



GDAŃSK UNIVERSITY OF TECHNOLOGY
FACULTY OF MANAGEMENT AND ECONOMICS

PRODUCT DIVERSIFICATION, RELATIVE SPECIALISATION AND ECONOMIC DEVELOPMENT: IMPORT-EXPORT ANALYSIS

Aleksandra Parteka*, Massimo Tamberi**

**GUT Faculty of Management and Economics
Working Paper Series A (Economics, Management, Statistics)
No 15/2013(15)**

August 2013

* Gdansk University of Technology, Faculty of Management and Economics,
aparteka@zie.pg.gda.pl (corresponding author)

** Università Politecnica delle Marche



PRODUCT DIVERSIFICATION, RELATIVE SPECIALISATION AND ECONOMIC DEVELOPMENT: IMPORT-EXPORT ANALYSIS

Aleksandra Parteka* and Massimo Tamberi**

This version: August 2013 (forthcoming in *Journal of Macroeconomics*)

Abstract

This paper contributes to trade diversification literature by comparing changes in relative (i.e. assessed in comparison with world patterns) heterogeneity of import and export structures in the process of economic development. In particular, by focusing on the diversification of imports, we add a missing piece to already analysed export trends. We use highly disaggregated trade statistics (4963 product lines) for 163 countries (1988-2010) and find that, despite differences in levels (imports being typically more diversified than exports, particularly at lower stages of economic development), they follow a similar path of evolution in the development process. Progressing relative diversification (despecialisation) of both import and export structures accompanies economic growth, while re-specialisation is plausible only in case of few specific countries (very rich, small ones, abundant in oil/petrol). We also show that even though while diversifying countries increase the degree of import-export similarity in terms of product categories, imported and exported goods differ in terms of within-product characteristics.

Keywords: diversification, specialisation, trade, economic development

JEL: F14, O11

*Gdansk University of Technology, Faculty of Management and Economics, Narutowicza 11/12, 80-233 Gdansk, Poland. tel. (+48 58) 348-60-04 (direct), fax (+48 58) 348-60-07. Corresponding author (aparteka@zie.pg.gda.pl).

**Università Politecnica delle Marche, Facoltà di Economia, Piazzale Martelli 8, 60121 Ancona, Italy.

The paper presents the results of the research project realised while A.Parteka was visiting at Universitat Pompeu Fabra (Barcelona, Spain) - financial support from the Polish Ministry of Science and Higher Education (Mobilnosc Plus 652/MOB/2011/0), and the hospitality of UPF Department of Economics and Business are gratefully acknowledged. Previous version of the paper was circulated under the title: Relative Product Diversification in the Course of Economic Development: Import-Export Analysis. We would like to thank: Paola Bustos, Antonio Ciccone, Luca de Benedictis, Giorgia Giovannetti, Jakub Growiec, Vanessa Strauss-Kahn, as well as the participants at: Warsaw Economic Seminar (Warsaw School of Economics, March 2012), LPD Seminar (Universitat Pompeu Fabra, May 2012), XII CIE Conference (University of Granada, June 2012), XIV ETSG Conference (University of Leuven, September 2012), XVII DEGIT Conference (University of Milan, September 2012), 53rd SIE Conference (Università della Basilicata, October 2012), Structural Change, Dynamics, and Economic Growth Conference (University of Pisa, September 2013) for their valuable comments and suggestions. We are responsible for all the remaining errors.

1. Introduction

This paper deals with empirical relationship between changes in the product diversity of imports (analysed simultaneously vs. export trends) and economic development process. In particular, we assess changes in the composition of import (and export) basket of countries in comparison with typical world patterns, which enables us to trace relative specialisation patterns.

Several arguments have influenced the choice of our research subject. Looking from the side of *exports*, their diversification (de-specialisation) has been analysed as a factor reducing risk and exposure to idiosyncratic shocks (Hesse, 2009). This is particularly important in the case of low income countries, which – as shown in recent literature on export diversification and development¹ (Cadot et al., 2011a; Minondo, 2011; Parteka and Tamberi, 2013) - often have very specialised (i.e. poorly diversified) economic structures and/or are dependent on natural resources. On the other hand, diversification of *imports* is directly related to the ability of countries to experience welfare and productivity gains resulting from increasing diversity of imported goods (for final consumption and in the form of inputs) in the spirit of Dixit-Stiglitz (1977). Throughout the paper we will show that the two sides of the phenomenon are linked, but only through a parallel import-export study we can fully refer to alternative (often contradictory) predictions of different theoretical models. They focus either on demand side (i.e. love-for-variety in consumption) or supply side (differentiated inputs) of the phenomenon and we capture both.

Conventionally, the ‘old’ trade theory (H-O and Ricardian frameworks) drew the attention to productivity gains resulting from trade specialisation. Failure to capture the extensive margin of trade (reflecting rising diversity of products) is also present in traditional computable general equilibrium (CGE) models based on the Armington’s (1969) model of nationally differentiated goods (i.e. goods differentiated by national origin, and countries specialising in various goods), in

¹ This stream of empirical literature stems from the seminal paper by Imbs and Wacziarg (2003) which boosted empirical research on evolution of diversification process in countries with respect to their level of income per capita. See working paper version of this work for a detailed literature review or a useful summary of the state of the art in Cadot et al. (2013) and Cadot et al. (2011b).

which each country is assumed to produce a fixed variety of goods.² In contrast, given our empirical results, we shall refer to monopolistic competition models present in ‘new’ trade theory (among others: Krugman 1980, 1981; Helpman and Krugman 1985 plus ‘new-new’ extensions by Melitz, 2003; Hummels and Lugovskyy, 2009) which focused on benefits obtained from international trade, stemming from expansion of available product variety. Our results are also in line with endogenous growth theory (Aghion and Howitt, 1999), and especially its stream on expanding product variety (e.g. Grossman and Helpman, 1991 – chapters 3, 8, 9), where growth goes hand in hand with diversification, both in consumption and in production, reflected in exports. In turn, Romer's (1994) model displays an extensive import margin.

Indeed, Figure 1 demonstrates that developed countries typically have more diversified import and export structures than developing economies and that, in general, major heterogeneity of exported goods is positively correlated with a highly diversified basket of imported products. While exploring in detail how the heterogeneity of trade baskets changes along with a rise in income per capita, we also have in mind nonhomothetic preferences (Markusen, 2013; Tarasov, 2012), implying that there is a systematic variation in the categories of goods demanded at different income levels and as such being at the basis of structural-change.

[Figure 1 about here]

Empirically, product diversity can be analysed from two different perspectives, depending whether we focus on *(i)* the degree of economic activity concentration assessed versus uniform distribution in a country (with no reference to world distribution) or *(ii)* relative specialisation (relative diversification) of economic structures of individual countries assessed with respect to

² The assumption that products are differentiated by their country of origin, but the number of varieties supplied by each country is fixed, is known as the ‘Armington assumption’. Hummels and Klenow (2002) point out that the predictions of Armington's model, ignoring extensive margin of trade, are at odds with the data (e.g. missing two-thirds of how larger economies export more and one-third of how they import more). Acemoglu and Ventura (2002) extend Armington's view, adding endogenous capital accumulation and endogenous number of varieties. In equilibrium the number of varieties produced by a country is proportional to its employment, so countries with more workers produce and export more varieties, but there is no room for extensive import margin (all countries import all varieties).

overall (world) benchmark³. Empirical literature on both aspects of diversification and its link to economic development has been expanding rapidly in the recent years (within the first stream: Imbs and Wacziarg, 2003; Koren and Tenreyro, 2007, Klinger and Lederman, 2006; Cadot et al., 2011a; Agosin et al., 2012; within the second stream, more closely related to our paper: de Benedictis et al., 2009; Parteka, 2010; Parteka and Tamberi, 2013). However, the main observation based on the examination of current state-of-the-art is that some aspects of the phenomenon still remain unexplored. Compared to the considerable effort made to investigate production or export diversification patterns across countries, empirical evidence concerning the diversification of *imports* is much more scarce, and an evident research gap can be observed here. In particular, there are no cross-country studies presenting product-wise import diversification in the context of development process (complementing the analysis concerning exports: Cadot et al., 2011a)⁴.

Consequently, the aim of this paper is to fill in this research gap and provide some extensions on export-import diversity. We draw on highly disaggregated import statistics (HS6 – subheadings), matching them with equally detailed data on exports⁵ and various country-specific characteristics. Our sample (163 countries observed over 23 years) enables us to compare patterns of relative import and export diversification visible at very different levels of economic development. By matching import and export product level data we are also able to trace changes in the degree of similarity between the two flows, as well as qualitative differences between imported and exported products subsequently added to trade portfolio as countries develop.

