turning point lies at around US$16,500 at constant 2000 international dollars at PPP (UNIDO 2009:12).

As a next step, Imbs and Wacziarg (2003:72-75) run parametric estimations, in particular fixed effects panel regressions of specialization on GDP per capita and squared GDP per capita. Including both a linear and a squared term of income per capita allows verification of the hypothesis of a U-curve as indicated by the non-parametric method, since a negative coefficient on the linear term and a positive one on the squared term add up to a U-shaped function. The right curve in Figure 1 presents the regression output of the fixed effects (within-country) panel regression over a backdrop of the pooled panel scatter plot. By using between- and within-country regressions, the existence and robustness of a U-shape as a phenomenon of differences between countries, as well as structural change within countries can be confirmed.[9]

Figure 1   Specialization of MVA over GDP per capita, on-parametric curve (left) and fixed effects regression
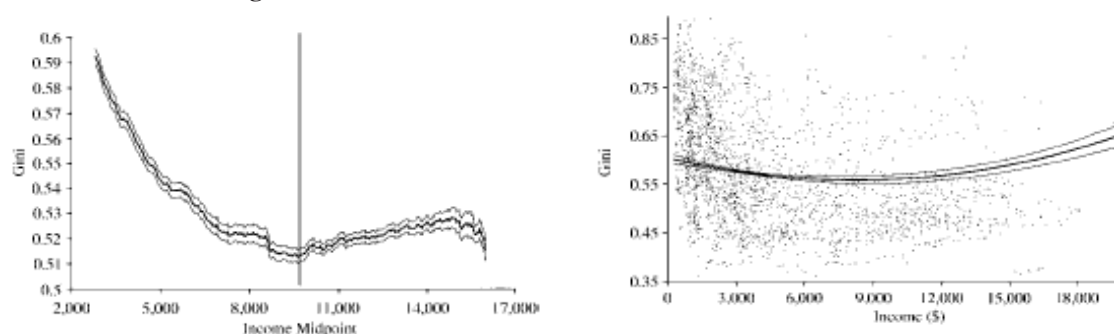
Source: Imbs and Wacziarg (2003:69).

The novelty lies in the fact that historic arguments, which are either exclusively for or exclusively against specialization and therefore seem to contradict each other, might *all* be correct, as "[…] each set of theories seems to be at play at different points in the development process." (Imbs and Wacziarg, 2003:64). These results are particularly surprising if one thinks within an oversimplified Ricardian model of specialization. "What is significant about this finding from our standpoint", Rodrik (2007:103) argues, "is that it goes against the standard intuition from the principle of comparative advantage. The logic of comparative advantage is one of specialization. It is specialization that raises overall productivity in an economy that is open to trade. […] Imbs and Wacziarg's findings suggest otherwise."

---

[9] Furthermore, Imbs and Wacziarg (2003:75-82) use several additional methods to prove the robustness of their results, such as focusing on individual countries and accounting for country size, periodic-specific and region-specific effects. Additional results and robustness checks are presented in Imbs and Wacziarg (2003), which supplements Imbs and Wacziarg (2003). It should be noted that Imbs and Wacziarg (2003) do not analyse any intra distribution dynamics – the same value of the Gini Index can result from different Lorenz curves – neither the "quality" of the distribution, or if the different locations within the specialization-GDP-space correspond to an optimal combination of goods in terms of impact on welfare, growth, employment, and so on.

Although it is tempting to replace the old oversimplified "rule" of economic growth via specialization with a new "rule" of economic growth via diversification and late re-specialization, it might still make sense to focus on the importance of diversification for developing countries, as Subramanian (2007:2) concludes from the U-curve: "[S]uccessful growth is accompanied by the private sector undertaking new, varied, and sophisticated activities […]. All economies start off agricultural, and the successful ones diversify away from agriculture toward manufacturing and, within manufacturing, from simple to more sophisticated activities. Diversification is thus intrinsic to development."

In an attempt to rationalize their empirical findings, Imbs and Wacziarg (2000)[10] propose a theoretical model which endogenizes the stages of diversification via trade forces. Each country produces only the subset of all potentially producible goods in which it is most productive, i.e. which can be produced cheaper than imported products. As a country catches up with the global technological frontier, its aggregate productivity rises as does the number of goods that can be produced domestically at competitive prices, and thus the country diversifies. But as infrastructure improves, transport costs fall, which leads to a decrease in the prices of imported goods. In other words, "[…] the presence of transport costs forces diversification beyond comparative advantage." (Imbs and Wacziarg, 2000:11). As a result, the number of domestically produced goods decreases and concentration rises again.

Linking the financial sector with the real sector represents a different theoretical explanation: When a country does not have access to global financial markets that could alleviate potential negative sectoral shocks, the only way left to absorb the potential effects of a sectoral shock is to diversify production across sectors. Once a country has gained access to the global financial market, sectoral diversification is no longer required to spread the risk. At this stage, the country can experience the advantage of specialization attributable to economies of scale via the international division of labour (Saint-Paul, 1992).[11]

---

[10] Imbs and Wacziarg (2000) is an early working version of Imbs and Wacziarg (2003).

[11] Saint-Paul (1992) finds the equilibrium with higher financial integration and specialization of technology more "appealing" than the equilibrium with lower financial integration and a diversified production technology, because "[…] we tend to think that financial markets are the most appropriate instrument for such a diversification." Saint-Paul (1992:764) However, this opinion can be questioned, especially in view of the 2008/2009 global financial crisis.

Batista and Potin (2007) apply a Heckscher-Ohlin type model[12] to explain the causal relations of the U-curve: countries with a low capital-labour ratio (i.e. primarily low-income countries) specialize in the production of labour-intensive goods, countries with a high capital-labour ratio, i.e. high-income countries, specialize in the production of capital-intensive goods, and countries with an intermediate capital-labour ratio produce both types of goods. In particular, for countries with a very low amount of capital, the value added per worker in the given labour-intensive good first increases sharply with rising capital intensity, while value added in the capital-intensive good remains at zero. After reaching a point of maximum value added per worker in the labour-intensive sector, the value added per worker in this sector declines with rising capital-labour ratio, while the value added per worker in the capital-intensive sector rises until the country is fully specialized in capital-intensive goods. Batista and Potin (2007) find that neoclassical factors – capital accumulation through the Rybczinski[13] effect, changes in relative prices, biased technological change – account for at least one third of the evolution of economic specialization, leaving the rest for other explanations like economies of scale and risk diversification.

When excluding the top quartile of the countries according to GDP per capita, Bebczuk and Berrettoni (2006) confirm a U-shaped relationship between export specialization and GDP per capita. Klinger and Lederman (2004) confirm this U-curve for all countries and report a turning point towards re-specialization at a higher level than that for production data found by Imbs and Wacziarg (2003). [14] This leads them to conclude that "[…] the pattern of economic diversification observed by Imbs and Wacziarg is probably driven by patterns of international trade flows" (Klinger and Lederman, 2004:21), which they denote as "trade-driven economic diversification". However, it might also be reasonable to conclude from this observation that export patterns *follow* production patterns: countries first develop their production portfolio nationally and then enter the global market, with certain sunset industries continuing to export despite declining production. A similar conclusion is also drawn in UNIDO (2009:12).

Harrigan (2007) confirms the existence of a U-curve between specialization and GDP per capita for MVA data when analysing a panel of 14 Asian economies for the period 1970-2005. By

---

[12] The HO-model they employ is based on several assumptions such as constant returns to scale, access to the same technology by all countries, small size of countries, free trade between countries, and a production where only two factors (capital and labour) are used to produce the two types of goods.

[13] The Rybczinski effect shows that an increase in the endowment of one production factor causes a more than proportional increase in the output of the good that uses this factor intensively (Krugman and Obstfeld, 2006:60).

[14] Klinger and Ledermann (2004) use export data from UN Comtrade (2008) at the SITC 3-digit level (around 175 commodity groups), HS 4-digit level (around 1,200 commodity groups) and HS 6-digit level (around 5,000 commodity groups) of aggregation.

running a pooled panel regression, he disregards country fixed effects, however, i.e. he does not account for whether the U-curve is driven by static differences between countries or by a movement of every country along a U-curve. Harrigan (2007) further argues that the observed U-curve might be driven by countries' sizes.

De Benedictis (2004) uses different settings for the non-parametric approach to show that the relation between diversification and GDP per capita is highly non-linear, with alternating phases of diversification and concentration along the path of rising GDP per capita. When using country fixed effects, countries always diversify along the path of rising GDP per capita, but this relationship is influenced by other explanatory variables, such as the size of the country (GDP or population), the level of openness and the quality of institutions. In a subsequent paper, but using the same dataset, De Benedictis, Gallegati and Tamberi (2008) employ a fixed effects generalized additive model to confirm the presence of diversification, but cannot find evidence for a re-specialization trend.

The relationship between production or export diversification and economic *growth* is of additional interest, as it adds a time-dynamic perspective to the otherwise static diversification-income-analysis. Al-Marhubi (2000) shows that when controlling for other determinants, countries with a relatively higher export diversification[15] experience faster growth. Lederman and Maloney (2003:15) provide further evidence of a positive effect of export diversification on the growth of GDP per capita that is robust to including other explanatory variables. Hesse (2008) finds a negative and linear relationship between export concentration and GDP per capita growth. Agosin (2007) develops and empirically tests a model of export diversification and economic growth, and finds that the introduction of new exports accounts for the main share of sources of economic growth in countries that are below the global technological frontier.

2.4 Diversification and product characteristics

As the discussed stages of diversification relate to general diversification patterns but do not discriminate between different types of products, a logical follow-up question concerns the *nature* of the sectors or products a country specializes in. UNIDO's 2009 Industrial Development Report (UNIDO, 2009) links the U-curve of specialization with the nature of the products being produced, characterized by their sophistication. "Sophistication" in this context is a new measure of the complexity of products, which was traditionally measured by the level

---

[15] He calculates a simple form of relative diversification using export data at SITC 3-digit level, excluding countries which account for less than 0.3 percent of the country's total exports.

of technology. Lall (2000) emphasizes the relevance of the technological composition[16] of a country's export basket for industrial development. An export structure with higher technological intensity offers better prospects for future growth, because the growth of trade in high-technology products tends to be greater due to higher income elasticity, creation of new demand, faster substitution of older products, greater potential for further learning and larger spillover effects.

Although this concept of "technology" serves well for many purposes, it is a static, ad-hoc concept, and is limited to aggregated product groups. To overcome these shortcomings, the concept of "sophistication" was developed by Lall, Weiss and Zhang (2006) and Hausmann, Hwang and Rodrik (2007): a given good is classified as more sophisticated the higher the average income of its exporter. The rationale for this index is that products which are exported by high-income countries have characteristics that allow high-wage producers to compete in world markets. These characteristics include, inter alia, technology, transport costs, natural resource availability, marketing and infrastructure quality. The sophistication index therefore represents an indirect measure of an amalgam of a varying set of influences. In UNIDO (2009), this concept of measuring sophistication is combined with the U-shaped curve of specialization and GDP per capita. Countries diversify by moving towards sophisticated products, and reach the highest level of diversification by producing low- and medium-sophisticated products. High-income countries specialize by producing highly sophisticated products (Figure 2).

Kalemli-Ozcan, Sorensen and Yosha (2003) use the level of *risk*, measured by the insurance against shocks to specialization, to classify those sectors that countries specialize in at the low-income and high-income section of the U-curve. They quantify income insurance indirectly by measuring the deviation of the movement of GDP from the average movement within a wider geographical group of regions. Hence, if a region is insured via capital markets, its GDP fluctuation should not deviate from the fluctuation of the risk-sharing group. They use data on sectoral manufacturing output, value added and employment for several OECD countries and regions within countries to calculate a country's specialization as the sum of all sectoral deviations from the group mean. By regressing specialization on income insurance, linear and squared GDP per capita and several other potential determinants[17] of specialization, they find a

---

[16] Lall (2000:34-35) divides all products in the SITC Revision 2 (3-digit level) into primary products, resource-based manufactures, low-technology manufactures, medium-technology and high-technology manufactures.

[17] They include trade volume, factor endowments, distance, shipping cost, customs union, education and population as additional control variables.

U-shaped relationship between GDP per capita and specialization, and a positive correlation between risk-sharing and specialization.

Figure 2   Specialization and sophistication
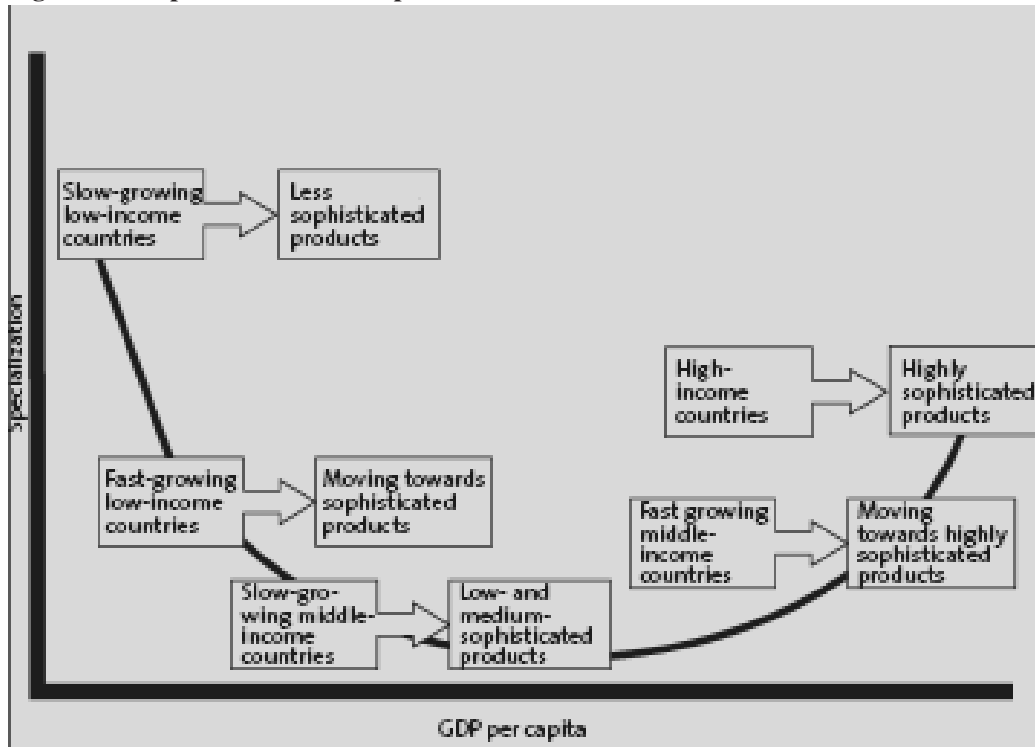
Source: UNIDO (2009:18).

Koren and Tenreyro (2004) further investigate the idea of risk by modelling the economy as a portfolio of sectors with different risk intensities referring to volatility, inter-sectoral correlation and broadness of sectors.[18] They calculate several measures of sectoral risk which are then used as weights in the sectoral specialization measures to extend the results of Imbs and Wacziarg (2003). They find that at low income per capita, countries are relatively concentrated in high-risk sectors. As income increases countries extend their production towards low-risk sectors, thus experiencing a decrease in specialization. Finally, while the trend to diversify becomes weaker and eventually evolves into re-specialization with rising GDP per capita, sectoral risk continues to decline.

---

[18] High-technology sectors are more disaggregated than low-technology sectors and agriculture.

14

2.5   Diversification by discovery

It is important to distinguish whether diversification takes place within existing sectors or by introducing new sectors that have not been exported before and are discovered to be profitable. The term "discovery" in this context was established by Hausmann and Rodrik (2003) to denote the production of a new good that does not necessarily stem from innovation but from entrepreneurial copying from abroad. This notion is particularly relevant in the context of developing countries.

Klinger and Lederman (2004) find that economic discoveries in the 1990s did not only take place in "modern" sectors, but also in sectors that are considered "traditional", such as foodstuffs and agriculture, with the highest level of discoveries in chemicals.[19] The relationship between discoveries and GDP per capita appears to be an inverted U-curve, but highly skewed to the left, indicating that the initial stage of diversification is driven by discoveries, while the subsequent stage of diversification is driven by dispersing production among goods that are already being produced. Surprisingly, sectoral discoveries and income per capita do not significantly differ across industries. This indicates that economic discoveries are *not* driven by the process of structural transformation, as this would only be the case if discoveries in "traditional" labour-intensive sectors peak at low levels of development, whereas discoveries in "modern" capital-intensive sectors peak at high levels of development. Klinger and Lederman (2004:29) conclude that "[…] developing countries are not limited to discoveries in certain sectors based on their level of development."