³Note that high degree of absolute product diversification implies low degree of product ‘concentration’; similarly ‘relative diversification’ (de Benedictis et al., 2009) is opposed to ‘(relative) specialization’ (Amiti, 1999), so throughout the paper these terms are used as antonyms.

⁴ Most studies concerning diversification of imports are country-specific and deal with micro-level consequences of increased variety of imported products in terms of welfare and/or productivity gains (see Cadot et al., 2013 for a review). In panel data setting Jaimovic (2012) and Cadot et al. (2011c) explore *geographical* side of imports diversification process (concerning concentration of imports across origin countries) but not the diversity of imported products.

⁵ In order to overcome the well-known problems in self-reported export flows, we use mirrored data for exports. Our analysis can be read in parallel with similar export-focused research on diversification (e.g. Cadot et al., 2011a, also drawing on HS0 mirrored product level data).

The rest of the paper is structured in the following way. In Section 2 we present the data and the methodology used to analyse trade diversification patterns, along with some descriptive statistics. Section 3, being the core of the paper, is devoted to the exploration of changes in import and export diversification (specialisation) along the path of economic growth. Finally, we summarize our contribution and present the final conclusions in Section 4.

2. Empirical Setting

2.1 The data

In order to calculate synthetic measures of import and export diversification (defined below) we use trade statistics from UNComtrade database (accessed through WITS) at the highest level of disaggregation available within the Harmonized System of goods classification (HS6 corresponding to sub-headings). We use the data for the years 1988-2010 and, after performing product lines concordance exercise across various revisions of the HS data and eliminating the goods which are never traded (see Data Appendix for all the details), we are left with 4963 product lines. Country-level alternative indices of both import *and* export relative diversification are then matched with data on levels of income per capita (PPP adjusted, const. 2005 int.USD) obtained from the World Bank's World Development Indicators database. Our panel⁶ is composed of 1905 country-year observations: 163 countries (number of observations by country group is reported in Table 1) for which it was possible to match import and export data observed across the years 1988-2010).⁷

Additional country-specific characteristics, potentially important in explaining diversification patterns (such as population size, share of fuel exports – are obtained from the World Bank WDI database or geographical data obtained from Gallup et al., 1999).

⁶ See working paper version for a detailed description of panel composition and the list of the countries. Altogether, the countries included in our analysis (annual mean 1988-2010) correspond, on average, to 76% of world trade (imports): from 22 % (in 1988) to 92% (in 2006).

⁷ The actual number of countries and observations used in the estimations is slightly lower - in the estimations we do not take into consideration the evident outliers (defined, on variable-per-variable basis, as observations below 1st or above 99th percentile), corresponding to possible errors in the original data, misreported trade values or countries with extremely low/high income per capita

2.2 Adopted trade diversification measures and emerging stylised facts

2.2.1 Number of imported/exported goods, relative extensive margin of trade

Following the contribution of Hummels and Klenow (2002), recent literature related to our study acknowledges the fact that diversification of traded goods can take place at alternative margins. The *intensive* margin of trade describes diversification changes within a set of already traded goods, while the *extensive* margin concerns newly traded (or disappearing) goods. To put it simply, import (export) diversification at the extensive margin can be measured by a change in the number of active import (export) lines. Therefore, in the first case we calculate, for each country and year, the number of active product lines with non-zero import (export) flow $N_{imp_{it}}$ ($N_{exp_{it}}$). As an alternative, we relate the number of active import lines $N_{imp_{it}}$ reported by country i to the number of products effectively imported at the world level, obtaining $RelN_{imp_{it}} = N_{imp_{it}} / N_{imp_{WLDt}}$. In the same way, we obtain $RelN_{exp_{it}}$. However, not all the goods are of the same importance in the world trade, so apart from crude counting of active import/export lines (i.e. assigning them equal importance regardless of their share in world imports/exports), we compute measures of extensive trade margin (as in Hummels and Klenow, 2002): $EM_{imp_{it}}$ and $EM_{exp_{it}}$. In Table 1 we report averages of these indices, separately for imports and exports, in all sample and in the subgroups of countries divided by income (along with their crucial characteristics).

[Table 1 about here]

As for the average number of active import lines in the whole sample of 163 countries (column 1 of Table 1), it is relatively high, with 3917 non-zero import lines per country per year (theoretical max=4963 HS0 lines) which corresponds to 80% of products imported at the world level. In comparison, the average number of exported products is much lower (2713 per country per year) and reaches only 56% of the set of goods actively exported in the world. Also, the average measure of the extensive export margin is lower than in the case of imports (0.70 versus 0.85).

Looking at different countries from the perspective of their level of development (columns 2-6, groups presented in ascending order of average income per capita), it is evident that, mainly with regard to exports, diversification stages can occur. The number of exported products tends to grow together with income per capita: average $N_{exp_{it}}$ ranges from 1015 (21% of world set of exported products) in low income countries to 4066 (83% of world set of exported products) in high income – OECD members countries. Similarly, looking at the variation in $EM_{exp_{it}}$ across income groups, it can be seen that, on average, the share of products belonging to export portfolio in world exports of low income countries is approximately two times lower (0.42 versus 0.87) than in high income (OECD) countries. Please note that the considerable difference in the number of exported products and the extensive margin of exports between high income-OECD and non-OECD members, despite their similar levels of average income per capita, is due to oil-abundance and/or very small size of countries (usually implying concentrated export structures).

Actually, the comparison between product variety of exports and imports at various levels of development is very informative – while there is substantial room for export diversification at the extensive margin (by increasing the number of exported products) in the case of developing countries, the degree of import diversification is much more stable across income groups. Quite surprisingly, even low income countries (column 2 of Table 1) already cover in their import portfolio 63% of products imported at the world level (3063 product lines), while they export only 21% of all the goods exported internationally. It is not surprising, however, that the number of imported products expands as countries move to higher income categories. High income OECD countries are characterised by the biggest average number of imported products (4467 active import lines corresponding to 91% of international set of imported goods), but the process is not as evident as in the case of exported goods.

2.2.2 Synthetic measures of relative diversification (specialisation) of imported and exported products

In order to measure the degree of trade diversification (or its opposite, i.e. concentration or specialisation), researchers usually rely on inverse indices based on inequality/concentration indicators⁸. Standard indices, such as: Herfindahl index, Gini index or Theil index provide the information on how far trade flows of individual countries are from the uniform distribution (i.e. they indicate whether the imports/exports of a country are concentrated or not, in terms of product shares). At the same time, being *absolute* measures of diversification, they have the disadvantage of isolating country-specific trade patterns from those typical for world structure of trade. Our choice, instead (following Amiti, 1999; Cowell, 1995, de Benedictis et al, 2009; Parteka, 2010; Parteka and Tamberi, 2013) is to apply their variation in the form of *relative* measures: Relative Theil index, Relative Gini index and Dissimilarity Index.⁹ They involve measuring the degree of trade flow diversification in a country with respect to the world structure of trade (denoted as *WLD*) and thus account for changes in relative importance of products and indicate how different a country's distribution of product shares is from the distribution of shares typical in world structure of trade.

Our preference for relative measures is justified by several factors. The main advantage is that by using them we do not calculate mere degree of product concentration in an 'isolated' country (as conventional Theil or Gini index in its absolute form would do) but we are able to account for trends common to all countries (i.e. changes in prices of commodity goods), and we do not detach country-specific trade diversification (specialisation) patterns from those typical for the world trade. Moreover, we deal with highly disaggregated trade data (HSO, 6-digit data, subheadings) across 23 years, when changes in the relative importance of products undoubtedly occur at the world level, which would not be captured if we used 'absolute' indices retaining *n*

⁸The discussion on potential characteristics of "ideal" measures and real limits of existing ones can be found in (among others): Bickenbach et al. (2010); Palan (2010).