In a subsequent paper, Klinger and Lederman (2006) further investigate the idea of diversification by dividing diversification into inside-the-frontier-innovation, i.e. discovering the profitability of an existing product, and on-the-frontier-innovation, i.e. invention of new products measured by new patents. Figure 3 presents their empirical results of the evolution of discoveries (dark solid line), patents (dashed line) and overall specialization (grey line) in relationship with GDP per capita: low-income countries mainly introduce new products through discoveries, but as GDP per capita grows, the amount of discoveries decreases while the number of new patents rises. In parallel, the overall specialization follows a U-curve.

---

[19] Klinger and Lederman (2004) refer to a "discovery" when the export level of a product was below US$10,000 in 1992 and above US$1,000,000 in the period 2000-2002.

Figure 3   Diversification and innovation
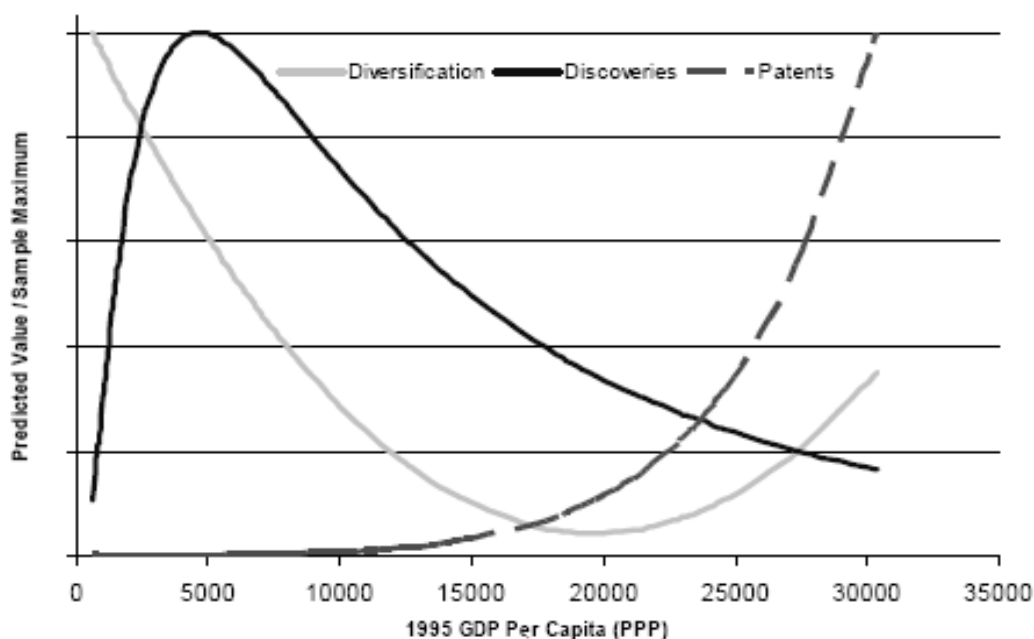
Source: Klinger and Lederman (2006:15).

A similar approach is to directly divide exports into differences at the extensive margin, i.e. differences in the number of product lines, and the "intensive margin", i.e. differences in the amount exported of the same number of product lines. Hummels and Klenow (2005) find that the extensive margin accounts for 62 percent of the greater exports[20] of larger economies. Carrère, Strauss-Kahn and Cadot (2007) further investigate this idea by dividing not just the export *value* but also export *diversification* into diversification at the extensive and intensive margins, i.e. diversification attributable to the addition of new product lines and to a more equal distribution within a constant number of product lines.[21] They confirm a U-shaped relationship between national income and export specialization for pooled and between regressions, but not for within-country regressions, with the turning point found to vary among specialization measures.[22] As a separate exercise, Carrère, Strauss-Kahn and Cadot (2007) show that the U-curve is dominated by changes in the extensive margin, e.g. the opening and closing of new product lines. Unsurprisingly, countries with a higher share of raw material exports are more specialized, providing some support for the Dutch Disease hypothesis. However, controlling for raw material exports does not affect the turning point, which shows that the non-linear shape is

---

[20] Measured at the HS 6-digit level for 126 exporters and 59 importers.

[21] Carrère, Strauss-Kahn and Cadot (2007) use export data from UN Comtrade (2008) on the HS 6-digit level and calculate the Gini, Theil and Herfindahl indices.

[22] Carrère, Strauss-Kahn and Cadot (2007) suggest that the non-existence of a turning point when accounting for country fixed effects may result from the short time span covered.

indeed a feature of the overall economic development process rather than a reflection of the existence of primary product exports.

In their analysis of product discovery, Klinger and Lederman (2004:34-35) find that export growth has a significant and positive impact on product discovery, indicating that export-promoting strategies are also discovery-promoting. Nevertheless, absorptive capacity is negatively and barriers to entry are positively correlated with discovery, which is surprising, and lends support to the market failure hypothesis described in Hausmann and Rodrik (2003).

2.6   Economic policies for diversification

The policy implications resulting from recent empirical findings on the stages of diversification are immense. The common notion to specialize in "what one does best" as a means to achieve economic prosperity and hence poverty reduction seems to be fundamentally wrong. "Whatever it is that serves as the driving force of economic development", Rodrik (2007:103) concludes from Imbs and Wacziarg (2003), "it cannot be the forces of comparative advantage as conventionally understood. The trick seems to be to acquire mastery over a broader range of activities, instead of concentrating on what one does best."[23] The misconception of comparative advantage in this context is the idea of interpreting it as a static rather than a dynamic process. These insights might not come as a big surprise, but according to Rodrik (2005), they challenge what has been taught in economics doctorate programmes at North American universities over the past four decades, namely, that a country must specialize in accordance with its comparative advantages in order to achieve prosperity and free itself from poverty.

The seminal study of Imbs and Wacziarg (2003), therefore, has not only had a huge impact on the academic discussion, but also influences international development policy. In UNIDO's 2009 Industrial Development Report (UNIDO, 2009), the U-curve serves as one out of several arguments that what a country manufactures affects growth – the "new structuralist" view. The crucial question is why and how diversification in low- and middle-income countries is taking place. Is it a result of market forces that might stimulate diversification through competition or trade, or can diversification only be achieved through public economic policy? In the latter case, market forces might better serve as an explanation of the flattening of the curve and the proposed eventual turn of diversification. After diversifying an economy through industrial policy and thereby reaching a certain level of GDP per capita, the influence of market forces increases and inefficient branches of the economy shrink. Also, Rodrik (2007:99-152) uses the

---

[23] See also Rodrik (2007c:9-10) for a similar conclusion.

findings of Imbs and Wacziarg (2003) to justify industrial policy that promotes the diversification of the production portfolio.

To achieve diversification, the current paradigm of liberalizing an economy and investing in human capital to exploit the existing comparative advantages in simple activities may generate economic growth in the short run, but once the initial "easy" stage of exporting is completed, significant technological upgrading and deepening are required to continue the growth trend (Lall 2000:30). The resulting question is therefore why some countries manage to develop a broader range of products than others. The answer to this question will shed light on the matter of the "right" industrial policy.[24] Rodrik (2007:100, italics in original) proposes that "[t]he nature of industrial policies is that they complement – opponents would say 'distort' – market forces: they reinforce or counteract the allocative effects that the existing markets would otherwise produce. […] [The] analysis of industrial policy needs to focus not on the policy *outcomes* – which are inherently unknowable ex ante – but on getting the policy *process* right."

To learn more about the "right" policy process, we first need to understand why industrial policy is needed for diversification as opposed to simply letting market forces act. One answer may lie in market failures through externalities. Hausmann and Rodrik (2003) apply the idea of information externalities to problems faced by developing countries. Their challenge is not to develop *new* products or processes, but to discover that a certain product or process, which is already well established in world markets, can be produced locally at low cost. Most knowledge is tacit, meaning that it cannot be formalized and transferred to other countries. The innovators of new goods and processes can be protected through an intellectual property rights regime, but an investor who discovers the profitability of an existing good does not receive such protection, so the private returns on investments in discoveries lie below the social returns – a laissez-faire policy would therefore initially create too little investment. Entrepreneurial effort – and therefore investment – is also required to adapt a product that is already being produced domestically to the "taste" of potential foreign markets. An example is wine production in Chile. Wine has been produced in Chile since the 16th century, but has only been exported since 1985 after entrepreneurs introduced modern techniques and uncovered foreign demand patterns (Agosin and Bravo-Ortega, 2007:11-25).

Rodrik (2007:107-109) describes coordination failures as another type of externality relevant to the discussion of economic diversification. Starting a new economic activity depends to a large

---

[24] Rodrik (2007:100) generally defines "industrial policy" as "policies for economic restructuring".

extent on the surrounding infrastructure and other supporting institutions, which have a high level of fixed costs. Individual producers might not know in advance whether their investment will be profitable and may therefore be reluctant to invest in upstream and downstream activities. A similar way of interpreting coordination failures in the context of development theory is linked to economies of scale: The "big push" models of economic development assume that low-income countries find themselves in a trap of low productivity created by an absence of economies of scale. A third way of combining coordination failures with the thinking of economic diversification is the cluster approach, which describes the instruments to be used by governments to foster the development of specific sectors of the economy. Again, a process of clustering of new emerging economic activities can theoretically be achieved within the private sector alone, but this is probably not the case for low-income countries. One way to overcome these coordination failures is for government to provide subsidies in case the investments are not profitable. If this expectation of a bailout in case of failure is credible, then investments will be made available and the likelihood of an unprofitable investment will be low ex-post, which will make the promised ex-ante subsidies obsolete.[25] Obviously, this method of promoting economic diversification is open to moral hazard and abuse, and such policy instruments were blamed for the Asian financial crisis of 1997. It is noteworthy that the same policies which were attributed as being the root cause of the Asian crisis, were previously praised as the reason for the preceding substantial growth rates (Raffer and Singer, 2004:148). Above all, successful policies for economic diversification cannot consist of a top-down process with a static set of rules for the private sector. As only the private sector is fully informed about existing problems, economic policy needs to establish some form of strategic collaboration and coordination between the public and the private sectors (Rodrik, 2005:20-21).

2.7 Controversies within the debate

Besides the general outcome of the debate that both diversification and specialization forces are active in the global market, but that diversification is more relevant for low-income countries, there are some contentious issues. These areas of conflict are:

- *Short-run versus long-run*

  De Benedictis (2004) argues that there are two different phenomena at work – diversification and structural change – which should be treated separately. *Structural change* means – in general terms – diversifying away from specialization in the agricultural

---

[25] This solution also applies to overcome information externalities, where the key lies in encouraging investments in the modern sector ex ante, but to rationalize production ex-post (Hausmann and Rodrik, 2003:7).

sector by entering industrial activities and eventually specializing in services. This movement emerges as a U-curve, because it includes a stage of diversification characterized by more or less equal shares of the three sectors. *Diversification*, in De Benedictis' (2004) point of view, is a short- to medium-term process of variations within sectors considered at a high level of disaggregation.

- *Production versus exports*

    De Benedictis, Gallegati and Tamberi (2003) assert that production and export forces are different but interlinked phenomena. While a country may specialize in exports in line with its comparative advantage according to Ricardian theory, this does not automatically imply that its overall degree of specialization in production will increase. This is particularly the case if home and foreign demand patterns for goods substantially differ from each other. An additional problem mentioned by De Benedictis, Gallegati and Tamberi (2003) is the use of employment shares to calculate specialization, since this implies the assumption of identical production functions across countries – i.e. same shares of employment correspond to the same production pattern – which is questionable. They therefore apply export data[26] only and confirm a U-shaped relationship between export specialization and GDP per capita for an average country, albeit with weak re-specialization. In particular, they find that the level of GDP per capita affects the level of specialization for low-income countries, while the average intercept term – and therefore the country-specific effects – is more important when explaining the level of specialization in medium- and high-income countries.

- *Absolute versus relative specialization*

    Another relevant criticism of the standard measures of absolute specialization is that they implicitly measure the deviation from a rectangular distribution, where each sector or product has the same share in total production, exports or employment. This shortcoming can be resolved by using the concept of relative diversification, which is measured by the deviation from the global production or export structure (Amiti, 1999). Given the current interdependence of national economies, the idea of relative diversification might be more meaningful than absolute diversification, because it is sensitive to global changes, reflecting the fact that the economic situation of a country changes when the global situation does, e.g. due to changes in preferences or in technology. Amiti (1999) adds that relative specialization is a better option because trade theories predict that trade liberalization will

---

[26] SITC 4-digit level, 786 total sectors but restricted to 539 manufacturing sectors, 1985-1998, 39 countries, GDP per capita in constant PPP dollars.

lead to higher differences between a country and its trading partners in terms of production and exports. Technically, absolute diversification is just a specific form of relative specialization with an unrealistic assumption about the global demand pattern. Nevertheless, De Benedictis, Gallegati and Tamberi (2007:11-12) confirm that relative diversification is higher with a rising GDP per capita, while no re-specialization at higher GDP per capita levels is evident.

- *Whole economy versus manufacturing*

  In the existing studies on the stages of diversification, the data used sometimes include the whole economy, and sometimes only the manufacturing sector. Imbs and Wacziarg (2003) use both approaches and confirm the existence of the U-curve in both settings, verifying a robustness of the U-curve. However, a manufacturing U-curve is a different phenomenon than an overall U-curve, as De Benedictis (2004) stresses, unless the former is just a part of the latter. Specifically, a U-curve in the overall economy corresponds to the traditional view of structural change as a shift from agriculture to manufacturing to services, while a U-curve exclusively within manufacturing is a different kind of structural change and thus cannot be compared. Moreover, the disaggregation of product lines differs among these three key sectors, thus comparisons must be treated with caution.

- *Other determinants of diversification*

  The discussion on the stages of diversification has also stimulated new research on the *determinants* of export specialization and diversification. In this regard, Bebczuk and Berrettoni (2006) find that richer, more efficient, more stable and more open countries tend to specialize rather than diversify their exports.[27] Surprisingly, variables typically associated with good macroeconomic performance [28] are correlated with specialization, as if entrepreneurs are seeking to take advantage of specialization-based economies of scale when macroeconomic risks are low. As countries, on average, move towards increased diversification over time despite different degrees of macroeconomic stability, Bebczuk and Berrettoni (2006:15) conclude that it is not volatility per se that drives diversification, but the "[…] desire to unburden themselves from the primary product dependence." Contributing to this thinking, Osakwe (2007) shows that infrastructure and institutions have a positive impact on diversification, while aid has a negative impact.

---

[27] Using 2-digit SITC export data (69 sectors).
[28] GDP per capita, exports to GDP ratio, investment rate, credit and infrastructure.

3. Methods and data

3.1 Measuring absolute product specialization

In this study, several absolute specialization/diversification measures are computed for each country and year and for each export and production dataset. The most commonly used measure of diversification in the literature is the Gini coefficient. A Gini value of zero equals perfect equality, whereas a value of one indicates a maximum unequal distribution, i.e. one sector or product line accounts for the total value of production or exports, whereas all other sectors (or product lines) have zero values. Geometrically, the Gini coefficient can be interpreted as the ratio of the area between the line of perfect equality and the Lorenz curve to the area of the triangle between the line of perfect equality and the diagram axes.

Figure 4 shows two stylized Lorenz curves for the income of population groups – for which the concept of the Lorenz curve was initially developed (cf. Lorenz, 1905) – where *L(1)* represents a more equal distribution of income than *L(2)*. In the subsequent analysis, sectors or products are used instead of population groups, and MVA or export values instead of income.

Figure 4 Stylized Lorenz curve
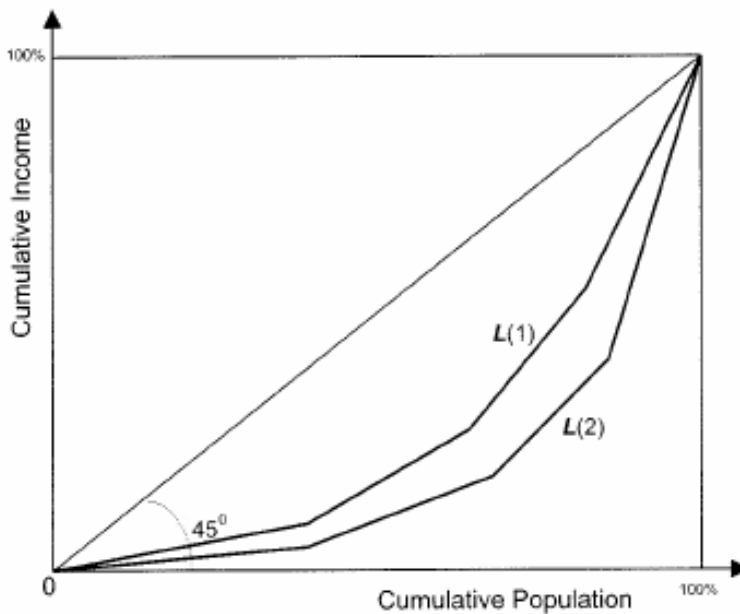
Source: Ray (1998:180).