⁹For comparison between the application of absolute and relative diversification measures to the study of economic diversification see the discussion in Parteka (2010).

constant as the number of products which are theoretically subject to international trade and present in the classification scheme.¹⁰

We calculate the indices as follows.¹¹ Denoting product by $k=1, 2, \dots, n$ ¹²; country by i and time t we first compute import shares of a given product in total country imports ($s_imp_{ikt} = imp_{ikt} / IMP_{it}$) and the ‘typical’ share of product k in total imports reported at the world level ($w_imp_{kt} = imp_{WLDkt} / IMP_{WLDt}$).¹³ Then, Theil relative index for imports is given by

$$(Parteka \text{ and } Tamberi, 2013): \quad RelT_imp_{it} = \sum_{k=1}^n \left(s_imp_{ikt} \cdot \ln \left(\frac{s_imp_{ikt}}{w_imp_{kt}} \right) \right). \quad \text{Our second}$$

measure - relative Gini index of import diversification is approximated by (Amiti, 1999):

$$RelG_imp_{it} = \sum_{j=0}^{n-1} (p_imp_{jt} - q_imp_{ijt}) / \sum_{j=1}^{n-1} (p_imp_{jt}) \quad \text{where}$$

$$p_imp_{jt} = \sum_{k=1}^j w_imp_{kt}, \forall j = 1, \dots, n \quad \text{and} \quad q_imp_{ijt} = \sum_{k=1}^j s_imp_{ikt}, \forall j = 1, \dots, n. \quad \text{Finally,}$$

Dissimilarity Index of relative product diversification in imports is obtained as (Parteka, 2010):

$$DI_imp_{it} = \sum_{k=1}^n |s_imp_{ikt} - w_imp_{kt}|. \quad \text{We perform calculations in the same way with regard to}$$

exports. All the three indices are positively related to the degree of relative product-wise specialisation and inversely related to the degree of relative diversification. Consequently, higher values of $RelT_imp_{it}$, $RelG_imp_{it}$, DI_imp_{it} (or $RelT_exp_{it}$, $RelG_exp_{it}$, DI_exp_{it}) are associated with

¹⁰Actually, Bickenbach et al. (2010) show that only relative measures can account for general long-term changes in size distribution across industries (in our case – products). Still, it is always possible to demonstrate that absolute and relative diversification measures differ in levels but are strongly correlated (we calculated that in our case the coefficient of correlation between absolute and relative Gini index equals to 0.90 in case of imports and 0.97 in case of exports).

¹¹ See also working paper version for full exposition of the formulas and specific details on indices computation.

¹²We consider the theoretical number of all products listed within our level of disaggregation (HS subheadings) in our cleared data set (see Data appendix), so that $n=4963$.

¹³ Given unbalanced nature of our panel, our reference point is given by the world and not the sample of countries in our dataset. As values of world trade of any particular type of goods we use direct data from UNComtrade on imports of this product, where reporter = all countries. This value does not always coincide with the sum of imports of all the countries treated separately, which could influence the values of w_imp_{kt} . As an alternative to taking direct data on world imports from UNComtrade we could have summed the imports of every product across reporting countries. However, we have compared a series of w_imp_{kt} obtained with the use of these two alternative ways: they are highly correlated (0.99) and the differences between them are negligible.

relatively less diversified (more specialised) structures of imports (or exports) with respect to the world composition of trade.

In Table 2 we show summary statistics of these three synthetic indices of relative diversification (accompanied by statistics referring to the number of imported/exported products), calculated for the whole sample of countries (excluding outliers). The comparison between mean values referring to imports and exports clearly proves that, in general, imports are characterised by a much higher degree of relative product heterogeneity than exports. An average country in our sample (outliers excluded) covers, in its import portfolio, 81% of all goods actively traded in the world; on the export side - around half (55%). Similarly, average values of synthetic indices of relative import diversification are always lower than those of exports. Table 3 contains analogous summary statistics but calculated within five distinct income groups (listed in ascending order of income per capita). Synthetic measures of both import and export specialisation (inversely related to the degree of product diversification) tend to diminish as we move towards higher income country groups while the number of imported/exported products rises. Hence, trade diversification process proceeds as income levels rise. Generally speaking, economies characterised by higher level of development have typically more diversified import and export structures.

[Table 2 about here]

[Table 3 about here]

3. Stages of import and export relative diversification in the process of economic development

3.1 Revealed relative diversification curve - unconditional non-parametric results

In order to verify how trade structures change along the path of economic growth, in the first instance we match synthetic inverse measures of relative diversification of countries from our

sample with their corresponding levels GDP per capita. *A priori*, given potential nonlinearity¹⁴, we do not want to impose any functional form on the relationship between variables of interest. We plot unconditional non-parametric lowess curves, where each plot corresponds to a non-parametric equation in the following form:

$$DIV_imp_{it} = f(GDPpc_{it}) + \varepsilon_{it} \quad (1)$$

where DIV_imp (and, in a parallel model, DIV_exp) denote one of the synthetic inverse measures of relative diversification (respectively for imports and exports) i refers to countries and t to time period, ε is an error term; $GDPpc$ is a proxy of the development level (real income per capita) while $f(.)$ is an unspecified function estimated with the use of the lowess smoother (Cleveland, 1979; Hastie and Tibshirani (1990, p.30) and represented graphically.¹⁵

In Figure 2 we demonstrate non-parametric relationship (lowess curves) between the degree of import product-wise relative diversification (employing our three baseline relative measures of diversification) and levels of economic development (left plots) – representing our original contribution, compared to analogous patterns obtained with the use of export data (right plots). Each plot shows two lowess lines: the grey one, obtained from all the country-year observations (with the exception of evident outliers only) and the black one, obtained from the observations corresponding to levels of income per capita ranging below 40,000 USD (PPP, const.2005). Why have we decided to split the sample in such a manner?

[Figure 2 about here]

We were interested in checking whether the U shaped pattern of diversification, found by Imbs and Wacziarg (2003) for production or by Cadot et al. (2011a) for exports and obtained with absolute measures of diversification is confirmed in our sample which matches import and export data in a relative setting. In fact, when we consider the whole group of countries, we also obtain the U curve of trade diversification (in Figure 2 it is represented by the grey line, which

¹⁴ This is in line with Cohen-Cole et al. (2012) who provide evidence on the importance of nonlinearities in growth process.

¹⁵ All the plots are corrected for the presence of evident outliers - defined as observations below 1st or above 99th percentile (possible errors in the original data, misreported trade values etc.).

partly overlaps with the black one). Independently, on the index, the line is first decreasing (which suggests a relative diversification course at lower levels of economic development) and then, having passed the level of income per capita around 30,000 USD (PPP, const. 2005) it starts to increase (which suggests relative re-specialisation of trade structures at higher stages of development). This pattern is similar both for exports (plots on the right) and imports (plots on the left).

However, following some data mining, we actually found out that the upward rising part of the U curve might be guided by a limited number of observations (namely: only 44 out of 1828 in our sample without outliers), corresponding to a few specific countries with the levels of income per capita roughly above 40,000 USD (PPP, const. 2005). They are marked in Figure 2 with triangles. We list them in ascending order of income per capita, along with their crucial characteristics in Table 4. The U.S.A. is one of these countries: as we will show later on, it has quite a distinct pattern of trade, as it follows the path of (slight) relative specialisation with respect to the overall benchmark. Among other countries responsible for the upward part of the non-parametric U curve, there are mainly quite specific economies, i.e. either those of oil-exporting countries (such as Norway, Kuwait or Brunei Darussalam) and/or small countries (in terms of population) which, in general, are likely to have more concentrated (specialised) trade structures than other countries. Lowess curves of relative diversification based on the remaining observations (97% of all the observations in our sample) are clearly decreasing, which shows that, in general, a vast majority of countries tend to diversify both their import and export portfolio as they develop.

[Table 4 about here]

3.2. Country-specific patterns of relative import and export diversification

Individual countries greatly differ from each other because of their size, institutions, geographical characteristics (which influence, among other things, the kind and variety of

agricultural products), natural resources, endowments, economic policies (openness policies), etc., and all of these factors can eventually influence the shape of the curve obtained from the data.¹⁶

Examination of country-specific patterns of relative trade diversification¹⁷ demonstrates that specificity of countries can indeed play a role in the diversification process. To illustrate this, we show in Figure 3 an example based on two developed countries (AUS – Australia, and the USA) and two developing ones (ROM – Romania and CHN – China), included in our sample. We simultaneously show how the degree of relative import diversification (measured inversely by Relative Theil index: $RelIT_imp$) and relative export diversification ($RelIT_exp$) evolve with respect to the level of income per capita in each country. We maintain the same scale for all the countries to show changes in diversification indices in a comparative setting.

[Figure 3 about here]

It is evident that specific countries may display specific patterns of changes in trade composition and the level of income per capita is not sufficient to explain it. In the case of ROM, CHN and AUS imports are more diversified than exports ($RelIT_exp > RelIT_imp$; the pattern already revealed in the whole sample), in the USA the difference is negligible (very low values of both $RelIT_exp$ and $RelIT_imp$ indicate the degree of export and import diversification which is already very high). Another highly developed country –AUS - is characterised by a very diversified import structure, but relatively concentrated exports (a hint that its location may influence trade costs and hamper export diversification opportunities). In terms of evolution along the path of growth, Romania demonstrates quite an evident diversification trend of both imported and exported products as its income per capita grows. Another developing country, China, was initially diversifying and later (having reached the income per capita of approx. \$4000) very slight re-specialisation of exports began. A very slight U-shaped pattern is typical also for

¹⁶An analogy can be found here with another very famous U-curve in economics; the inverted-U ‘Kuznets curve’. One of the strong objections to the existence of the inverted-U shaped curve derived from the use of country-fixed effects in the estimation. The inverted U-turn was first empirically questioned by Deininger and Squire (1998), who found that the use of country-specific effects largely wiped out the bell shape of the inequality-development relationship; after that, the use of country-specific effects became unavoidable.

¹⁷Taking into consideration the unbalanced nature of our panel we have examined the cases for which we dispose of more than 18 observations within the years 1988-2010.

AUS, but clearly with an entirely different turning point than in the case of CHN. Changes in product diversity in the USA are negligible when compared to other countries, and only a slight increase in $RelT_{exp}$ and $RelT_{imp}$ after reaching the level of \$40,000 – please note that possible re-specialisation is very small in magnitude.

3.3 Incorporating country fixed effects into general relative diversification curve: semi-parametric results

Given just the presented importance of country specificity, which must be taken into account when revealing the general diversification curve (demonstrating a ‘typical’ trend in the sample), we first consider the most natural extension of the unconditional lowess model, i.e. incorporation of country-fixed effects into the estimation of non-parametric diversification curves (as in: de Benedictis et al., 2009 and Parteka, 2010). In order to do so, we apply semi-parametric Generalised Additive Model (GAM, see Hastie and Tibshirani, 1987, 1990) of the following form:

$$DIV_imp_{it} = f(GDPpc_{it}) + D_i + \varepsilon_{it} \quad (2)$$

where D_i denotes country dummies and $f(.)$ is an unspecified function estimated with the use of the lowess smoother, estimated from the data through a “backfitting” procedure, formally represented as Gauss-Seidel algorithm and consisting of iteratively smoothed partial residuals (see Buja et al., 1989 for formal details on the procedure), represented in a graphical form as a plot of partial residuals.¹⁸ All the other components are as in eq.(1). In a parallel model, we employ DIV_exp as left hand side variable.

[Figure 4 about here]

In Figure 4 we show plots of partial residuals of the GAM model obtained with alternative synthetic measures of relative import and export diversification. With regard to both imports and exports, semi-parametric curves are decreasing (with only minor differences in

¹⁸In the robustness check section we additionally check the robustness of model (2) to the inclusion of both country- and time-fixed effects. Please note that in principle relative measures of diversification which we already use shall already correct for time-specific and common to all countries changes in the diversification process.

shapes depending on the measure used) – an increase in income per capita goes hand in hand with a decrease in the indices of relative specialisation. Hence, the general result is that, when we correct unconditional non-parametric lowess curves for the presence of country-fixed effects, trade diversification is the predominant tendency in our wide sample of countries, with only a negligible possibility of trade re-concentration at higher levels of development (more visible in the case of imports). In other words: we do not exclude the possibility of re-specialisation course with regard to single countries, but the trend based on the predominant mass of country-year observations argues in favour of treating import and export relative diversification as a ‘typical’ tendency in the course of economic growth.

3.4 Parametric results – country-specific factors and relative diversification process

In order to provide further confirmation for the robustness of our findings concerning progressing relative diversification of imports and exports in the process of economic growth, in the following step we confront the revealed diversification curves obtained through GAM with parametric estimates of the following model:

$$\ln DIV_imp_{it} = \alpha + \beta(\ln GDPpc_{it}) + D_i + D_t + \varepsilon_{it} \quad (3)$$

where DIV_imp (alternated with DIV_exp) and $GDPpc$ (income per capita) are introduced in natural logarithms in order to mimic the shape of semi-parametric diversification curves. D_i stands for country-specific fixed effects and, additionally (in alternative formulations), time dummies D_t are additionally included to control for common business cycle effects (for example the ones affecting world prices of imported/exported products) and ε stands for the error term. We address the issue of potential endogeneity between DIV and $GDPpc$ or additional country-level controls in the model in the robustness check section.

Table 5 presents results of FE estimates¹⁹ referring to relative diversification of imports (upper panel) and exports (lower panel). Alternative models employ the use of different inverse indices of diversification; specifications (4) to (5) include time effects. In both cases – imports

¹⁹ OLS results are reported in working paper version. They do not differ significantly from FE ones.

and exports – the estimated β coefficient is negative and statistically significant, which confirms the trend visible in relative diversification curves revealed through semi-parametric estimation.

[Table 5 about here]

3.5 Interpretation of the results and import-export similarity in the diversification process

We have already demonstrated that diversification process of both exports and imports accompanies economic growth which is in line with models emphasising ‘love-for-variet’y both in consumption and in production, and with contributions underlying the importance of extensive margin of trade (see the Introduction). This implies that especially at higher income per capita levels, single countries despecialize and actually export *and* import the majority of goods which are traded around the world. Hence, they import and export the same categories of goods. So, the intuitive view that countries export what they can produce and import what they cannot produce can be very misleading. In order to prove it, we measured the degree of similarity between import and export structures, considering two classes of measures: first based on the binary data (assigning 1 to active product lines, being the lines with non-zero trade flow, and 0 otherwise; see Finch, 2005 for discussion), and second employing the comparison between the shares of active exported goods in total exports (s_exp_{ikt}) with the shares of active imported goods in total imports (s_imp_{ikt}). Within the first class (import-export *similarity* indices), we compute: Russel-Rao index (RR), Jaccard index (J), as well as Sokal and Michener index (SM). As shown in [Table 6](#) they are all positively correlated with income per capita levels and negatively correlated with our synthetic measures of import and export diversification. Additionally, within the second class of measures, we consider such dissimilarity indices as: Michaely-Krugman Index, MK and Battarcharyya Index, B which higher values indicate higher *dissimilarity* between import and export structures. In turn, as expected, MK and B being the opposite of similarity indices, are negatively correlated with GDP per capita and positively correlated with the degree of relative trade diversification. To sum up, as income per capita grows and trade baskets become less

specialized, the similarity between import and export structures increases and countries tend to export and import the same categories of products.

[Table 6 about here]

Does this mean that after a successful diversification countries import and export exactly the same goods? Not necessarily. As argued by Schott (2004), there is a difference between across-product and within-product specialisation in international trade: he shows that United States increasingly sources the same product categories from different (both high- and low-wage) countries, but specialisation (in terms of characteristics) occurs within products. We can add another piece to this story: we show that as the diversification process accompanies economic development, countries increasingly export and import the same categories of goods, but imported and exported goods vary in terms of characteristics. To prove this we employ Revealed Factor Intensities from UNCTAD (Shirotori et al., 2010) to compute average human capital content (as a proxy of quality – see Schott, 2004) of: (i) product lines being *active* in import/export portfolio, (ii) of product lines being annually *added* to trade baskets (new import/export lines)²⁰. As shown in Figure 5 (left plot), in line with nonhomothetic preferences’ approach (Markusen, 2013; Tarasov, 2012), consumption in richer countries deviates from that in poorer countries which here is reflected in rising quality of goods demanded internally and externally (imports). Importantly, at almost all levels of income per capita, human capital content of imported goods is higher than that of exported products (i.e. countries import goods of higher quality than those they are able to produce and export). Similarly, as countries develop, they add to their export and import portfolio (Figure 5, right plot) goods of higher and higher quality, but human capital content of products new in expanding export portfolio is always lower than that of new imported goods. Countries may diversify by exporting and importing the same product categories, but not exactly the same goods in terms of quality characteristics.

[Figure 5 about here]

²⁰ In order to define new product lines we use moving window procedure (with two-year cutoff to define the spell) similar to the one adopted in Cadot et al. (2011a) and based on trade survival concept.