The Gini, Theil and Herfindahl indices are calculated according to Carrère, Strauss-Kahn and Cadot (2007:6-7). For each country and year, the products are sorted in increasing order of their value added (or export value) e such that $e_k < e_{k+1}$ for all product lines k, with K denoting the

total number of product lines. Omitting country and time subscripts, the formula for the Gini coefficient is given by

$$G = \left| 1 - \sum_{k=1}^{K} (E_k - E_{k-1})(2k-1)/K \right|$$ (1)

where E_k, the cumulative value added (or export) shares, can be written as

$$E_k = \sum_{i=1}^{k} e_i \bigg/ \sum_{i=1}^{K} e_i$$ (2)

The second measure, the Herfindahl Index, can be computed according to

$$H = \frac{\sum_{k=1}^{K} (s_k)^2 - \dfrac{1}{K}}{1 - \dfrac{1}{K}}$$ (3)

where

$$s_k = e_k \bigg/ \sum_{i=1}^{K} e_i$$ (4)

is the share of product or product line k in total value added or exports. This Herfindahl Index is normalized to range between zero and one.

As a third measure, the Theil entropy index takes the form

$$T = \frac{1}{K} \sum_{k=1}^{K} \frac{e_k}{\mu} \ln\left(\frac{e_k}{\mu} \right)$$ (5)

with

$$\mu = \frac{1}{K} \sum_{k=1}^{K} e_k$$ (6)

The Hirschman Index is similar to the Herfindahl Index, which is given by

$$HI = \sqrt{\sum_{k=1}^{K} (s_k)^2}$$ (7)

Additionally, the number of active product lines, i.e. observations with a value greater than zero, is also used to capture the extensive margin of diversification.

The calculated values of export specialization are expected to be rather higher, and much higher than the levels of production specialization measures. The reason for this is related to the high

level of disaggregation of export data, i.e. the large number of product lines in which trade data gets reported. In such a disaggregated dataset as used in this study, a significant number of export lines exhibit very small trade values, while a small number of export lines account for most of total exports, particularly in the case of developing countries (see also Carrere, Strauss-Kahn and Cadot 2007:7-8). For the purpose of this study, however, it is the change and not the actual level of specialization that is of interest.

3.2 Measuring relative product specialization

Absolute specialization measures as described above describe deviations from uniform distributions. To capture deviation from the global average distribution, certain *relative* specialization measures can be applied. De Benedictis and Tamberi (2004) use the Balassa Index (BI) of revealed comparative advantage:

$$BI = \frac{e_{c,k}}{e_c} \bigg/ \frac{e_{w,k}}{e_w}$$

(8)

where e denotes exports, c denotes a specific country, w the whole world and k a specific product. The BI therefore measures the ratio of each sector's share within a country relative to the global share of that sector within total exports.

The BI can also be written as

$$BI = \frac{e_{c,k}}{e_{w,k}} \bigg/ \frac{e_k}{e_w}$$

(9)

which can be interpreted as the ratio of a country's share of sectoral export to the country's share in world exports. A BI above one indicates that the sector has a comparative advantage. To calculate a country's overall BI, the respective medians of the sectoral BIs are more suitable than their means due to the skewness of a country's BIs. A high median BI implies that a country has a comparative advantage in a large share of sectors.

Two additional measures of overall relative specialization are calculated following De Benedictis, Gallegati and Tamberi (2007). First, after sorting the observations of each country by the BI in ascending order, the relative Gini coefficient[29] is given by

---

[29] Also called "country Gini" in some studies.

$$relGini = \frac{\sum_{k=1}^{K-1} (p_k - q_k)}{\sum_{k=1}^{K-1} p_k} \qquad (10)$$

where q_k and p_k are cumulated shares of the numerator and denominator of BI (equation 8), i.e. the cumulated national and world sectoral shares. Second, the Theil Index can also be modified to account for relative specialization:

$$relTheil = \sum_{k=1}^{K} \left[\frac{e_k}{e} \ln\left(\frac{e_k/E}{f_k/F} \right) \right] \qquad (11)$$

where e_k denotes country exports in sector k, f_k is world exports in sector k, E is total exports in the country and F is total world exports. The relative Theil Index is therefore a weighted sum of the logarithms of sectoral Balassa Indices, with weights represented by country sectoral shares.

Al-Marhubi (2000), UNCTAD (2008) and Albaladejo (2007) use the sum of the differences of the national and world sectoral shares as a measure of relative specialization:

$$DI_c = \frac{1}{2} \sum_k |h_{k,c} - h_k| \qquad (12)$$

where $h_{k,c}$ is the share of commodity k in total exports of country c, and h_k is the share of commodity k in world exports.

The general parametric fixed effects specification[30] to verify the existence of a U-curve is to regress specialization on GDP per capita and squared GDP per capita is

$$Spec_{ct} = \alpha_0 + \alpha_c + \beta_1 \cdot GDPpC_{ct} + \beta_2 \cdot GDPpC_{ct}^2 + \varepsilon_{ct} \qquad (13)$$

where β_1 is expected to be negative and β_2 to be positive, as the sum of a negative linear function and a positive squared function equals a U-curve, though only when the turning point lies within the range of GDP per capita. The turning point is calculated by setting the first derivative of (13) with respect to GDP per capita to zero and rearranging it:

$$GDPpC^{turn} = -\frac{\beta_1}{2\beta_2} \qquad (14)$$

---

[30] For the pooled panel case, the country fixed effects α_c are assumed to be jointly zero, while the between estimator shows the coefficients of regressing time-invariant means of $Spec$ on $GDPpC$ and $GDPpC^2$.

3.3 Data used in this study

Overall, this study brings together four different datasets to measure specialization, with one dataset for industrial production and three datasets for exports. Data on MVA are taken from UNIDO (2006) using the ISIC Revision 2 nomenclature at the 3-digit level, corresponding to 28 different manufacturing sectors, where data are available from 1963 to 2003.[31] Export data are obtained from the UN Comtrade (2008) database at the 5-digit level (935 non-zero product lines) of the SITC Revision 1 classification, and the 6-digit level (5,018 non-zero product lines) of the HS 1989/92 system, as well as the dataset of Feenstra et al. (2005), who created a database at SITC Revision 2 (4-digit level, 1,069 non-zero product lines) that has been corrected for errors by comparing export and import data and by including national databases. Each of these datasets has particular advantages and disadvantages: The HS dataset has the highest level of disaggregation, but covers the shortest period (1989-2005), the Feenstra dataset has been corrected for errors and covers a longer period (1962-2000), but is much more aggregated, while the SITC dataset covers the longest time span (1962-2006).

For MVA data, the dataset is modified so that the number of sectors available over time for each country is constant, which requires excluding observations for some sectors when these were not available for a given country for all years (cf. Imbs and Wacziarg, 2003:3). This approach needs to be taken since missing values do not necessarily indicate zero values, but may represent non-reported entries, because UNIDO data are based on data obtained from national surveys which do not always cover all economic activities. However, some countries report an aggregation of specific sectors – for example, food products and beverages – into one larger sector. These combinations of aggregations differ between countries, and there are several ways to deal with this issue. For our study, all aggregated sectors were deleted, leaving us with countries that do not aggregate sectors at all and hence report on all 28 sectors.[32] Second, as certain combinations of sectors appear in a substantial number of countries, one can aggregate these sectors in *all* countries. Following the combinations suggested by Koren and Tenreyro (2004:15-16), the dataset we are left with has 19 sectors. This latter method results in more observations than when using 28 sectors only, but it also means that the data of countries reporting more than 19 sectors are contracted, hence, some information is lost.

---

[31] In this dataset, data from countries that already report their data in ISIC Revision 3 have been converted to ISIC Revision 2 data (Yamada, 2005).

[32] Imbs and Wacziarg (2003) include countries with 27 reported sectors in their fixed effects regressions.

Handling missing values is different for export data, because most missing values in fact represent zero exports.[33] Contrary to employment and MVA data, missing values have to be replaced by zero values to obtain a rectangular dataset (cf. Carrère, Strauss-Kahn and Cadot, 2007:7).

Data on GDP per capita are derived from the World Bank (2008), using both constant 2000 US dollars to exclude effects stemming from inflation and constant 2005 international dollars at Purchasing Power Parity (PPP) for better comparability across countries.[34]

A number of caveats of this dataset should be noted. First, all data obviously only measure formal economic activity, disregarding the informal sector. Second, MVA data are based on national surveys and estimates, hence perhaps not capturing the full (formal) economic activity of an economy in contrast to export data, which are directly measured. Third, the dataset is limited to the years up to 2007, i.e. right before the start of the recent financial crisis. Further research could make of use of data after 2007, though the general conclusions are expected to hold, in particular when potential crisis effects have been accounted for.

4.   Results

4.1   Production

4.1.1   Descriptive statistics

For MVA data, the absolute Gini, Theil, Herfindahl and Hirschman indices are computed as described in Section 3. Table 1 shows the ten most diversified and ten least diversified countries according to their Gini value in 1994.[35] Portugal, Austria and Argentina are the most diversified countries, while Kuwait, Senegal and Gabon are the most specialized countries. This is somehow consistent with the U-curve hypothesis, as the list is not entirely headed by highest-income countries, although high-income countries, like Austria and the United Kingdom, still rank relatively high. However, as expected, the most specialized countries are low-income countries.

---

[33] Gleditsch (2002) conducts an in-depth analysis on the issue of missing values in trade datasets: some missing export values are actually non-reported positive values, which can be obtained from the corresponding import values, as being done in the Feenstra dataset. A minority of missing values can be replaced by positive values through time-series methods such as interpolating and estimating lags and leads.

[34] The availability of GDP data in constant PPP dollars is smaller than in constant US dollars, however, so this study concentrates on using constant 2000 US dollars, while the results for PPP dollars are available upon request.

[35] After this year, the number of countries that report all 28 sectors declines rapidly per year.

Table 1   Ten most and ten least specialized countries, MVA, 1994

Rank	Country	Gini	Theil	Herfindahl	Hirschman
1	Portugal	0.44	0.33	0.03	0.25
2	Austria	0.46	0.35	0.03	0.25
3	Argentina	0.48	0.39	0.03	0.26
4	Korea, Rep. of	0.49	0.40	0.04	0.27
5	United Kingdom	0.49	0.40	0.03	0.26
6	Canada	0.49	0.42	0.03	0.26
7	Turkey	0.50	0.41	0.03	0.26
8	United States	0.50	0.42	0.03	0.26
9	Macedonia, FYR	0.51	0.43	0.03	0.26
10	Chile	0.51	0.47	0.05	0.28
32	Ethiopia	0.67	0.84	0.08	0.34
33	Honduras	0.69	0.93	0.11	0.38
34	Oman	0.70	0.91	0.09	0.35
35	Panama	0.72	1.09	0.15	0.43
36	St. Lucia	0.73	1.05	0.11	0.37
37	Ecuador	0.76	1.26	0.18	0.46
38	Iceland	0.76	1.34	0.25	0.52
39	Gabon	0.77	1.19	0.11	0.38
40	Senegal	0.79	1.33	0.19	0.47
41	Kuwait	0.80	1.66	0.38	0.64

Source: Author's own calculation based UNIDO (2006) database.

Table 2 presents summary statistics of the diversification measures and the economic development measures. The variance of the Gini coefficient is quite high, ranging from 0.36 to 0.94, which indicates that the sample includes economies that are very specialized in their production as well as those with a highly diversified production structure. The population size of countries ranges from 140,000 to over one billion. In terms of annual constant GDP per capita, the sample ranges from very low (US$92) to very high figures (US$45,000).

Table 2   Summary statistics, MVA

Variable	Observations	Mean	Std. Dev.	Min	Max
Gini	1,869	0.58	0.11	0.36	0.94
Theil	1,869	0.65	0.35	0.21	2.77
Herfindahl	1,869	0.08	0.08	0.02	0.75
Hirschman	1,869	0.32	0.09	0.23	0.87
Population	1,866	4.17e+07	1.13e+08	1.39e+05	1.05e+09
GDP per capita	1,757	7,012	7,887	92	45,391
GDP per capita (PPP)	866	11,768	9,341	388	59,893

Source: Author's own calculation based UNIDO (2006) and World Bank (2008) databases.

The pair-wise correlation coefficients between the specialization measures in Table 3.1 indicate that all four measures are highly correlated among each other, indicating that they quantify the same phenomenon. However, the graphical correlation matrix (not shown) reveals that the

specialization measures are not correlated in a linear manner. For example, the Theil Index discriminates more finely between countries with lower specialization, while the Theil and especially the Herfindahl Index discriminate more between countries with higher specialization, and computing the correlation of ranks is therefore more appropriate. Table 3.2 presents rank correlations [36] instead of correlations between values, and indeed indicates a very high correlation between the specialization measures.

Table 3.1 Correlation table, MVA					Table 3.2 Rank correlation table, MVA				
	Gini	Theil	Herfindahl	Hirschman		Gini	Theil	Herfindahl	Hirschman
Gini	1.00				Gini	1.00			
Theil	0.96	1.00			Theil	1.00	1.00		
Herfindahl	0.83	0.95	1.00		Herfindahl	0.97	0.98	1.00	
Hirschman	0.90	0.98	0.98	1.00	Hirschman	0.97	0.98	1.00	1.00

Source: Author's own calculation based UNIDO (2006) database.

4.1.2 Non-parametric results

Figure 5.1 shows a pooled panel scatter plot of export diversification measured by the Gini coefficient and a country's income per capita level, i.e. all years and countries are in one plot, for all countries that report 28 sectors together with a simple locally weighted scatterplot smoothing (Lowess) curve.[37] The same relationship is presented in Figure 5.2 for the sample of countries that report data for 19 industrial sectors. Countries with the lowest GDP per capita are relatively specialized, and this specialization decreases rapidly with economic growth. The decrease in specialization then becomes flatter and turns towards increased specialization at high levels of GDP per capita. This supports the assumption of a U-shaped relationship between concentration in production and economic development, although the shape of the curve resembles an "L" rather than a "U" shape. However, any preliminary conclusions have to be treated with caution, because the Lowess curve smoothes over the pooled panel of countries and years.

---

[36] The Spearman rank correlation was used throughout this study.
[37] Observations from all available years are included, but countries with a population of less than one million people are excluded to avoid specialization effects that are purely a result of small country size.

Figure 5.1 Specialization of MVA, 28 sectors

Figure 5.2 Specialization of MVA, 19 sectors

Source: Author's own calculation based UNIDO (2006) and World Bank (2008) databases.

Figures 6.1 (for 28 sectors) and 6.2 (for 19 sectors) show the relationship between specialization and GDP per capita measured in constant PPP dollars[38], a measure commonly used in the existing literature (in particular, Imbs and Wacziarg, 2003). Although the resulting Lowess curve has a much more distinctive U-shape when compared with the results using constant US dollars, the upward-sloping part seems to be driven by a small number of highly specialized high-income countries. Note that the coverage of PPP dollars is lower than for constant US dollars.

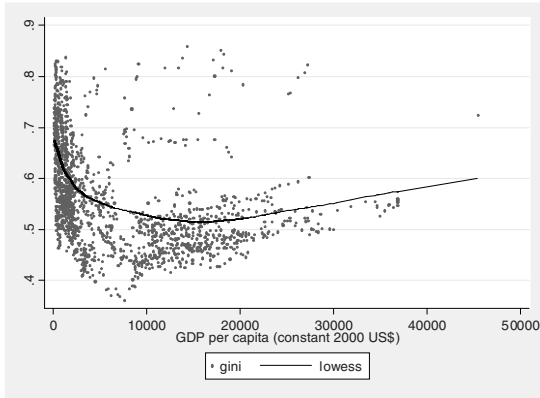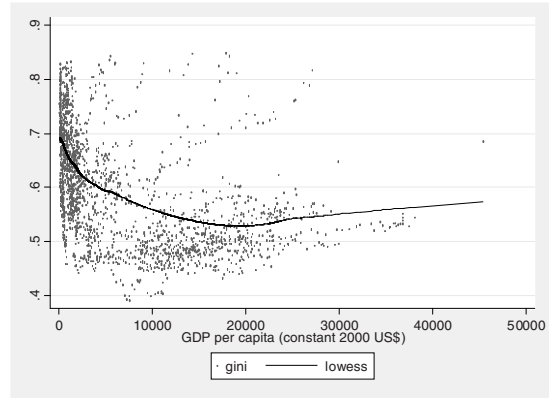**Figure 6.1   Specialization of MVA, 28 sectors, PPP dollars**
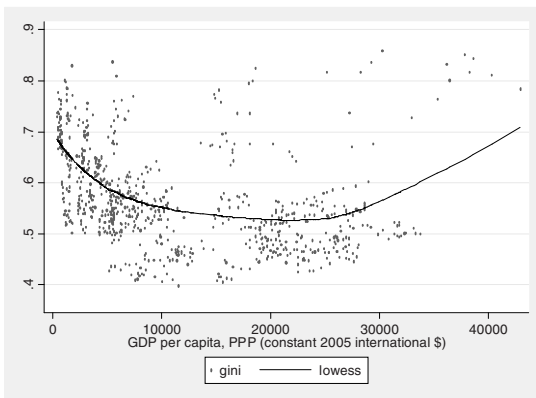
Figure 6.2   Specialization of MVA, 19 sectors, PPP dollars
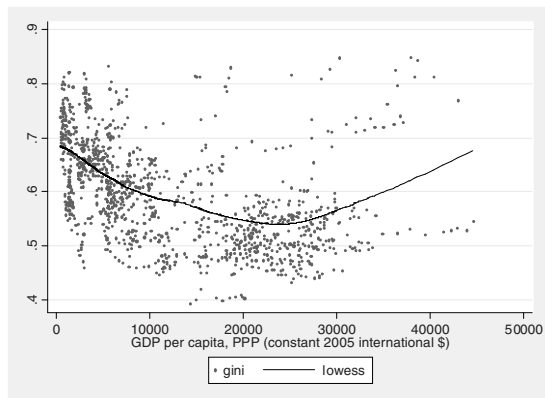
Source: Author's own calculation based UNIDO (2006) and World Bank (2008) databases.