3.6 Robustness checks

We have checked the robustness of our main finding – progressing relative diversification/despecialisation of imports (and exports) along the path of economic growth – in several ways.²¹

With regard to the calculated indices of diversification, we calculated standard measures of export concentration (absolute Gini index, absolute Theil index) with our set of product lines and compared them to the series of indices employed in Cadot et al. (2011a) study, which were obtained with the use of a wider set of 4991 products.²² There are slight differences in levels (our indices are usually higher), but the correlation coefficients between our measures and those of Cadot et al. (2011a) are very high (0.9951 for Theil index and 0.9974 for Gini index, $n=2797$). Secondly, we recalculated our measures of trade diversity excluding oil/petrol-related commodities.²³ The two types of measure (*i.* calculated with all product lines and *ii.* without oil/petrol) are strongly correlated – between 0.95 and 0.99.

With respect to the estimated models we modified, first of all, the non-parametric and semi-parametric estimates of relative diversification curve by altering the span parameter (lowering it to 0.5 in lowess models and rising to 0.7 in GAM estimations). Next, we checked the shape of the benchmark semi-parametric curves for the inclusion of time fixed effects D_t (complementing D_i). We also run parametric regression models for sub-sample of countries – eliminating from eq. (3) the microstates (countries with population below 1mln) and oil/petrol-abundant countries (with oil/petrol exports above 50% of total exports). We have also estimated the extended equation (augmented by other country-specific characteristic):

$$\ln DIV_imp_{it} = \alpha + \beta(\ln GDPpc_{it}) + \sum_{k=1}^K \delta_k X_{ik(t)} + D_t + \varepsilon_{it} \quad (4)$$

²¹ All the results referring to this section can be consulted in the Appendix.

²²We thank Céline Carrère for providing us with absolute diversification indices used in Cadot et al. (2011a) paper.

²³In order to do so, we removed from our set of products the following HS0 product lines: 27900, 271011, 271119, 271129, 271210, 271311, 271312, 271320, 271390.

[and similarly employing DIV_exp] where $X_{ik(t)}$ denotes a set of k country-specific factors (time variant such as GDP or time invariant, such as geographical features). In line with the empirical literature (see Cadot et al., 2013, section 4) and suggestions of the theoretical models we consider: country size - measured alternatively in terms of gross output - GDP and population - POP (bigger countries should have more diversified economic structures), $DISTANCE$ being the distance from the main markets (directly affecting access to foreign market and trade costs) and $PETROL$ - the importance of oil/petrol in economic structure of countries (expecting it to be positively linked to trade concentration, at least on the side of exports).²⁴ All of the coefficients have the expected signs and the general result (significant and negative β parameter) holds even when additional controls are included into the model.

We run additional regression eliminating oil/petrol-related products from the set of goods used in calculation of diversification measures; we also checked whether the elimination of observations from the total sample, performed due to the necessary matching between import and export data sets influenced the results. On top of that, we considered a quadratic model with income per capita squared as a right hand side variable (as in: Imbs and Wacziarg, 2003 or Cadot et al., 2011a). The estimated turning point is very high (between approx. \$29,000 PPP2005 for imports and approx. \$35,000 PPP2005 for exports). The percentage of country-year observations lying above is very low; up to 98% of data points and 160 countries lie on the decreasing side of the hump shaped curve. Finally, following the findings of Hesse (2009) we addressed the issue of potential endogeneity in the parametric model, resulting from plausible two-way relations between diversification and economic development.

None of these modifications alter significantly the conclusions drawn with regard to the benchmark results.

²⁴Please note that in this case we do not include country-fixed effects as they wipe out most of the cross-country variability and, additionally, are collinear with time invariant characteristics, such as market distance.

4. Summary of findings and concluding remarks

The main aim of this paper was to contribute to empirical trade diversification literature by adding a missing piece – i.e. imports - to the already analysed export trends. In particular, we compare the evolution of import and export diversification in the process of economic development, assessed in a relative setting with respect to changes in the world distribution of product shares. Consequently, our study constitutes natural extension of literature on relative trade diversification (de Benedictis et al., 2009; Parteka, 2010; Parteka and Tamberi, 2013) in the sense that: *(i)* apart from providing evidence for exports we also analyse relative diversification of imports which allows us to test alternative theoretical models, *(ii)* we perform our analysis with highly disaggregated product level data (rather than sector level data), *(iii)* the study is not limited to manufacturing and *(iv)* it is performed for a much larger panel of international economies than it used to be done.

We merge country-level data on income per capita with a wide set of alternative relative diversification indices calculated with highly disaggregated import and export statistics (4963 product lines) for 163 countries across the time period 1988-2010. To the best of our knowledge, this is the first attempt to explore *jointly* the patterns of diversification of imported and exported goods in the context of development process in a wide set of countries.

We find that diversification of imports and exports differ with regard to levels, but not in terms of evolution along the economic growth path. As far as the differences in levels are concerned, we confirm intuitively expected result: since the very beginning of the development process, product heterogeneity of imports is much higher than that of exports. Even poor countries import a lot of varieties; sets of goods imported by low income countries are three times more heterogeneous than the sets of products they export – they already cover, in their import portfolio, 63% of products imported at the world level, while they export only 21% of all goods exported internationally. Consequently, at low levels of development there is room for activity at extensive margin of exports (in line with Hummels and Klenow, 2002 and Cadot et al.,

2011a) and the process of adding new lines to the export portfolio is much more dynamic than in the case of imports. In the light of our results, even poorer countries, not capable of producing (and exporting) a wide range of goods, augment (in line with ‘love-for-variety’ concept present in Krugman-type models) products variety available on the market by importing them. We can also interpret this result in the light of non-homotheticities in international trade (Markusen, 2013; Tarasov, 2012): the number of desired goods (or indeed of goods that can be produced domestically) become endogenous and depend on aggregate income per capita. This finding has also important policy implications: only relatively free trade allows poorer countries to import intensively and to increase the welfare of citizens.

In terms of evolution along the economic growth path, we find a robust tendency towards progressing relative diversification (despecialisation) of both import and export structures. Even though slight re-specialisation course is possible with regard to some specific countries, the predominant tendency revealed by the data is different: the pattern of trade diversification is more marked as countries grow rich. This finding is not sensitive to changes in the methodological setting, which we proved by numerous robustness checks. We also demonstrate that even though while diversifying countries increase the degree of import-export similarity in terms of product categories, imported and exported goods differ in terms of within-product characteristics (we thus disentangle across-product and within-product specialisation as in Schott, 2004).

With regard to the guidelines for future research, this import-export study could serve as a starting point for even deeper exploration of the topic. We focused on diversification-development nexus, leaving aside other causes of diversification, such as the process of trade integration around the world and the resulting increase in trade openness. Moreover, our findings confirm the importance of ensuring clear distinction between various ways of measuring the diversification process (relative trends vs. absolute changes in the degree of product concentration). We do not negate the results of Cadot et al. (2011a) and export re-concentration

track they reveal, but rather suggest that there could be some specific patterns concerning changes in the world structure of trade. Country specific studies on diversification of imports and exports, and reconcentration tendency could be performed for selected countries (i.e. US) to uncover what lies behind it. Finally, qualitative aspect of diversification calls for further exploration. For the time being, these issues are beyond the scope of this paper.

References

- Acemoglu D., Ventura J. (2002). The World Income Distribution. *Quarterly Journal of Economics*, 117: 659-694.
- Aghion P., Howitt P. (1999). *Endogenous Growth Theory*, MIT Press
- Agosin M.R., Alvarez R., Bravo-Ortega C. (2012). Determinants of export diversification around the world: 1962-2000. *The World Economy* 35(3):295–315.
- Amiti M. (1999). Specialisation patterns in Europe. *Review of World Economics*, 135(4): 573-593.
- Armington P. (1969). A theory of demand of for products distinguished by place of production. IMF Staff Papers, XVI: 159-178.
- Bickenbach F., Bode E. , Krieger-Boden C. (2010). Closing the Gap between Absolute and Relative Measures of Localization, Concentration or Specialisation. *Kiel Working Paper* No.1660.
- Buja A., Hastie T., Tibshirani R. (1989). Linear smoothers and additive models (with discussion). *Annals of Statistics*, 17, 453-555.
- Cadot O., Carrere C., Strauss-Kahn V. (2011a). Export diversification: What's behind the hump? *The Review of Economics and Statistics*, 93(2): 590-605.
- Cadot O., Carrere C., Strauss-Kahn V. (2011b). Trade diversification: drivers and impacts. [in:] *Trade and Employment. From myths to facts*. Jansen H., Peters R., Salazar-Xirinachs J.M. (eds.), ILO, Geneva, 253-307.
- Cadot O., Carrere C., Kukenova M., Strauss-Kahn V. (2011c). OECD Imports. Diversification of suppliers and quality search. The World Bank Policy Research Working Paper No.5258.
- Cadot O., Carrere C., Strauss-Kahn V. (2013). Trade diversification, income, and growth: what do we know? *Journal of Economic Surveys*. 27(4): 790-812.
- Cleveland W. (1979). Robust locally weighted regression and smoothing scatter plots. *Journal of American Statistic Association*, 74(368): 829-836.
- Cohen-Cole E., Durlauf S.N., Rondina G. (2012). Nonlinearities in growth : from evidence to policy. *Journal of Macroeconomics*, 34(1): 42-58.
- Cowell F. (1995). *Measuring inequality*. Prentice Hall/Harvester Wheatsheaf, London (2nd edition).
- de Benedictis L., Gallegati M., Tamberi M. (2009). Overall trade specialisation and economic development: countries diversify, *Review of World Economics*, 145(1): 37-55.
- Deininger K., Squire L. (1998). New ways of looking at old issues: inequality and growth, *Journal of Development Economics*, 57: 259-287