---

[38] United Arab Emirates is excluded when using constant PPP dollars, as it is an extreme outlier.

When countries are classified according to their income group[39], an interesting picture emerges (Figure 7). The observations of low-income countries are scattered across a wide range of specialization values, ranging from below 0.5 to above 0.8. Lower middle-income countries are similarly spread, but with fewer observations at higher levels of diversification. Upper-middle income countries seem to be much more diversified on average than lower middle-income countries, and high-income OECD countries are also more highly diversified, with a slight upward trend. Only high-income non-OECD countries do not fit into this picture. These observations are scattered across the whole spectrum, including many observations in the high-income and high specialization group, which follows from the fact that many high-income non-OECD countries are oil exporters and consequently achieve high income levels without economic diversification.

Figure 7   **Specialization of MVA, 28 sectors, highlighted country groups**
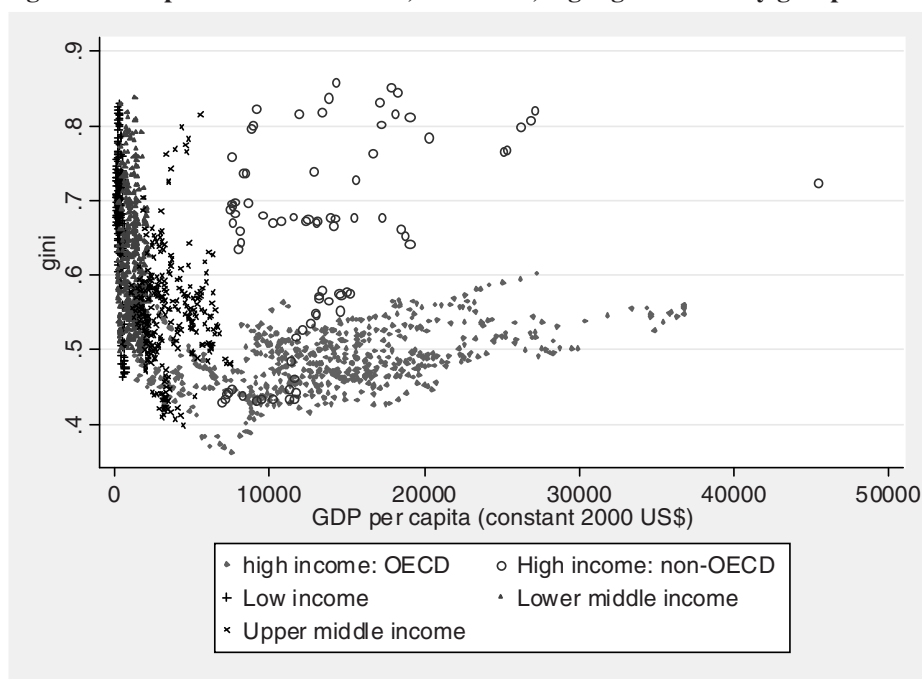
Source: Author's own calculation based UNIDO (2006) and World Bank (2008) databases.

Figure 8.1 presents the levels of specialization of six countries – Germany, Chile, the Republic of Korea, Mexico, Ghana and the USA[40] – over time. Nigeria is the most specialized country within this group, although it was almost as diversified as Chile around 1970. The Republic of

---

[39] Income groups are classified according to the World Bank (2008).
[40] These particular countries were mainly selected for illustrative purpose: The USA as the world's largest economy, Mexico as a developing Latin American country neighbouring the USA, Chile as another Latin American country with a different economic structure than Mexico, the Republic of Korea as a former low-income country that has transformed into a high-income country, and Germany and Ghana to include one country from each continent.

Korea was relatively specialized in the first half of the 1960s, then diversified until the 1980s, and showed a strong trend towards re-specialization in the 1990s, thus following the U-curve relationship, given that the Republic of Korea's GDP per capita increased during this period. Germany experienced a movement in the opposite direction, as it was more diversified than the Republic of Korea in the 1960s, but became more specialized in the 1970s, which might represent the upward-sloping part of a U-curve. The USA also shows a slight trend towards specialization. Ghana's economy followed a distinct U-curve relationship as well, but not in the sense of Imbs and Wacziarg (2003), as its GDP per capita actually *declined* from US$270 to US$198 between the 1960s and 1985, while its level of specialization fell in the 1960s and then rose again until the 1980s, thus indicating what might be called a "backward U-curve". Figures 8.2 and 8.3 present specialization on the left vertical axis (solid line) and GDP per capita on the right vertical axis (dotted line) for the Republic of Korea and Ghana. In a panel data regression, both countries would strengthen the significance of a U-curve relationship, although the underlying dynamics differ significantly – economic growth in the Republic of Korea and economic decline in Ghana.

Figure 8.1 Specialization of MVA over time
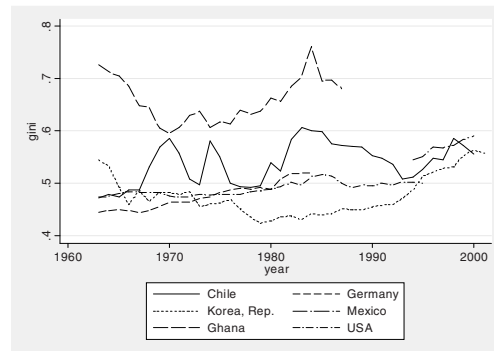
Figure 8.2 GDP per capita and specialization of MVA over time, Ghana
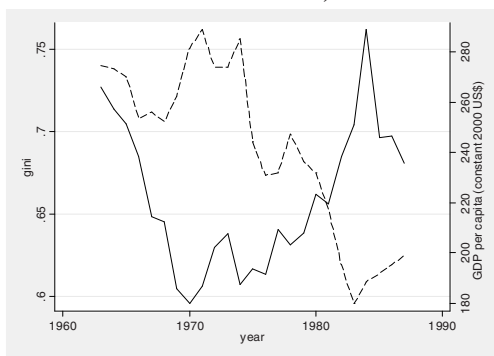
Figure 8.3 GDP per capita and specialization of MVA over time, Republic of Korea
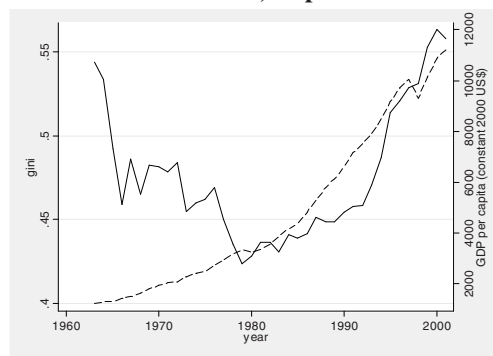
Source: Author's own calculation based UNIDO (2006) and World Bank (2008) databases.

To summarize, only vague conclusions can be drawn on the existence of a distinct U-shaped relationship between specialization and economic development, but it can be confirmed that – when analysing a pooled panel – rising GDP per capita goes hand in hand with a *diversification* of the productive activities for most countries, not with *specialization*. A slight trend towards specialization can be observed at the highest income levels, though this is potentially driven by outliers. The within-country variation shows a marginal trend of countries towards higher levels of specialization.

To further investigate the patterns of specialization and diversification, and to verify the assumption of a U-curve beyond the tentative conclusions based on non-parametric pooled panel analysis, parametric panel data regression methods are applied in the following chapter.

4.1.3   Regression results

The analysis in the previous chapter is based on a non-parametric analysis of pooled panel plots, which do not distinguish whether two different observations are two different countries or a single country in two different periods, hence without the ability to account for time-invariant heterogeneity. Nevertheless, the presented preliminary findings indicate a non-linear relationship between specialization and GDP per capita. The following regression analysis aims at verifying the U-curve hypothesis using parametric panel data regression methods, which include a linear and a squared term of GDP per capita to verify the existence and determine the location of the proposed U-curve, controlling for time-invariant fixed effects. A U-curve can said to be present if (i) a negative coefficient on the linear term, (ii) a positive coefficient on the squared term, and (iii) a turning point that lies within the data range is observable.

The parametric regression specification that corresponds to the non-parametric regression presented above would be a simple pooled regression. In this case, the U-curved relationship is indeed significant at the 1 percent level for all four measures of specialization, both levels of aggregation (28 and 19 sectors), and both measures of GDP per capita (Table 4, column 1 for the Gini coefficient and column 2 for the Herfindahl index)[41]. The same results are obtained when the Hirschman and Theil coefficients are used as measures of specialization.[42] The U-curve is also significant when only a single year is included. When constant PPP dollars are used, the results remain the same. The turning point of the U-curve, when significant, lies at a

---

[41] Note that the estimated coefficients appear to be small, because GDP per capita is measured by the level of GDP per capita in dollars.
[42] Results are available upon request.

GDP per capita level between US$17,000 and US$22,000, so re-specialization indeed only takes place at very high economic levels.

Table 4   Regression results for specialization of MVA, 28 and 19 sectors

ISIC 28 sectors	(1) Gini Pooled	(2) Herfindahl Pooled	(3) Gini FE	(4) Herfindahl FE	(5) Gini BE	(6) Herfindahl BE
GDP per capita	-1.6e-05***	-5.2e-06***	9.0e-07	8.8e-07*	-1.5e-05***	-4.8e-06***
	(8.2e-07)	(5.0e-07)	(1.0e-06)	(5.1e-07)	(3.0e-06)	(1.8e-06)
GDP per capita squared	4.7e-10***	1.4e-10***	4.5e-11*	-1.3e-11	3.8e-10***	1.2e-10**
	(3.5e-11)	(1.8e-11)	(2.3e-11)	(1.1e-11)	(9.1e-11)	(5.4e-11)
Constant	6.2e-01***	8.8e-02***	5.5e-01***	6.3e-02***	6.4e-01***	9.6e-02***
	(3.4e-03)	(2.1e-03)	(4.7e-03)	(2.8e-03)	(1.5e-02)	(8.9e-03)
Observations	1,637	1,637	1,637	1,637	1,637	1,637
Number of countries	82	82	82	82	82	82
R-squared	0.26	0.08	0.08	0.07	0.25	0.08
Turning point	16,896	18,028			19,647	19,587
ISIC 19 sectors						
GDP per capita	-1.7e-05***	-1.2e-05***	3.0e-06***	9.3e-07*	-1.8e-05***	-1.2e-05***
	(6.6e-07)	(6.1e-07)	(8.9e-07)	(5.2e-07)	(3.1e-06)	(3.2e-06)
GDP per capita squared	4.5e-10***	2.9e-10***	6.1e-12	6.0e-12	4.5e-10***	2.8e-10**
	(2.7e-11)	(2.3e-11)	(2.0e-11)	(1.2e-11)	(1.1e-10)	(1.1e-10)
Constant	6.6e-01***	1.7e-01***	5.7e-01***	1.2e-01***	6.8e-01***	1.9e-01***
	(2.8e-03)	(3.1e-03)	(4.2e-03)	(2.7e-03)	(1.3e-02)	(1.3e-02)
Observations	2,155	2,155	2,155	2,155	2,155	2,155
Number of countries	98	98	98	98	98	98
R-squared	0.34	0.21	0.23	0.15	0.34	0.21
Turning point	18,855	20,139			19,532	21,996

Robust standard errors in parentheses (non-robust standard errors for BE)
*** p<0.01, ** p<0.05, * p<0.1
Source: Author's own calculation based UNIDO (2006) and World Bank (2008) databases.

As country-specific unobserved characteristics exist, it is appropriate to consider a fixed effects panel data estimation. Moreover, this method is relevant as it describes how the production structure of an "average" country evolves alongside the economic development process.[43] The results show a less distinct picture than for the pooled regression (columns 3 and 4 in Table 4). When using 28 sectors, no U-curved relationship can be observed. The upward-sloping part is only significant for the Gini-coefficient (column 3), and only at the 10 percent level. The coefficient on the linear term even has an unexpected positive sign of GDP per capita for the Herfindahl (column 4) and Hirschman Index, though only at the 10 percent significance level. When using 19 sectors, thereby increasing the countries with applicable MVA data from 74 to 91 (lower part of Table 4), the unexpected positive sign on the linear term becomes significant

---

[43] The Hausman test rejects the null hypothesis of no correlation between the fixed effects and the explanatory variables, thus the fixed effects model is preferred over a random effects model.

at the 1 percent significance level for the Gini (column 3) and Theil Indices, and at the 10 percent significance level for the Herfindahl (column 4) and Hirschman Index. When constant PPP dollars are used, the within-country U-curve does not become more significant.

Since the results above seem to indicate that the U-curve in the pooled panel is not mainly driven by within-country effects, it is of interest to investigate the between-country effects, i.e. the correlation between countries' means of specialization and GDP per capita. The between effects estimates (columns 5 and 6) show a highly significant (i.e. at the 1 percent level in most cases) U-curve for all measures of diversification and for both levels of aggregation (28 and 19 sectors), meaning that the U-curve is, to a large extent, a given time-invariant global structure, and not primarily a development path of countries. This also holds for other measures of diversification and when using constant PPP dollars.

This study has thus far concentrated on relatively aggregated data with a maximum of 28 different sectors. The level of aggregation is likely to have significant influence on the observed shape of specialization. Given the level of aggregation, specialization may occur within rather than between sectors, but only the latter form of specialization is observed in the data. As a result, the data observed thus far may not provide a true reflection of the actual levels of specialization. Value added data are not available at more disaggregated levels for low- and middle-income countries, but trade data are available at various aggregation levels for a large set of countries. The following chapter therefore employs export data to further investigate the patterns of diversification.

4.2 Exports (absolute specialization)

4.2.1 *Descriptive statistics*

Looking at export patterns might reveal more about the specialization/diversification path of countries, as trade data are available at a much higher level of disaggregation and for a larger number of countries than production data. In addition, trade data are available for agricultural as well as manufactured products, allowing for a broader analysis compared to using MVA as in the previous section.[44]

The country with the most diversified export structure in 2005, measured at the highest disaggregation level (HS 6-digit) is Italy, followed by the USA and Germany, with the remaining seven countries being OECD members plus China. Out of the ten least diversified

---

[44] Results for manufactured exports alone are available upon request.

countries, African countries occupy the bottom seven positions, with Mauritania, Gabon and Sudan at the lower end of the scale (Table 5.1).

This is consistent with the idea that high-income countries are more diversified than low-income countries, though a trend towards re-specialization cannot be ruled out, as the most diversified countries do not fully correspond to countries with the highest levels of GDP per capita.