- Dixit A. K., Stiglitz J.E. (1977). Monopolistic Competition and Optimum Product Diversity. *American Economic Review*, 67(3): 297-308.
- Finch H. (2005). Comparison of Distance Measures in Cluster Analysis with Dichotomous Data. *Journal of Data Science*, 3: 85-100.
- Gallup J.L., Sachs J., Melinger A.D. (1999). Geography and Economic Development. *International Regional Science Review*, 22 (2): 179-232; database available at www.cid.harvard.edu/ciddata
- Grossman G.M., Helpman E. (1991). *Innovation and growth in the global economy*. Cambridge: MIT Press.
- Hastie, T.J., Tibshirani, R.J. (1987). Generalized Additive Models. Some Applications. *Journal of American Statistical Association*, 82, 371-386.
- Hastie T.J., Tibshirani R.J. (1990). *Generalised Additive Models*. Chapman and Hall, London.
- Helpman E., Krugman P. (1985). *Trade policy and market structure*. MIT Press
- Hesse H. (2009). Economic diversification and economic growth. [in:] Newfamer R., Shaw W., Walkenhorst P. (eds). *Breaking into new markets. Emerging lessons for export diversification*. Washington, DC: The World Bank. Chapter 3 (55-80).
- Hummels D., Klenow P. (2002). The variety and quality of a nation's trade. NBER Working Paper No. 8712.
- Hummels D., Lugovskyy V. (2009). International Pricing in a Generalized Model of Ideal Variety, *Journal of Money, Credit and Banking*, 44(s1): 3–33,
- Imbs, J., Wacziarg, R. (2003). Stages of diversification, *American Economic Review*, 93(1): 63-86.
- Jaimovich E. (2012). Import diversification along the growth path. *Economics Letters*, 117(1), 306-310.
- Klinger B.; Lederman D. (2006). Diversification, innovation, and imitation inside the global technology frontier, World Bank Policy Research Working Paper No. 3872.
- Koren, M.; Tenreyro, S. (2007). Volatility and development, *Quarterly Journal of Economics*, 122(1): 243-287.
- Krugman, P. (1980). Scale Economies, Product Differentiation, and the Pattern of Trade. *American Economic Review*, 70 (5): 950-959.
- Krugman, P. (1981). Intraindustry Specialisation and the Gains from Trade, *Journal of Political Economy*, 89(5): 959-973.
- Markusen, J.R. (2013). Putting per-capita income back into trade theory, *Journal of International Economics*, 90(2): 255-265.
- Melitz, M.J. (2003). The Impact of Trade on Intra-industry Reallocations and Aggregate Industry Productivity, *Econometrica*, 71: 1695–1725.
- Minondo A. (2011). Does comparative advantage explain countries' diversification level? *Review of World Economics*, 147(3): 507-526.
- Palan N. (2010). Measurement of specialisation - the choice of indices. *FIW Working Paper* No.62.
- Parteka A. (2010). Employment and export specialisation along the development path: some robust evidence. *Review of World Economics*, 145(4): 615-640.
- Parteka A., Tamberi M. (2013). What determines export diversification in the development process? Empirical assessment. *The World Economy*. 36(6): 807–826.

- Romer, P. (1994). New Goods, Old Theory, and the Welfare Costs of Trade Restrictions, *Journal of Development Economics*, 43: 5-38.
- Schott P. (2004). Across-Product versus Within-Product Specialisation in International Trade. *Quarterly Journal of Economics*, 119(2): 647-678
- Shirotori M., Tumurchudur B., Cadot O. (2010). Revealed Factor Intensity Indices at the product level. Policy Issues in International Trade and Commodities Study Series No. 44. UNCTAD, Geneva.
- Tarasov, A. (2012). Per capita income, market access costs, and trade volumes. *Journal of International Economics*, 86(2), 284-294.

FIGURES

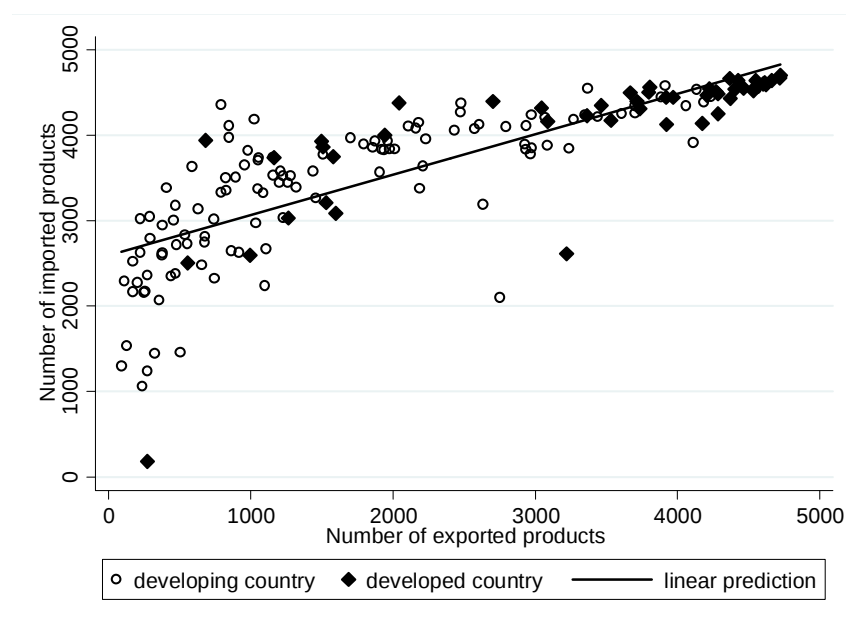


Figure 1. Diversification of imports versus diversification of exports – countries divided according to the development level

Note: Plot shows average number of imported products versus average number of exported products (theoretical max: 4963 product lines) - average values are calculated across time for 163 countries included in our sample (listed in Appendix 1A). Segregation of countries into developed and developing ones is based on World Bank (2011) classification using GNI per capita 2010 (developing countries = low and middle income countries)

Source: own elaboration based on trade data obtained from UNComtrade, 2011 (HS0, 4963 product lines).

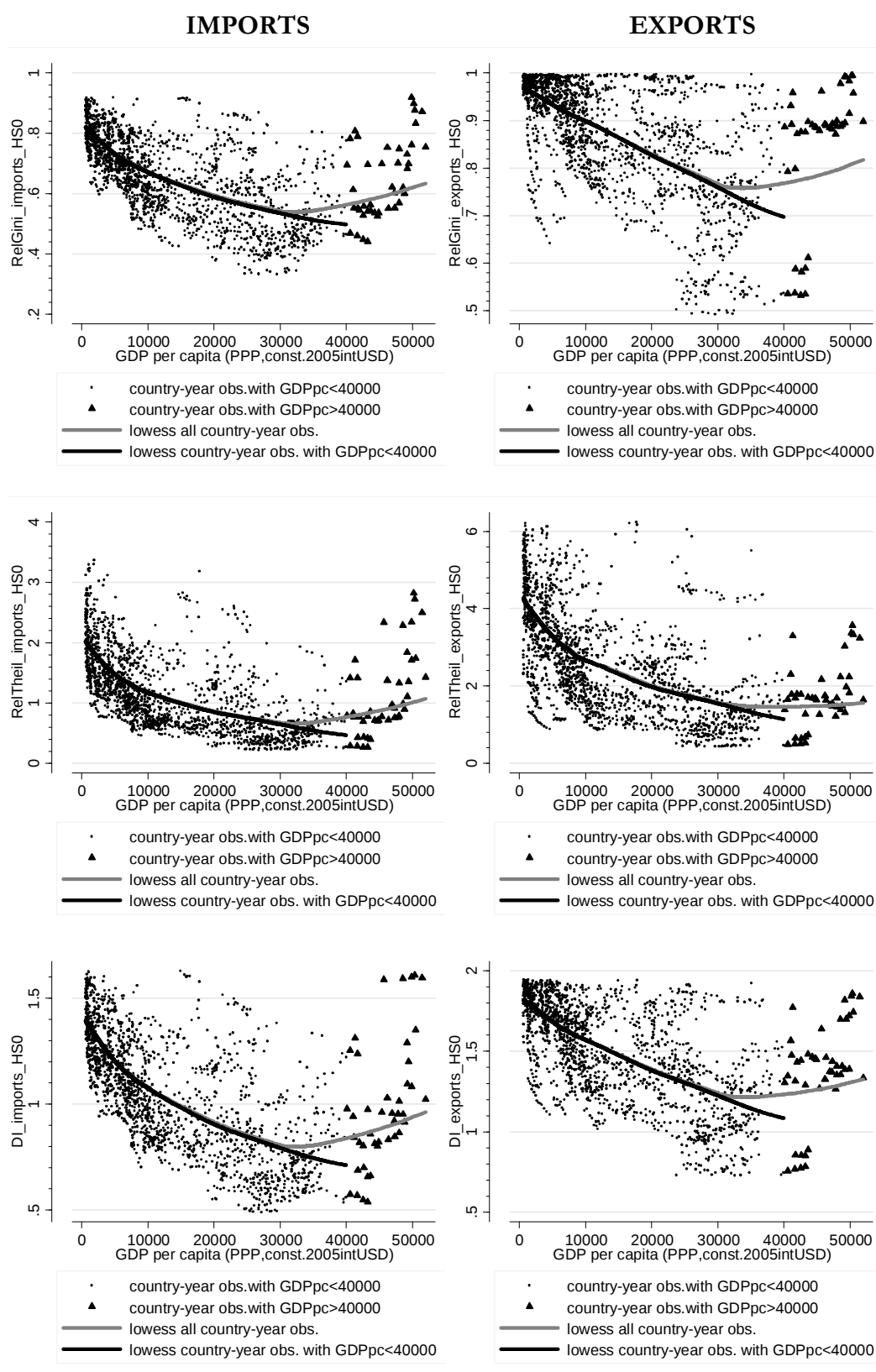


Figure 2. Degree of diversification of imported and exported products versus development level (plots on the left: imports, plots on the right: exports), alternative indices

Source: own elaboration based on trade data obtained from UNComtrade, 2011 (HS0, 4963 product lines) and GDP per capita – data obtained from WB WDI (2011).

Note: lowess - span=0.8, sample without outliers defined as observations below 1st or above 99th percentile (number of obs.=1828)

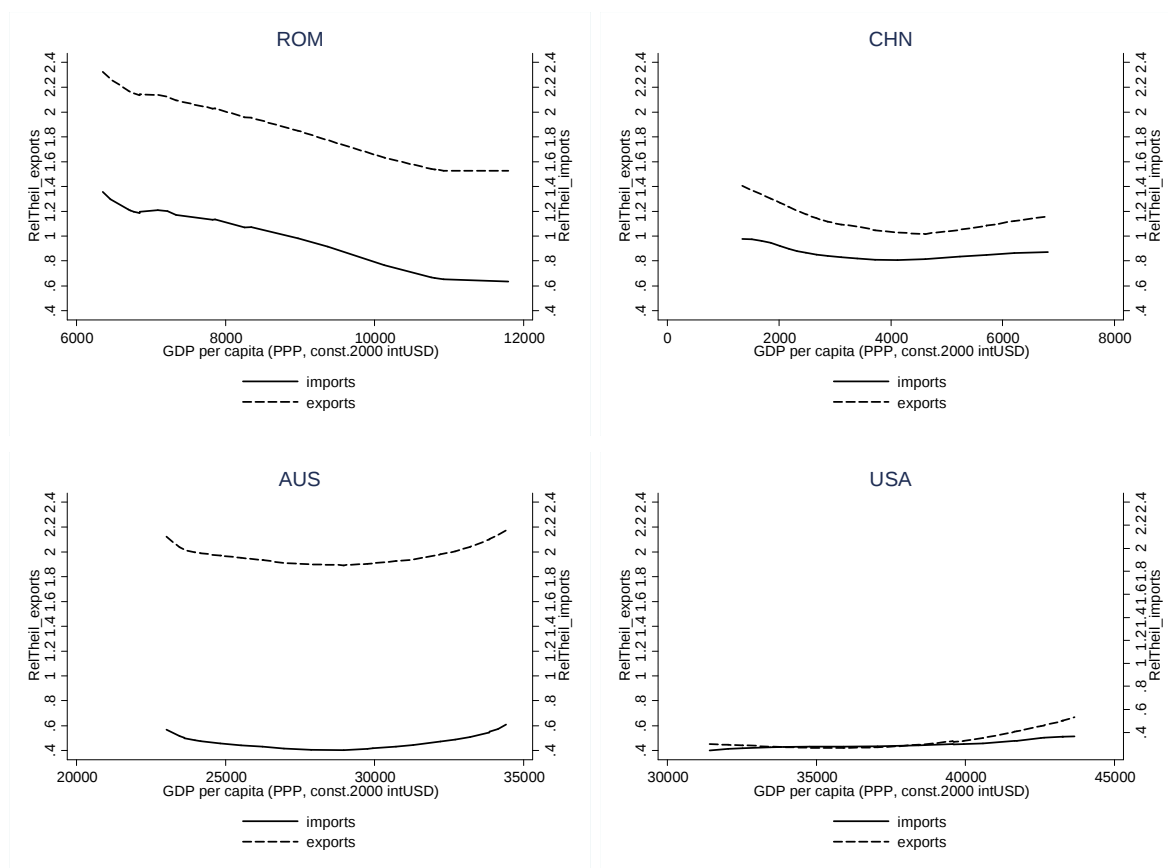


Figure 3. Degree of relative diversification of imported and exported products versus development level – selected country specific trends

Note: lowess - span=0.8. Left axis: relative Theil index of export concentration; Right axis: relative Theil index of imports concentration.

Source: own elaboration based on trade data obtained from UNComtrade, 2011 (HS0, 4963 product lines) and GDP per capita – data obtained from WB WDI (2011).