Table 5.1   Ten most and ten least specialized countries, HS 6-digit exports, 2005

Rank	Country	Gini	Theil	Herfindahl	Hirschman	Number of export lines
1	Italy	0.83	1.71	0.00	0.06	4,746
2	Germany	0.84	1.96	0.01	0.08	4,666
3	United States	0.85	2.01	0.01	0.08	4,831
4	France	0.86	2.09	0.01	0.09	4,683
5	Spain	0.86	2.20	0.01	0.09	4,765
6	China	0.87	2.18	0.01	0.09	4,743
7	Belgium	0.88	2.43	0.01	0.12	4,750
8	Netherlands	0.88	2.38	0.01	0.11	4,741
9	United Kingdom	0.88	2.47	0.01	0.11	4,756
10	Austria	0.89	2.16	0.01	0.07	4,520
107	Oman	1.00	7.21	0.51	0.72	985
108	Azerbaijan	1.00	6.67	0.33	0.57	1,122
109	Yemen	1.00	7.60	0.72	0.85	1,111
110	Central African Republic	1.00	6.52	0.19	0.43	115
111	Mali	1.00	7.26	0.48	0.69	488
112	Burundi	1.00	7.17	0.40	0.63	292
113	Algeria	1.00	7.03	0.34	0.58	935
114	Sudan	1.00	7.62	0.70	0.84	212
115	Gabon	1.00	7.66	0.69	0.83	730
116	Mauritania	1.00	7.88	0.60	0.78	15

Table 5.2 presents the most important summary statistics for the HS 6-digit and the SITC 5-digit data. The HS 6-digit dataset covers the widest span of export lines, with up to 4,976 recorded export lines, but the data only go back to 1988. The SITC 5-digit dataset reports a maximum of 921 different export lines, but with data going back to 1962.

Table 5.2   Summary statistics of specialization in exports, HS 6-digit and SITC 5-digit

HS6-digit	Mean	Std. Dev.	Min	Max
Number of export lines	2,363	1,620	9	4,976
Gini	0.96	0.04	0.79	1.00
Theil	4.56	1.69	1.59	8.47
Herfindahl	0.14	0.19	0.0025	0.9873
Hirschman	0.31	0.22	0.0521	0.9936
Population	4.27E+07	1.47E+08	40,740	1.31E+09
GDP per capita	7,505	9,772	100	54,178
GDP per capita (PPP)	11,811	11,878	319	73,277
SITC 5-digit				
Number of export lines	388	279	1	921
Gini	0.9576	0.0488	0.7698	0.9989
Theil	4.00	1.50	1.26	6.85
Herfindahl	0.24	0.29	0.0060	1
Hirschman	0.42	0.26	0.0842	1
Population	3.12E+07	1.10E+08	40740	1.31E+09
GDP per capita	6,332	8,305	92	54,178
GDP per capita (PPP)	10,956	11,361	319	79,032

Source: Author's own calculation based UN Comtrade (2008) and World Bank (2008) databases.

Table 6.1 presents the pair-wise correlation between the specialization/diversification measures for the SITC dataset. Some pairs are relatively weakly correlated, e.g. the correlation between the Gini and the Herfindahl Index is only 0.57 compared with 0.85 between the Gini and the Theil Index. As with the production data, however, the correlations (not shown in the table) show a high degree of non-linearity between some indices. At the higher level of disaggregation, the export shares of many sectors are close to zero, meaning that the squared shares are even closer to zero, and consequently, the entire sum of shares is pushed towards zero. For this reason, it is appropriate to look at the pair-wise rank correlations (Table 6.2), which are indeed much higher than the standard pair-wise correlations, thus revealing the similarity of the measures. In Tables 6.1 and 6.2, the number of export lines is negatively correlated with the other measures, as expected, since more export lines correspond to less specialization, and therefore lower values of the Gini, Theil, Herfindahl and Hirschman indices.

Table 6.1   Correlation between export specialization measures, SITC 5-digit

	Gini	Theil	Herfindahl	Hirschman	Number of export lines
Gini	1.00				
Theil	0.85	1.00			
Herfindahl	0.57	0.89	1.00		
Hirschman	0.70	0.96	0.97	1.00	
Number of export lines	-0.82	-0.76	-0.51	-0.61	1.00

Table 6.2   Rank correlation between export specialization measures, SITC 5-digit

	Gini	Theil	Herfindahl	Hirschman	Number of export lines
Gini	1.00				
Theil	0.99	1.00			
Herfindahl	0.95	0.98	1.00		
Hirschman	0.95	0.98	1.00	1.00	
Number of export lines	-0.85	-0.78	-0.71	-0.71	1.00

Source: Author's own calculation based UN Comtrade (2008) and World Bank (2008) databases.

4.2.2   Non-parametric results

Figures 9.1-9.4 plot the Gini coefficient of the pooled panel against GDP per capita, along with a non-parametric Lowess. The Lowess curve clearly shows a U-curve for the 5-digit data (Figure 9.1), which is robust to lower bandwidths of the Lowess procedure. However, the observations are quite scattered, with a large number of observations some distance from the Lowess curve. The implication of this result is that the U-curve is indeed a feature of the data, but does not entirely describe the variation of the Gini coefficient. For the Feenstra dataset (Figure 9.2), the upward-sloping part is even less distinct, and for HS 6-digit data (Figure 9.3), which covers the smallest time range, the upward-sloping part seems to be non-existent. A large number of observations are concentrated in the upper-left part, i.e. countries with low GDP per capita and a high degree of specialization. On the right side of the figure, the density is very low, and it seems reasonable to assume that the observed upward-sloping part might be driven by outliers. When applied to the SITC 5-digit dataset, the outlier detection procedure developed by Hadi (1992, 1994) suggests that 31 observations are outliers. Figure 9.4 highlights those observations if they belong to countries with a population of above one million inhabitants, along with a new Lowess curve, which excludes these observations and which now shows no increase in specialization at higher levels of GDP per capita.

Figure 9.1   Export specialization, SITC 5-digit   **Figure 9.2   Export specialization, HS 6-digit**
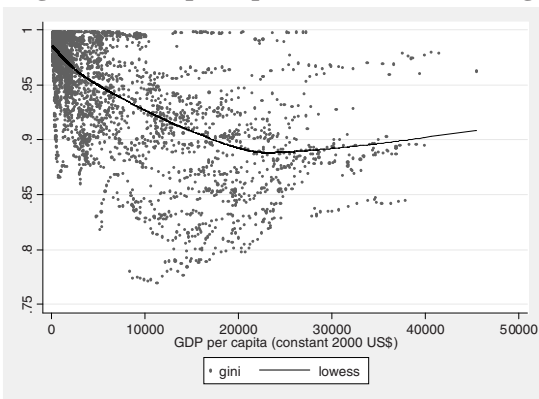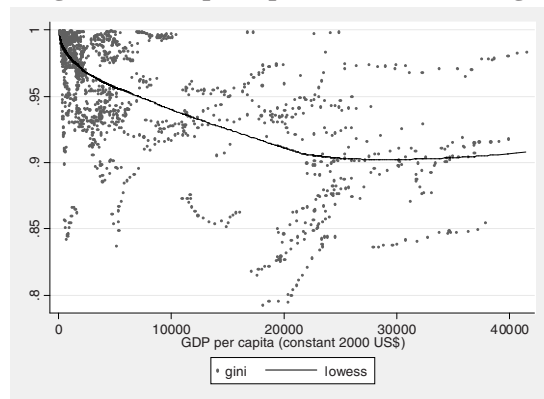
Figure 9.3   Export specialization,
Feenstra 4-digit
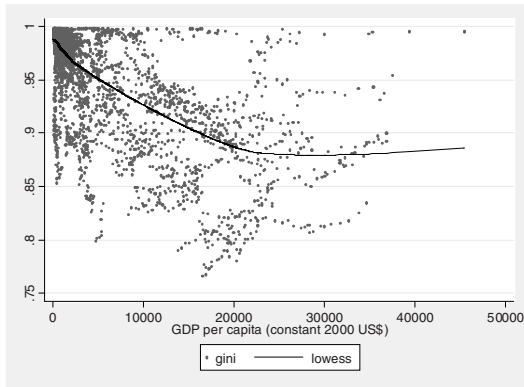
Figure 9.4   Export specialization, SITC 5-digit,
excl. outliers
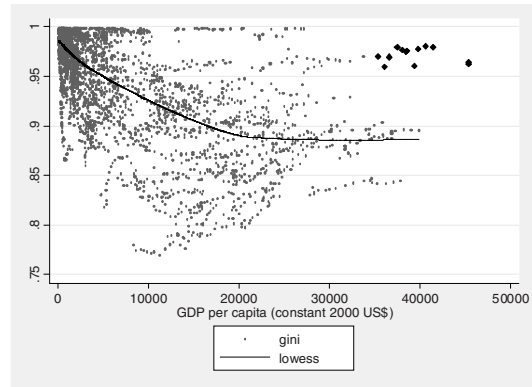
Source: Author's own calculation based UN Comtrade (2008), Feenstra et al. (2005) and World Bank (2008) databases.

The downward-sloping part of the supposed U-shape is also visible when using constant PPP dollars (Figure 10.1), but the upward-bending part again seems to be driven by very few observations.

As the observations on the upper-left part of the scatter plot are very dense, Figure 10.2 shows the same relationship as Figure 9.1, but uses a logarithmic scale on the horizontal axis to allow for a better visual observation of low-income countries. It reveals that some countries remain specialized, while others experience a decrease in specialization with a rising GDP per capita.

Figure 10.1   Export specialization, SITC 5-digit,
PPP dollars

Figure 10.2   Export specialization, SITC
5-digit, logarithmic scale

Source: Author's own calculation based UN Comtrade (2008) and World Bank (2008) databases.

To provide information on the location of particular countries, Figure 11.1 shows country abbreviations as data markers for the year 2000 for the SITC 5-digit dataset. Countries which are highly specialized and have a very high GDP per capita are typically oil-exporting countries, such as Kuwait (KWT), United Arab Emirates (ARE) or Norway (NOR). Thus, the re-specialization part might not be a globally valid stylized fact of economic development; instead, it only reflects the ability of oil-abundant countries to reach high levels of GDP per capita without diversifying their economies.

Figure 11.2 distinguishes between five different income groups, with the plot showing slightly less distinction between the different groups than the corresponding plot for MVA. Low- and middle-income countries are concentrated on the upper left side of the panel, while high-income countries are broadly spread across the plot. Contrary to the MVA data, OECD countries overlap much more with the non-OECD countries.

Figure 11.1 Export specialization, SITC 5-digit, 2005

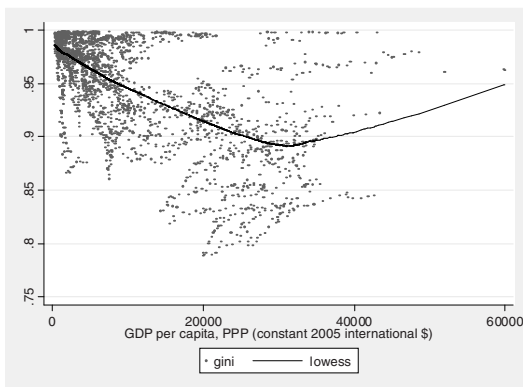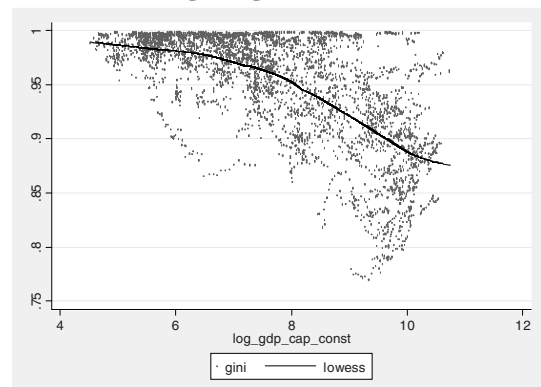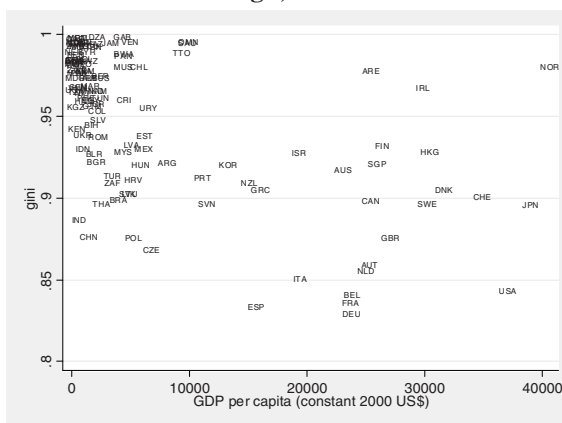**Figure 11.2 Export specialization, with country groups**
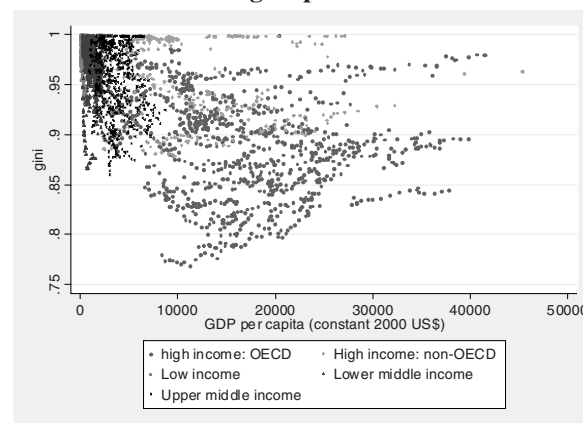
Source: Author's own calculation based UN Comtrade (2008) and World Bank (2008) databases.

Figure 12 illustrates the evolution of specialization over the period 1963 to 2007 for six selected countries – Chile, Germany, Republic of Korea, Mexico, Nigeria and the USA. Of these, Germany is the most diversified economy, but with a trend towards more specialization. The USA began with a much higher level of specialization than Germany, but its economy diversified in the 1960s and at the end of the 1980s, leaving its specialization level similar to that of Germany. Mexico experienced different phases of specialization and diversification, with a slight increase in specialization during the past ten years. The Republic of Korea and Chile

started from almost the same level in 1963, but underwent different evolutions of specialization: The Republic of Korea increased its level of diversification until the mid-1990s when a trend towards re-specialization occurred. Chile also showed a slight trend towards diversification for much of the period, but only on a small scale. Ghana shows a constant trend towards specialization, but – as is the case for MVA data – it should be noted that Ghana faced an economic decline during that period. This pattern therefore does not denote a re-specialization trend, but rather specialization with disadvantageous economic consequences.

Figure 12   Export specialization of over time
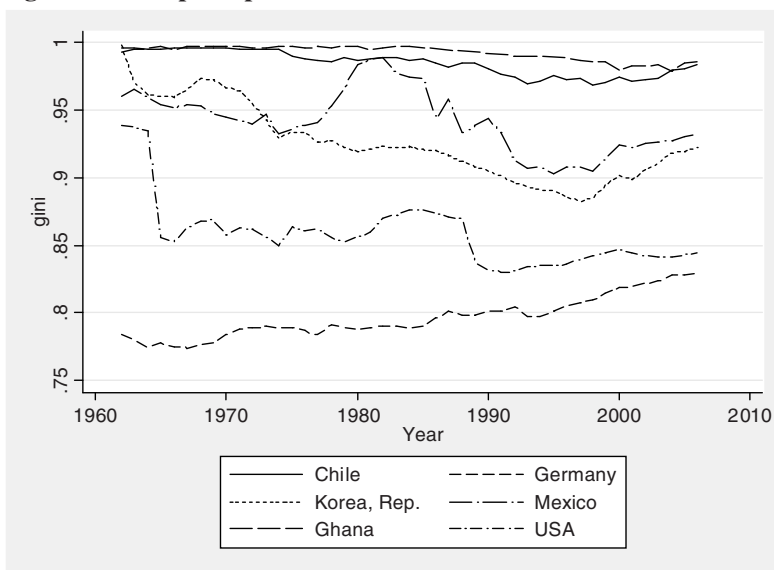
To summarize, although the observations are not clearly aligned along a U-shaped curve as indicated by some studies, some conclusions can already be drawn from the simple scatter plots and non-parametric curves presented. Countries with a low GDP per capita have a highly specialized export structure which, on average, becomes more diversified with a growing GDP per capita. The trend to diversify decreases at higher levels of GDP per capita, and even slightly reverses at the highest levels of GDP per capita. As with the pattern when using production data, the trend towards re-specialization for high-income countries is not symmetric to the specialization trend of low-income countries, so the resulting shape might be called an "L-curve" instead of a "U-curve". Including countries with a population of less than one million inhabitants does not lead to different conclusions. The slope of the within-country variation depends on the dataset used, showing some form of non-linearity in the long run, but a monotonous within-trend towards specialization in the short run. These tentative conclusions

41

based on a non-parametric analysis, however, need to be further explored using panel data regression techniques in the next chapter.

4.2.3   Regression results

The parametric regression analysis in Table 7 includes a linear and a squared term for GDP per capita as explanatory variables, and excludes countries with fewer than one million inhabitants.

Table 7   Regression results for export specialization, SITC 5-digit, HS 6-digit, Feenstra 4-digit

	(1) Gini Pooled	(2) Herfindahl Pooled	(3) Gini FE	(4) Herfindahl FE	(5) Gini BE	(6) Herfindahl BE
SITC 5-digit						
GDP per capita	-9.2e-06***	-1.9e-05***	-4.4e-06***	-4.5e-06*	-1.0e-05***	-1.9e-05**
	(2.5e-07)	(1.3e-06)	(4.5e-07)	(2.5e-06)	(1.3e-06)	(9.5e-06)
GDP per capita squared	2.0e-10***	4.0e-10***	1.1e-10***	1.6e-10**	2.4e-10***	4.4e-10
	(8.8e-12)	(4.8e-11)	(1.2e-11)	(7.3e-11)	(5.2e-11)	(3.7e-10)
Constant	9.9e-01***	2.8e-01***	9.7e-01***	2.2e-01***	9.9e-01***	2.9e-01***
	(6.0e-04)	(5.9e-03)	(1.6e-03)	(8.6e-03)	(3.8e-03)	(2.8e-02)
Observations	3,862	3,862	3,862	3,862	3,862	3,862
Number of countries	142	142	142	142	142	142
R-squared	0.50	0.07	0.48	0.02	0.49	0.07
Turning point	22,754	23,050	19,879	13,792	20,804	
Feenstra 4-digit						
GDP per capita	-9.5e-06***	-2.5e-05***	-8.6e-06***	-1.1e-05***	-1.2e-05***	-2.9e-05***
	(2.5e-07)	(1.2e-06)	(3.9e-07)	(1.1e-06)	(1.5e-06)	(8.9e-06)
GDP per capita squared	2.1e-10***	6.6e-10***	1.6e-10***	3.2e-10***	3.1e-10***	8.9e-10**
	(9.7e-12)	(5.0e-11)	(9.1e-12)	(3.1e-11)	(6.2e-11)	(3.7e-10)
Constant	9.9e-01***	2.7e-01***	9.9e-01***	2.3e-01***	9.9e-01***	2.8e-01***
	(5.5e-04)	(4.7e-03)	(1.3e-03)	(3.6e-03)	(4.1e-03)	(2.5e-02)
Observations	4,008	4,008	4,008	4,008	4,008	4,008
Number of countries	135	135	135	135	135	135
R-squared	0.51	0.13	0.51	0.12	0.49	0.12
Turning point	23,084	18,923	27,659	16,708	18,476	16,501
HS 6-digit						
GDP per capita	-6.7e-06***	-1.2e-05***	1.3e-06***	-1.4e-06	-6.6e-06***	-1.3e-05*
	(3.7e-07)	(1.4e-06)	(4.6e-07)	(1.7e-06)	(1.1e-06)	(6.6e-06)
GDP per capita squared	1.3e-10***	2.3e-10***	1.5e-11*	4.8e-11	1.3e-10***	2.7e-10
	(1.2e-11)	(4.2e-11)	(9.4e-12)	(3.4e-11)	(3.7e-11)	(2.3e-10)
Constant	9.9e-01***	1.7e-01***	9.4e-01***	1.3e-01***	9.9e-01***	2.0e-01***
	(1.1e-03)	(6.5e-03)	(2.2e-03)	(7.9e-03)	(3.9e-03)	(2.4e-02)
Observations	1,612	1,612	1,612	1,612	1,612	1,612
Number of countries	134	134	134	134	134	134
R-squared	0.43	0.08	0.34	0.01	0.43	0.08
Turning point	25,698	25,108			25,963	
SITC 5-digit without outliers						
GDP per capita	-9.0e-06***	-2.5e-05***	-5.2e-06***	-6.9e-06***	-8.7e-06***	-1.7e-05
	(2.7e-07)	(1.2e-06)	(4.1e-07)	(2.1e-06)	(1.5e-06)	(1.0e-05)
GDP per capita squared	1.9e-10***	5.4e-10***	1.3e-10***	2.3e-10***	1.8e-10***	3.4e-10
	(9.8e-12)	(4.5e-11)	(1.1e-11)	(6.2e-11)	(6.1e-11)	(4.1e-10)
Constant	9.8e-01***	4.8e-01***	9.7e-01***	2.3e-01***	9.8e-01***	2.9e-01***
	(6.1e-04)	(5.3e-03)	(1.5e-03)	(7.6e-03)	(4.0e-03)	(2.8e-02)
Observations	3,849	3,853	3,849	3,858	3,849	3,858
Number of countries	142	142	142	142	142	142
R-squared	0.50	0.15	0.48	0.04	0.50	0.07
Turning point	23,485	23,337	19,287	14,692	23,932	

Robust standard errors in parentheses (non-robust standard errors for BE)

*** p<0.01, ** p<0.05, * p<0.1

Source: Author's own calculation based UN Comtrade (2008), Feenstra et al. (2005) and World Bank (2008) databases.

The results are shown for the SITC 5-digit, Feenstra 4-digit and HS 6-digit data, basically confirming the conjectured U- or L-shaped relationship between specialization and GDP per capita.

Columns 1 and 2 in Table 7 show the results of the simple pooled panel regression, where the time and cross-country dimensions of every observation are treated without methodological difference. The U-curve, i.e. a negative coefficient on the linear part and a positive coefficient on the squared part,[45] is highly significant at the 1 percent level for both indices and all three datasets. The estimated turning points of the U-curve vary a great deal across the different models, ranging from US\$14,000 to US\$28,000. The turning points of the fixed effects estimates, describing the turning point of an average country, are of particular interest but even these turning points are not consistent between datasets and measures.

The results of the fixed effects regressions (columns 3 and 4) are of particular interest in this context, showing a significant U-shaped relationship for the SITC and Feenstra datasets. For the HS dataset, the estimates do *not* show a U-curve, but this might be attributable to the short time span covered by the HS dataset, which is 19 years compared with 44 years for the SITC dataset and 38 years for the Feenstra dataset.

Columns 5 and 6 present the results of the between effects regression analysis, which estimates the correlation between the country *means* of diversification and the means of GDP per capita. The U-shaped relationship is found to be significant at the 1 percent level for all datasets when using the Gini Index, but for the Herfindahl Index, the estimates are surprisingly insignificant for the SITC and HS dataset.

The lower part of Table 7 uses SITC 5-digit data, but excludes potential outliers.[46] Since the sign and significance of the coefficients is consistent with those when outliers are included, it can be concluded that the U-curve is *not* driven by outliers.

Taking a closer look at the number of export lines provides additional insights into the diversification debate (Figure 13). As economies grow, new product lines are added to their exports, or, more precisely, products in a certain classification system which previously were zero now turn into a positive value.

---

[45] Except for the number of export lines as a measure of diversification, for which an inverted-U-curved shape is expected.

[46] Outliers are identified using the approach of Hadi (1992, 1994).

Figure 13   Number of export lines, SITC 5-digit
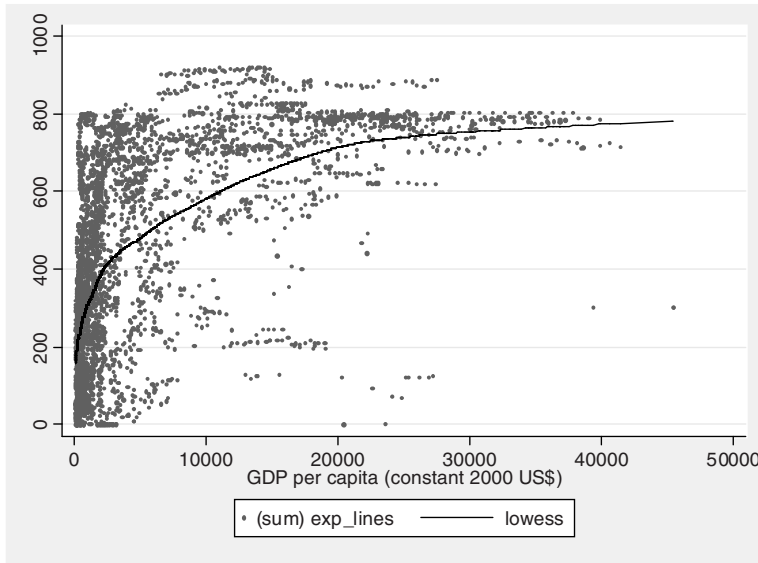
Source: Author's own calculation based UN Comtrade (2008) and World Bank (2008) databases.

This chapter has so far analysed *all* traded goods, hence diversification in this context reveals the well-known structural change by shifting exports away from agricultural products into processed products. Nevertheless, any change in the structure of the economy away from agriculture will appear as an increase in diversification, whether diversification has taken place or not, because in the current classification systems, agricultural goods are less disaggregated than non-agricultural goods. To examine whether the results are sensitive to this distinction and to allow for a more ready comparison with the results using MVA data, the export data is restricted to manufactured products only.

When only manufactured products are considered, the U-curve is *more* significant than when non-manufactured products are included as well. In addition, the turning points are – albeit not constant among the different datasets and indices – on average higher than for MVA. This could indicate that export patterns indeed follow production patterns. Overall, these results clearly show that diversification is the dominant force for low-income countries.

However, the approach of applying a simple specialization measure like the Gini coefficient is questionable because it measures the deviation from an equal distribution – a highly artificial concept. As outlined in the section on literature in this field, it is more realistic to measure the deviation from a global average. The following section will account for this idea.

4.3 Exports (relative specialization)

4.3.1 Descriptive statistics

The idea of relative diversification can be graphically represented in several ways. Following De Benedicts, Gallegati and Tamberi (2007:4-5), the sectoral market shares can be presented by country and year – Figures 14.1 and 14.2 show the data for the Republic of Korea using SITC 5-digit data as an example. Each bar in Figure 14.1 represents the ratio of the value of national total exports of a given sector to world exports in that sector in the year 2005. The horizontal line equals the share of the Republic of Korea's exports in world total exports. Figure 14.2 shows the same relationship for the year 1963. These figures clearly illustrate how the Republic of Korea's export structure has changed since 1963. In that year, the Republic of Korea's major export sector was "Ores & concentrates of nickel" (SITC code 28321), which accounted for 45 percent of world exports in that sector, while in 2005, the Republic of Korea's major export sector was "Special purpose vessels (e.g. light vessel dredgers)" (SITC 73592), accounting for 31 percent of global exports in that sector. Every value that exceeds the horizontal line indicates a sector with a so-called "revealed comparative advantage", i.e. this sector is exported on a higher scale than that country's average export intensity. Dividing each sectoral share by the share of the Republic of Korea's total exports in total world exports in the respective year yields the sectoral Balassa Indices.

Figure 14.1   Relative specialization of the Republic of Korea, sectoral shares and total share, 2005
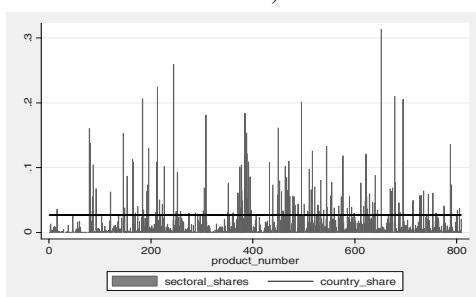
Figure 14.2   Relative specialization of the Republic of Korea, sectoral shares and total share, 1963
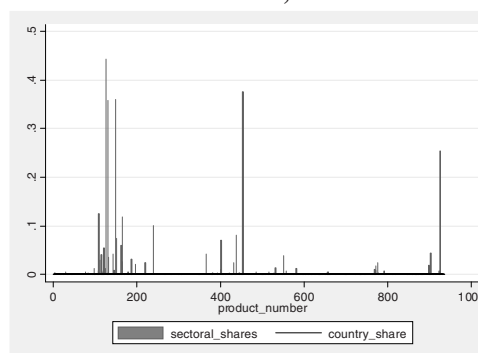
Source: Author's own calculation based UN Comtrade (2008) and World Bank (2008) databases.

As described in Section 3, the median of the Balassa Indices of a country in a particular year is a good indicator of diversification[47], with the Balassa Index also serving as a basis for the

---

[47] The median of the Balassa Index has the opposite sign to the other specialization indices.

construction of the relative Gini and Theil indices. The ten most and ten least specialized countries in terms of relative specialization are listed in Table 8. The world's largest economy, the USA, heads the list, followed by other OECD countries plus China. The ten most specialized countries are all located in sub-Saharan Africa, with the exception of Algeria. This ranking is probably an outcome of the construction of the relative specialization index itself, since big exporters, to a large extent, determine the benchmark – the global average export structure – and are by definition diversified. Whether this counteracts any trend for relative re-specialization is analysed in detail in this chapter.

Table 8   Ten most and ten least specialized countries, HS 6-digit, year 2005

Rank	Country	Relative Gini	Relative Theil	Relative DI	Median Balassa
1	United States	0.49	0.45	0.37	0.71
2	Germany	0.49	0.44	0.37	0.62
3	United Kingdom	0.51	0.50	0.36	0.47
4	Italy	0.55	0.72	0.46	0.68
5	France	0.55	0.56	0.41	0.57
6	Spain	0.61	0.79	0.47	0.51
7	Netherlands	0.61	0.67	0.45	0.37
8	China	0.61	0.84	0.53	0.54
9	Belgium	0.62	0.71	0.47	0.40
10	Austria	0.67	0.90	0.50	0.26
107	Gambia, The	0.99	5.81	0.96	0.00
108	Algeria	0.99	2.38	0.88	0.00
109	Gabon	0.99	2.91	0.91	0.00
110	Benin	0.99	5.50	0.96	0.00
111	Ethiopia (excludes Eritrea)	0.99	7.08	0.98	0.00
112	Sudan	1.00	3.29	0.95	0.00
113	Central African Republic	1.00	6.22	0.97	0.00
114	Burundi	1.00	5.38	0.98	0.00
115	Mali	1.00	5.33	0.97	0.00
116	Mauritania	1.00	6.76	1.00	0.00

Source: Author's own calculation based UN Comtrade (2008) and World Bank (2008) databases.

The summary statistics of the variables used in the following analysis are presented in Table 9. Compared to the absolute specialization indices (Table 5.2), the differences between the minima and maxima are much larger, as are the standard deviations, since the deviation from the global export distribution is smaller, on average, than the deviation from the artificial equal distribution that was implicitly used in absolute specialization measures.

Table 9   Summary statistics of relative export specialization, SITC 5-digit

Variable	Observations	Mean	Std. Dev.	Min	Max
Relative Gini	5,340	0.87	0.15	0.30	1.00
Relative Theil	5,340	2.99	1.88	0.24	10.67
Relative DI	5,340	0.77	0.18	0.24	1.00
Median Balassa	5,340	0.08	0.18	0.00	1.10
Population	4,960	3.12e+07	1.10e+08	40,740	1.31e+09
GDP per capita	4,567	6,332	8,305	92	54,178
GDP per capita (PPP)	3,042	10,956	11,361	319	79,032

Source: Author's own calculation based UN Comtrade (2008) and World Bank (2008) databases.

Table 10.1 presents the correlation between the various absolute and relative diversification indices, which varies a great deal between pairs and, in particular, between the relative and absolute diversification measures. While the absolute Gini Index and the relative Gini Index seem to be highly correlated, other measures such as the absolute and relative Theil Indices are less correlated. This difference is not attributable to non-linearities, since the rank correlations are also much lower between relative and absolute measures when compared with rank correlations within the two groups of measures (Table 10.2). These results provide an additional motivation to test whether a U-curved relationship can be observed in relative specialization, as the results from absolute specialization cannot be directly applied to relative specialization.

Table 10.1   Correlation between relative and absolute export specialization, SITC 5-digit

	Gini	Theil	Herfindahl	Hirschman	Rel. Gini	Rel. Theil	Rel.DI	Median Balassa
Gini	1.00							
Theil	0.85	1.00						
Herfindahl	0.57	0.89	1.00					
Hirschman	0.70	0.96	0.97	1.00				
Relative Gini	0.97	0.81	0.54	0.66	1.00			
Relative Theil	0.68	0.67	0.45	0.56	0.71	1.00		
Relative DI	0.92	0.80	0.52	0.65	0.96	0.83	1.00	
Median Balassa	-0.89	-0.61	-0.34	-0.46	-0.89	-0.52	-0.80	1.00

Table 10.2   Rank correlation between relative and absolute export specialization, SITC 5-digit

	Gini	Theil	Herfindahl	Hirschman	Rel. Gini	Rel. Theil	Rel. DI	Median Balassa
Gini	1.00							
Theil	0.99	1.00						
Herfindahl	0.95	0.98	1.00					
Hirschman	0.95	0.98	1.00	1.00				
Relative Gini	0.97	0.94	0.89	0.89	1.00			
Relative Theil	0.76	0.71	0.67	0.67	0.84	1.00		
Relative DI	0.87	0.83	0.77	0.77	0.93	0.94	1.00	
Median Balassa	-0.82	-0.77	-0.71	-0.71	-0.83	-0.79	-0.82	1.00

Source: Author's own calculation based UN Comtrade (2008) and World Bank (2008) databases.

4.3.2   Non-parametric results

Figure 15.1 plots relative specialization, measured by the relative Gini Index, against GDP per capita, together with a non-parametric Lowess curve. Most observations are scattered in the upper-left part of the plot, indicating countries with an export structure that highly diverges from the global export structure. The figure reveals that as GDP per capita rises, countries' export structures become more similar to the global export structure, but the additional increase in relative diversification diminishes at higher levels of GDP per capita, and eventually reverses at the highest levels of GDP per capita. Figures 15.2 and 15.3 display a similar pattern for the Feenstra and HS datasets, with greater divergence from the global export structure for countries with low GDP per capita, and increasing similarity towards the global distribution with increasing GDP per capita, but at a decreasing rate, with a slight tendency towards re-specialization. As with the results in the previous sections, it might be better to refer to these stylized facts as an "L-curve" rather than a "U-curve".

Figure 15.1   Relative export specialization, SITC 5-digit
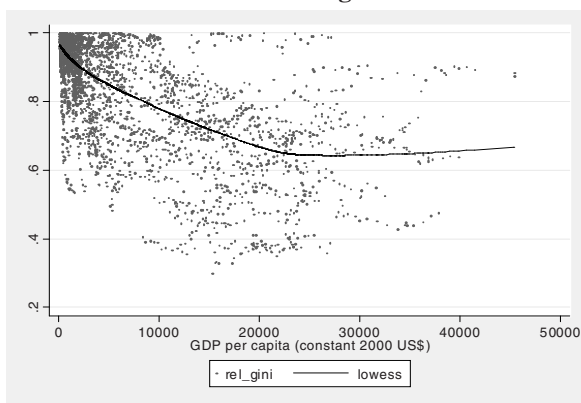

Figure 15.2   Relative export specialization, HS 6-digit
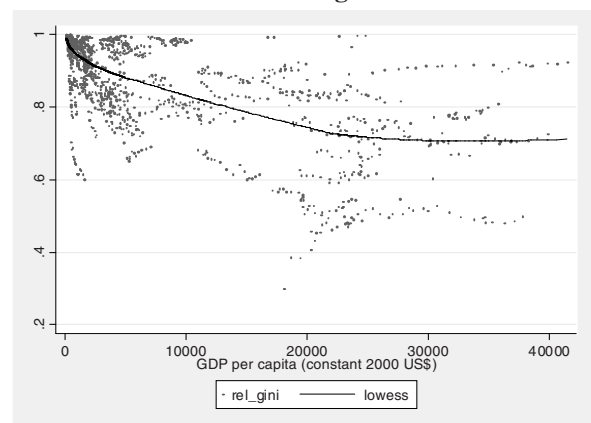

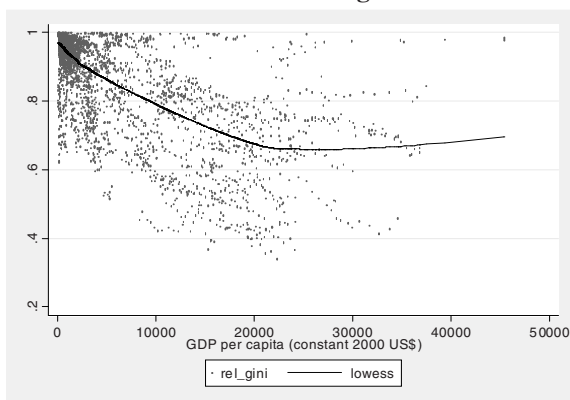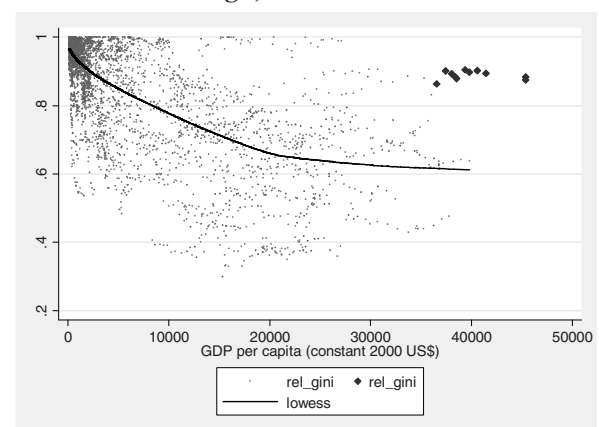**Figure 15.3   Relative export specialization, Feenstra 4-digit**

Figure 15.4   Relative export specialization, SITC 5-digit, Lowess excl. outliers

Source: Author's own calculation based UN Comtrade (2008), Feenstra et al. (2005) and World Bank (2008) databases.

The Hadimvo procedure identifies a number of outliers on the upper-right side of the scatter plot (see the marked observations in Figure 15.4). When these outliers are excluded from the Lowess calculation, the re-specialization part of the curve entirely disappears. This preliminary conclusion based on the pooled panel, however, remains to be tested by employing panel data regression methods.

Figure 16.1 plots the same relationships as above for the SITC dataset, but with data on GDP per capita using constant PPP dollars, which is the measure used by De Benedictis, Gallegatis and Tamberi (2007) in their analysis of relative diversification. The conclusions are the same as for constant US dollars, but the upward-sloping path is more significant, which is presumably attributable to a small number of observations with high GDP per capita and high relative specialization.

Figure 16.2 presents the scatter plot between relative specialization in the SITC dataset and the logarithm of GDP per capita to visualize observations with low GDP per capita. Even at the lowest levels of GDP per capita, it is evident that rising GDP per capita is connected with a higher variance of specialization through an increasing number of countries that converge to the global export structure, while some countries remain relatively specialized.

Figure 16.1 Relative export specialization, SITC 5-digit, PPP dollars	Figure 16.2 Relative export specialization, SITC 5-digit, logarithmic scale
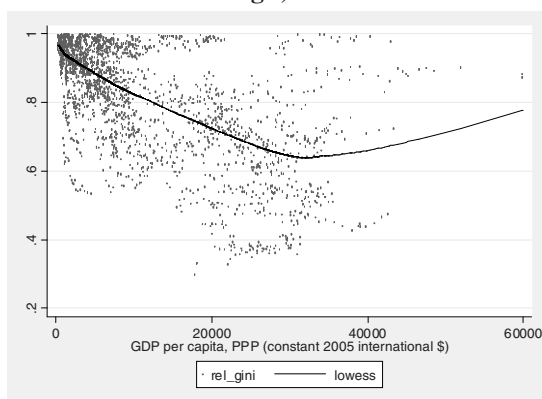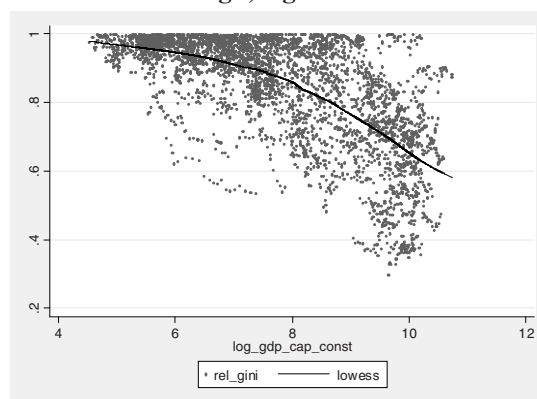
Source: Author's own calculation based UN Comtrade (2008) and World Bank (2008) databases.

Figure 17.1 presents SITC 5-digit data for the year 2000 to provide information about the location of specific countries. Possibly as a result of the way the relative specialization measure is calculated, as mentioned above, large economies such as the United Kingdom (GBR), Germany (DEU), France (FRA) and the USA are most diversified, as they probably highly

influence the benchmark global export structure. Some oil-exporting countries, like Kuwait (KWT), United Arab Emirates (ARE) and Norway (NOR), are highly specialized relative to all other countries, despite being in the middle or upper range of GDP per capita.

The evolution of specialization over the period 1963 to 2007 for six selected countries is shown in Figure 17.2. Among these, the Republic of Korea is the only country with an export structure emerging as a U-curve. Germany is the most diversified country over the whole period, and the USA is approaching Germany's export structure. Ghana is the country with the greatest distance to the global average export structure. Chile also remains highly specialized, while Mexico's relative structure varies a great deal, but shows a tendency to move towards the global export structure from the late 1980s onwards.

Figure 17.1 Export specialization and GDP per capita, 2000	Figure 17.2 Relative export specialization over time, selected countries
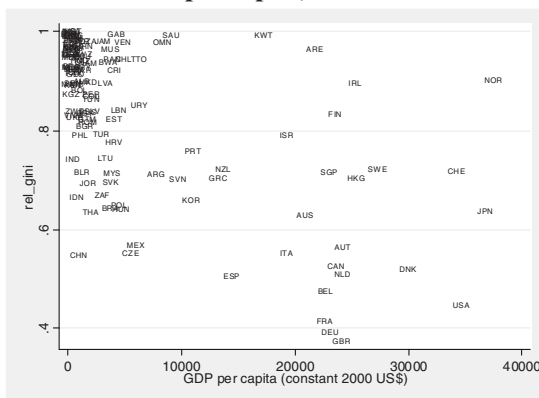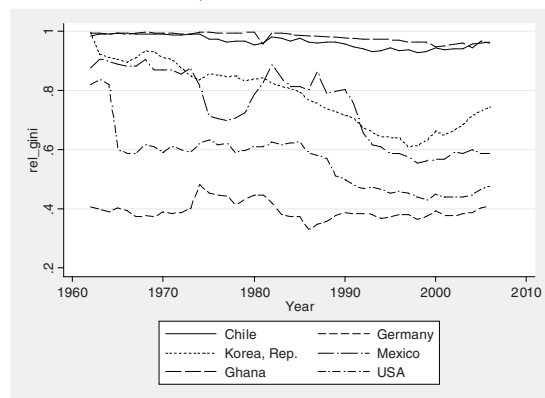
Source: Author's own calculation based UN Comtrade (2008) and World Bank (2008) databases.

4.3.3 Regression results

The above presented non-parametric estimates indicate a negative correlation between relative export specialization and GDP per capita, with the slope coefficient decreasing at higher levels of GDP per capita, but without a pronounced trend towards re-specialization, except for the within-country variation. The regression results (Table 11) show a significant and quite robust U-curved correlation between specialization and GDP per capita. The pooled regression (columns 1 and 2) is significant at the 1 percent level for the four measures and three datasets, as well as for constant US dollars and constant PPP dollars. The coefficients on the median Balassa Index occasionally return an unexpected sign for the Feenstra dataset for the between

effects regression with constant PPP dollars. However, the estimates for the vast majority of settings are as expected and highly significant.

Table 11   Regression results for relative export specialization, SITC 5-digit, HS 6-digit and Feenstra 4-digit

	(1) Relative Gini	(2) Median Balassa	(3) Relative Gini	(4) Median Balassa	(5) Relative Gini	(6) Median Balassa
SITC 5-digit	Pooled	Pooled	FE	FE	BE	BE
GDP per capita	-2.6e-05***	2.7e-05***	-2.4e-05***	1.9e-05***	-3.1e-05***	3.1e-05***
	(7.5e-07)	(1.3e-06)	(1.3e-06)	(1.3e-06)	(4.0e-06)	(5.3e-06)
GDP per capita squared	5.3e-10***	-5.0e-10***	4.4e-10***	-2.8e-10***	7.6e-10***	-6.8e-10***
	(2.7e-11)	(4.8e-11)	(3.5e-11)	(3.6e-11)	(1.6e-10)	(2.1e-10)
Constant	9.5e-01***	-1.3e-02***	9.5e-01***	1.6e-02***	9.6e-01***	-1.3e-02
	(1.9e-03)	(2.3e-03)	(5.0e-03)	(5.1e-03)	(1.2e-02)	(1.5e-02)
Observations	3,862	3,862	3,862	3,862	3,862	3,862
Number of countries	142	142	142	142	142	142
R-squared	0.49	0.39	0.49	0.38	0.47	0.39
Turning point	24,630	27,210	27,117	32,838	20,413	22,733
SITC 5-digit excl. outliers						
GDP per capita	-2.5e-05***		-2.4e-05***		-2.7e-05***	
	(8.0e-07)		(5.1e-06)		(4.4e-06)	
GDP per capita squared	4.9e-10***		4.5e-10***		6.1e-10***	
	(3.0e-11)		(1.3e-10)		(1.8e-10)	
Constant	9.5e-01***		9.5e-01***		9.5e-01***	
	(1.9e-03)		(1.9e-02)		(1.2e-02)	
Observations	3,851		3,851		3,851	
Number of countries	142		142		142	
R-squared	0.50		0.50		0.48	
Turning point	25,824		27,024		22,404	
Feenstra 4-digit						
GDP per capita	-2.8e-05***	3.0e-05***	-2.1e-05***	1.9e-05***	-3.4e-05***	3.4e-05***
	(7.3e-07)	(1.2e-06)	(9.9e-07)	(1.4e-06)	(3.9e-06)	(5.4e-06)
GDP per capita squared	6.2e-10***	-6.1e-10***	3.8e-10***	-3.3e-10***	9.0e-10***	-8.3e-10***
	(2.9e-11)	(4.6e-11)	(2.5e-11)	(3.4e-11)	(1.7e-10)	(2.3e-10)
Constant	9.7e-01***	-1.7e-02***	9.5e-01***	1.7e-02***	9.7e-01***	-2.0e-02
	(1.5e-03)	(1.6e-03)	(3.3e-03)	(4.6e-03)	(1.1e-02)	(1.5e-02)
Observations	4,047	4,047	4,047	4,047	4,047	4,047
R-squared	0.55	0.44	0.54	0.44	0.53	0.43
Number of countries	136	136	136	136	136	136
Turning point	22,712	24,769	27,283	28,445	18,705	20,810
HS 6-digit						
GDP per capita	-1.9e-05***	1.8e-05***	-9.5e-06***	5.6e-06**	-1.8e-05***	1.7e-05***
	(1.1e-06)	(1.7e-06)	(1.6e-06)	(2.2e-06)	(3.0e-06)	(3.8e-06)
GDP per capita squared	3.4e-10***	-3.0e-10***	1.7e-10***	-1.7e-10***	3.3e-10***	-2.6e-10*
	(3.6e-11)	(5.6e-11)	(3.4e-11)	(4.1e-11)	(1.0e-10)	(1.3e-10)
Constant	9.6e-01***	-7.4e-03**	9.2e-01***	6.4e-02***	9.7e-01***	-7.3e-03
	(2.8e-03)	(3.6e-03)	(7.6e-03)	(1.1e-02)	(1.1e-02)	(1.4e-02)
Observations	1,612	1,612	1,612	1,612	1,612	1,612
R-squared	0.472	0.351	0.47	0.11	0.47	0.35
Number of countries			134	134	134	134
Turning point	27,387	29,931	27,298	16,046	27,442	32,603

Robust standard errors in parentheses (non-robust standard errors for BE)

*** p<0.01, ** p<0.05, * p<0.1

Source: Author's own calculation based UN Comtrade (2008), Feenstra et al. (2005) and World Bank (2008) databases.

Note in particular that when excluding outliers in the SITC 5-digit dataset, the fixed effects regression results for the Gini, Theil and RDI figures become even more robust, with a slightly larger absolute value of the coefficients and slightly smaller standard errors.

These results indicate that the export structure of low-income countries becomes increasingly similar to the world average, but that this process is slowing down and that countries do not actually reach the world average structure. Eventually, high-income countries tend to diversify away from the global average, possibly due to specific products that can only be produced by countries with the highest incomes. These products are potentially high value-added products, which require a high level of technology and are as a result sold at high prices to a narrow range of consumers.

These conclusions also hold when non-manufacturing data are excluded from the regressions. In this case, the results are even more robust than when including non-manufacturing data.

The high level of relative diversification for high-income countries also has a dynamic interpretation. As the global export structure changes over time, the benchmark of relative specialization also changes. When global trade patterns become more diversified, high-income countries are able to adapt to new demand structures more quickly than low- and middle-income countries; otherwise they would – by definition – be more specialized.

5. Conclusions

This study has analysed the relationship between economic specialization and national income per capita. For development policy, it is crucial whether a country focuses on improving its production or export capacity within a certain narrow range of products for which it has a comparative advantage, or whether a country aims to diversify its production structure to be less vulnerable to economic shocks. In the related literature, arguments are put forward for both specialization and diversification, and the seminal econometric paper by Imbs and Wacziarg (2003) presents empirical evidence that both trends are taking place but at different economic stages. Their finding of a "U-curve" between specialization and income per capita has stimulated a large number of studies that have expanded their analysis by either criticizing the methodology or confirming the findings and adding more dimensions.

The main results from the analysis are as follows:

1. By combining the various methodological approaches, the econometric analysis presented in this study has shown that low-income countries *diversify* their export base as their GDP per capita grows. When production is considered instead of exporting, this trend is less distinct.

2. Some doubts have been raised on the issue whether the diversification trend eventually turns into re-specialization at higher levels of GDP per capita. Although this critique does not diminish the importance of diversification, it challenges the recent discussion about the existence of a U-curve relationship between specialization and income per capita. At least, the shape of the relation between specialization and GDP per capita can be illustrated by an "*L-curve*".

3. There is one crucial caveat in the idea of an L-curve. The comparison of within-country and between-country comparisons reveals that time-invariant differences in diversification levels account for a large share of the observed overall differences.

4. This study has discussed the relevance of the particular benchmark used for maximum diversification. Most studies use specialization measures that quantify the deviation from a distribution where every product has the same share. However, countries actually benefit by serving global demand regardless of how specialized or diversified global demand is structured. Therefore, the distribution of global demand itself should serve as a benchmark, and *relative specialization* should measure how much a country diverges from this benchmark.

5. As diversification is the driving force of economic development for low-income countries, *policymakers* need to design and implement policies that enable and foster this process.

Several issues have not been analysed empirically in this study and could be addressed by future research:

1. The theoretical rationale for the stages of the economic diversification result needs to be further elaborated. In particular, no strong conclusions about the directions of causality can be drawn from this study. Moreover, the drivers of the production structure might differ from the drivers of the export structure.

2. Given the missing theoretical rationale and thus the "true" model, the estimated relations might suffer from various forms of endogeneity biases. From a pure econometric point of

view, some forms of bias could be overcome by applying dynamic panel data models with instrumental variables.

3. This study aimed at drawing generic conclusions on the global level, i.e. all countries with available data have been analysed together, controlling for time-invariant heterogeneity. Further insights might be gained when looking at different country groups separately to better understand the relation between specific country characteristics and the corresponding pattern of diversification.

4. The nature of the products in which economies specialize or diversify has not been touched upon in this study, and other studies have so far only analysed the economic risk associated with specific sectors. Other relevant dimensions could include the technological content of products or their capital and labour intensities.

5. With time passing, more data will be available in future, so the hypotheses used in this study are subject to revision once new information is available. This is particularly relevant for production data, where the dataset used in this study was rather limited in terms of time coverage and level of disaggregation.

6. Diversification can be split into diversification at the extensive and the intensive margins, i.e. whether diversification takes place due to balancing the existing production or by producing entirely new products. As only a small number of studies have considered such margins, the analysis of the extensive and intensive margins should be reconsidered using different concepts and measures of specialization as well as different datasets.

7. The issue of *market* diversification, i.e. the pattern of geographic diversification of export destinations, has barely been researched in the existing literature. It could well be that diversification therefore should be seen as two-dimensional: diversification in products and diversification in markets.

8. The relationship between the structure of specialization and economic *growth* has been analysed by some authors, but not targeted in this study. Further effort has to be made – both theoretically and empirically – to investigate the relationship between the level and nature of specialization and a sustainable economic growth path.

6.   References

Acemoglu, Daron and Fabrizio, Zilibotti (1997): "Was Prometheus Unbound by Chance? Risk, Diversification, and Growth", Journal of Political Economy, vol. 105, no. 4, pp. 709-751.

Agosin, Manuel R. (2007): "Export Diversification and Growth in Emerging Economies", University of Chile, Department of Economics, Working Paper No. 233.

Agosin, Manuel R. and Bravo-Ortega, Claudio (2007): "The Emergence of New Successful Export Activities in Latin America: The Case of Chile", University of Chile, Department of Economics, Working Paper No. 236.

Albaladejo, Manuel (2007): "New Measures of Industrial Competitiveness: Indicators for UNIDO's CIP Index in IDR 2008", Presentation at 30 October 2007, Trade Analysis and Compliance Unit, Trade Capacity Building Branch. Vienna: United Nations Industrial Development Organization.

Al-Marhubi (2000): "Export Diversification and Growth: an Empirical Investigation", Applied Economics Letters, vol. 7, no. 9, pp. 559-562.

Amiti, Mary (1999): "Specialization Patterns in Europe", Review of World Economics (Weltwirtschaftliches Archiv), vol. 135, no. 4, pp. 573-593.

Batista, Catia and Potin, Jacques (2007): "Stages of Diversificaton and Specialization in an Heckscher-Ohlin World, 1976-2000", University of Oxford, Department of Economics Discussion Paper Series Number 356.

Bebczuk, Ricardo and Berrettoni, N. Daniel (2006): "Explaining Export Diversification: An Empirical Analysis", CAF Research Program on Development Issues.

Bhagwati, Jagdish N. and Srinivasan, T. N. (1983): "Lectures on international trade", Cambridge, Massachusetts: MIT Press.

Carrere, C., Strauss-Kahn, V. and Cadot O. (2007): "Export Diversification: What's behind the Hump?" CERDI Working Papers 200724.

Collier, Paul (2002): "Primary commodity dependence and Africa's Future", In: Annual Bank Conference on Development Economics, Proceedings 2002, edited by B. Pleskovic and N. Stern. New York: Oxford University Press. Available from: http://siteresources.worldbank.org/INTABCDEWASHINGTON2002/Resources/Collier-full.pdf [12 February 2009]

Corden W. Max, J. Peter Neary (1982): "Booming Sector and De-Industrialisation in a Small Open Economy", Economic Journal, vol. 92, no. 368, pp. 825-848.

Cuddington, John T., Ludema, Rodney and Jayasuriya, Shamila A. (2002): "Prebisch-Singer Redux", in: Daniel Lederman and William F. Maloney (eds.), "Natural Resources and Development: Are They a Curse? Are They Destiny?", World Bank/Stanford University Press, 2007. Cited version available from: http://inside.mines.edu/%7Ejcudding/papers/Prebisch_Singer/Cuddington_et%20al_PS_Redux-short_10.30.02.pdf [16 February 2009]

De Benedictis, Luca (2004): "Stages of Diversification: Comment", Available from: http://www.cide.info/conf_old/papers/1142.pdf [17 April 2009]

De Benedictis, Luca and Tamberi, Massimo (2004): "Overall Specialization Empirics: Techniques and Applications", Open Economies Review, vol. 15, pp. 323-346.

De Benedictis, Luca, Gallegati, Marco and Tamberi, Massimo (2003): "Non-Linearities in Specialization and Growth", Available from http://ssrn.com/abstract=412260 [22 May 2009]

De Benedictis, Luca, Gallegati, Marco and Tamberi, Massimo (2008): "Semiparametric Analysis of the Specialization-Income Relationship". Applied Economics Letters, vol. 15, no.4, pp. 301-306.

De Benedictis, Luca, Gallegati, Marco and Tamberi, Massimo (2007): "Overall Specialization and Income: Countries Diversify", Working papers 73, University of Rome "La Sapienza", CIDEI.

Deardorff, Alan V. (2005): "How Robust is Comparative Advantage?" Research Seminar in International Economics Working Paper No. 737. University of Michigan.

Dornbusch, Rudiger, Fischer, Stanley and Samuelson, Paul A. (1977): "Comparative Advantage, Trade, and Payments in a Ricardian Model with a Continuum of Goods", American Economic Review, vol. 67, no. 5, pp. 823-839.

Feenstra, Robert C., Lipsey, Robert E., Deng, Haiyan, Ma, Alyson C. and Mo, Hengyong (2005): "World Trade Flows: 1962-2000", NBER Working Paper No.11040, Cambridge: National Bureau of Economic Research Working Paper Series.

Gleditsch, Kristian Skrede (2002): "Expanded Trade and GDP data." Journal of Conflict Resolution, vol. 46, no. 5, pp. 712-724.

Graham, Frank D. (1923): "Some Aspects of Protection Further Considered", Quarterly Journal of Economics, vol. 37, no. 2, pp. 199-227.

Grossman, Gene M. and Helpman, Elhanan (1992): "Innovation and growth in the global economy", Cambridge, Massachusetts: MIT Press.

Hadi, Ali S. (1992): "Identifying Multiple Outliers in Multivariate Data", Journal of the Royal Statistical Society, Series B (Methodological), vol. 54, no. 3, pp. 761-771.

Hadi, Ali S. (1994): "A Modification of a Method for the Detection of Outliers in Multivariate Samples", Journal of the Royal Statistical Society, Series B (Methodological), vol. 56, no. 2, pp. 393-396.

Harrigan, Frank (2007): "Technology and Development in Asia", ERD Policy Brief Series No. 49, Economics and Research Department, Asian Development Bank.

Hausmann, Ricardo and Klinger, Bailey (2006): "Structural Transformation and Patterns of Comparative Advantage in the Product Space", Centre for International Development at Harvard University Working Paper No. 128.

Hausmann, Ricardo and Rodrik, Dani (2003): "Economic development as Self-Discovery", John F. Kennedy School of Government, Harvard University.

Hausmann, Ricardo, Hwang, Jason and Rodrik, Dani (2007): "What you export matters", Journal of Economic Growth, vol. 12, no. 1, pp. 1-25.

Hesse, Heiko (2008): "Export Diversification and Economic Growth", Commission on Growth and Development Working Paper 21.

Hirschman, Albert Otto (1958): "The strategy of economic development", New Haven, Conn.: Yale University Press.

Hummels, David and Klenow, Peter J. (2005): "The Variety and Quality of a Nation's Exports", American Economic Review, vol. 95, no. 3, pp. 704-723.

Imbs, Jean and Wacziarg, Romain (2000): "Stages of Diversification", Stanford University, Graduate School of Business Research Papers 1653.

Imbs, Jean and Wacziarg, Romain (2003): "Stages of Diversification", American Economic Review, vol. 93, no. 1, pp. 63-86.

Kalemli-Ozcan, Sebnem, Sorensen, Bent E. and Yosha, Oved (2003): "Risk sharing and industrial specialization: Regional and international evidence", American Economic Review, vol. 93, no. 3, pp. 903-918.

Klinger, B. and Lederman, D. (2004): "Discovery and Development: An empirical exploration of 'new' products", World Bank Policy Research Working Paper Series No. 3450.

Klinger, Bailey and Lederman, Daniel (2006): "Diversification, Innovation, and Imitation inside the Global Technological Frontier", World Bank Policy Research Working Paper 3872.

Koren, Miklos and Tenreyro, Silvana (2004): "Diversification and Development", Working Papers 03-3, Federal Reserve Bank of Boston.

Krugman, Paul R. and Obstfeld, Maurice (2006): "International Economics: Theory and Policy", 7. ed., Boston, Massachusetts: Pearson, Addison-Wesley.

Kuznets, Simon (1971): "Modern Economic Growth: Findings and Reflections. Prize Lecture", Lecture to the memory of Alfred Nobel, December 11, 1971.
Available from: http://nobelprize.org/nobel_prizes/economics/laureates/1971/kuznets-lecture.html [2 March 2009]

Lall, Sanjaya (2000): "The Technological Structure and Performance of Developing Country Manufactured Exports, 1985-1998", QEH Working Papers No. 44, Queen Elizabeth House, University of Oxford.

Lall, Sanjaya, Weiss, John and Zhang, Jinkang (2006): "The 'sophistication' of exports: A new trade measure", World Development, vol. 34, no. 2, pp. 222–237.

Lederman, Daniel and Maloney, William F. (2003): "Trade structure and growth", Policy Research Working Paper Series 3025, The World Bank.

Lorenz, M. O. (1905): "Methods of Measuring the Concentration of Wealth", Publications of the American Statistical Association, vol. 9, no. 70, pp. 209-219.

Lutz, Matthias G. (1999): "A General Test of the Prebisch-Singer Hypothesis", Review of Development Economics, vol. 3, no. 1, pp. 44-57.

Maneschi, Andrea (2004): "The True Meaning of David Ricardo's Four Magic Numbers", Journal of International Economics, vol. 62, pp 433-443.

Matsuyama, Kiminori (1991): "Agricultural Productivity, Comparative Advantage and Economic Growth", National Bureau of Economic Research Working Paper No. 3606.

Ocampo, Jose Antonio and Parra, Maria Angela (2004): "The Terms of Trade for Commodities in the Twentieth Century", International Trade 0402006, EconWPA.

Osakwe, Patrick N. (2007): "Foreign Aid, Resources and Export Diversification in Africa: A New Test of Existing Theories", United Nations Economic Commission for Africa, African Trade Policy Centre, Work in Progress No. 61.

Prebisch, Raúl (1950): "The Economic Development of Latin America and its Principal Problems", Economic Commission for Latin America. Lake Success, NY: United Nations Department of Economic Affairs.

Raffer, Kunibert (2004): "Handel und Unterentwicklung, Kritische Anmerkungen zur Freihandelsideologie", in: Karin Fischer, Irmi Maral-Hanak, Gerald Hödl, Christof Parnreiter (Hg): Entwicklung und Unterentwicklung, Eine Einführung in Probleme, Theorien und Strategien, pp. 107-126, Vienna: Mandelbaum.

Raffer, Kunibert and Singer, Sir Hans (2004): "The Economic North-South Divide: Six Decades of Unequal Development", 2nd edition, Cheltenham (UK)/Northampton (US): Edward Elgar.

Ramey, Garey and Ramey, Valerie A. (1995): "Cross-country evidence on the link between volatility and growth", American Economic Review, vol. 85, no. 5, pp. 1138-1151.

Ray, Debraj (1998): "Development Economics", Princeton: Princeton University Press.

Ricardo, David (1971): "On the Principles of Political Economy, and Taxation", Harmondsworth: Penguin Books.

Rodrik, Dani (2005): "Policies for economic diversification", CEPAL Review 87, pp. 7-23.

Rodrik, Dani (2007): "One Economics, Many Recipes. Globalization, Institutions and economic growth", Princeton University Press.

Rodrik, Dani (2007b): "Specialization or diversification", 16 May 2007. Dani Rodrik's weblog. Available from:
http://rodrik.typepad.com/dani_rodriks_weblog/2007/05/specialization_.html
[2 March 2009]

Rodrik, Dani (2007c): "Industrial Development: Some stylized facts and policy directions", In: Industrial Development for the 21st Century: Sustainable Development Perspectives. New York: United Nations Department of Economic and Social Affairs.

Ruffin, Roy J. (2002): "David Ricardo's Discovery of Comparative Advantage", History of Political Economy, vol. 34, pp. 727-748.

Sachs, J-D and Warner, A-M (1995): "Natural Resource Abundance and Economic Growth", National Bureau of Economic Research Working Paper 5398, Cambridge.

Sachs, Jeffrey D. and Warner, Andrew M. (2001): "The Curse of Natural Resources", European Economic Review, vol. 45, no. 4-6, pp. 827-838.

Saint-Paul, Gilles, (1992): "Technological choice, financial markets and economic development", European Economic Review, vol. 36, no. 4, pp. 763-781.

Singer, H. W. (1950): "The Distribution of Gains between Investing and Borrowing Countries", American Economic Review, vol. 40, no. 2, pp. 473-485.

Subramanian, Arvind (2007): "Capital Account Convertibility: A Neglected Consideration. Peterson Institute for International Economics", Paper presented at the Center for Policy Research, New Delhi, India. Available from: http://www.iie.com/publications/papers/subramanian04071.pdf [18 February 2009].

Turnovsky, Stephen J. (1974): "Technological and Price Uncertainty in a Ricardian Model of International Trade", Review of Economic Studies, vol. 41, no. 2, pp. 201-217.

UN Comtrade (2008): "United Nations Commodity Trade Statistics Database", Accessed through World Bank World Integrated Trade Solution, Available from: http://wits.worldbank.org/witsweb/ [17 April 2009]

UNCTAD (2008): "Handbook of Trade Statistics", Geneva: United Nations Conference on Trade and Development. Available from: http://www.unctad.org/en/docs/tdstat33_en.pdf. [26 February 2009]

UNIDO (2006): "INDSTAT3 2006. Industrial Statistics Database. 3-digit level of ISIC Code", Vienna: United Nations Industrial Development Organization.

UNIDO (2009): "Industrial Development Report 2009: Breaking In and Moving Up: New Industrial Challenges for the Bottom Billion and the Middle-Income Countries", Vienna: United Nations Industrial Development Organization.

World Bank (2008): "World Development Indicators", Washington DC: International Bank for Reconstruction and Development / The World Bank.

WTO (2009): "The case for open trade. Understanding the WTO: Basics", Available from: http://www.wto.org/english/thewto_e/whatis_e/tif_e/fact3_e.htm [2 March 2009]

Yamada, Tetsuo (2005): "Relevance and applicability of the UNIDO Industrial Statistics Database for Research Purposes", Expert Group Meeting on Industrial Statistics, New York, 19-23 September 2005. United Nations Department of Economic and Social Affairs, Statistics Division.