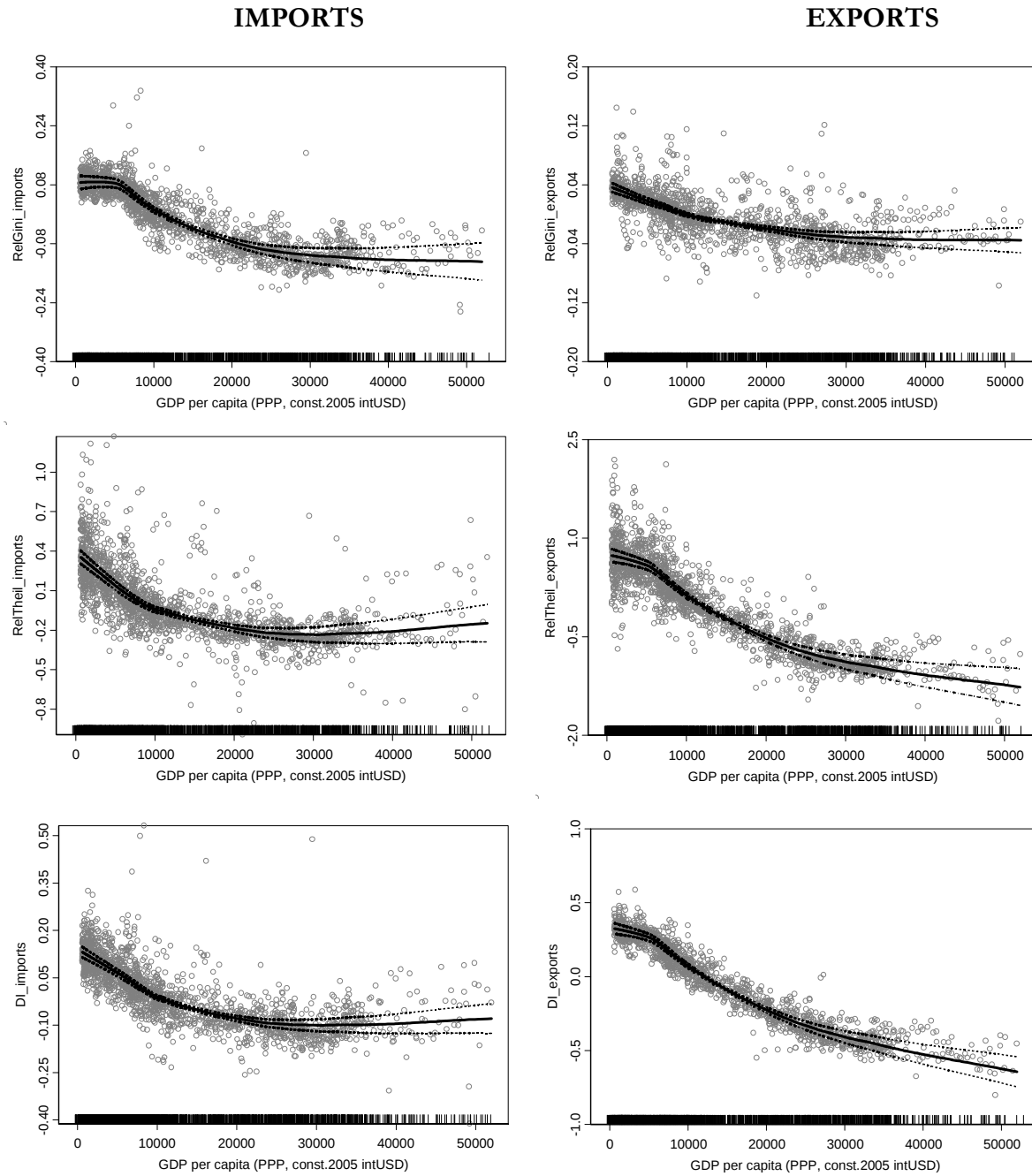


Figure 4. Degree of relative diversification of imported and exported products versus development level (plots on the left: imports, plots on the right: exports) – semi-parametric curves controlling country-fixed effects, alternative indices

Source: own elaboration based on trade data obtained from UNComtrade, 2011 (HS0, 4963 product lines) and GDP per capita – data obtained from WB WDI (2011).

Note: yaxis: inverse diversification measure - partial residuals of GAM model (eq. 5), span=0.5; sample without outliers defined as observations below 1st or above 99th percentile (number of obs.=1828). Due to FE inclusion only countries with multiple observations have been considered.

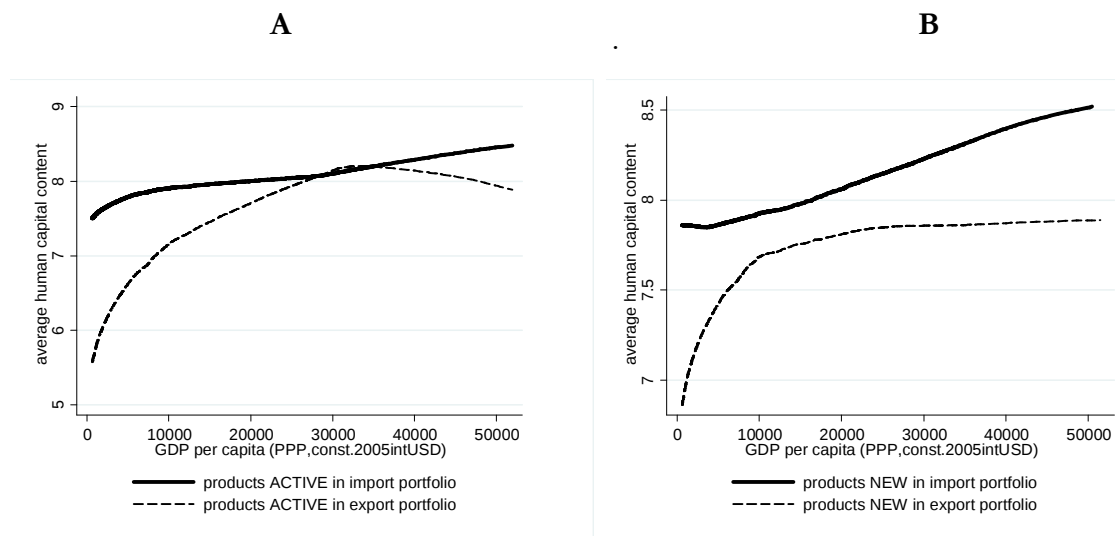


Figure 5. Average human capital content of active and new import/export lines vs. countries' income per capita levels (163 countries, 1988-2010)

Note: lines correspond to lowess approximation (outliers excluded). Active lines = lines with non zero trade value. New products defined as product lines not present in export (import) portfolio at time $t-1$, but exported (imported) at time t and $t+1$. Average human capital content goods is the average RHCI (Revealed Human Capital Index) calculated across active or new exported (imported) products, weighted by product share in export (import) basket. Source: own elaboration with trade data from UNComtrade, 2011 (HS0, subheading), income per capita from WB WDI (2011) and RHCI from Revealed Factor Intensity Indices Database from UNCTAD.

TABLES

Table 1. Number of imported/exported products, extensive margin of trade and new import/export lines – average values (overall and by income group)

	All countries	Countries divided by income group*				
	(1)	LI (2)	LMI (3)	UMI (4)	HI-nonOECD (5)	HI-OECD (6)
Total number of country-year obs.	1905	178	333	609	195	590
Number of countries	163	24	43	48	17	31
GDPpc						
[PPP, const.2005 int.USD]	14291	922.52	2971.01	8173.75	26384.50	27031.40
Population [1000]	48401	25302	80446	67029	3717	32987
Share of fuel exports [%]	13.2	3.2	14.4	16.6	29.2	6.3
Imports:						
N_imp_{it}	3917	3063	3476	3931	3746	4467
$RelN_imp_{it}$	0.80	0.63	0.71	0.80	0.77	0.91
EM_imp_{it}	0.85	0.77	0.82	0.87	0.81	0.90
Exports:						
N_exp_{it}	2713	1015	1685	2547	2303	4066
$RelN_exp_{it}$	0.56	0.21	0.35	0.52	0.47	0.83
EM_exp_{it}	0.70	0.42	0.57	0.70	0.69	0.87

Note: average values across country-year observations (163 countries, 1988-2010) *country groups according to the World Bank (2011) classification: LI- low income (2010 GNI per capita of \$1,005 or less) , LMI - lower middle income (\$1,006 - \$3,975), UMI - upper middle income (\$3,976 - \$12,275), HI- high income (\$12,276 or more; OECD and non-OECD members).Low income and middle income countries are classified by the World Bank as “developing countries”.

Source: own elaboration based on trade data obtained from UNComtrade (2011); GDP per capita, population and share of fuel exports (as % of all merchandise exports) – data obtained from WB WDI (2011).

Table 2. Summary statistics of import and export diversification measures – all countries (without outliers)

Import diversification measures							Export diversification measures						
index	group	obs	mean	sd	min	max	index	group	obs	mean	sd	min	max
N_imp_{it}	all countries	1867	3939	693	1257	4868	N_exp_{it}	all countries	1867	2690	1461	171	4858
$RelN_imp_{it}$	all countries	1867	0.81	0.14	0.26	0.99	$RelN_exp_{it}$	all countries	1867	0.55	0.30	0.03	0.99
$RelT_imp_{it}$	all countries	1867	1.18	0.64	0.23	3.38	$RelT_exp_{it}$	all countries	1867	2.63	1.37	0.44	6.24
$RelG_imp_{it}$	all countries	1867	0.66	0.13	0.33	0.91	$RelG_exp_{it}$	all countries	1867	0.88	0.12	0.49	1.00
DI_imp_{it}	all countries	1867	1.06	0.27	0.49	1.63	DI_exp_{it}	all countries	1867	1.53	0.31	0.73	1.95

Note: Sample without outliers defined as observations below 1st or above 99th percentile.

Source: own elaboration

Table 3. Summary statistics of import and export diversification measures – countries by income group

Import diversification measures							Export diversification measures						
index	group	obs	mean	sd	min	max	index	group	obs	mean	sd	min	max
N_imp_{it}	LI	173	3127	661	1257	4197	N_exp_{it}	LI	168	1067	772	178	3482
	LMI	325	3538	593	1444	4669		LMI	326	1695	1134	177	4641
	UMI	605	3952	599	1819	4826		UMI	608	2532	1296	171	4767
	HI-nonOECD	193	3782	787	1419	4778		HI-nonOECD	194	2302	1279	193	4622
	HI-OECD	571	4454	290	2435	4868		HI-OECD	571	4035	714	377	4858
$RelN_imp_{it}$	LI	173	0.64	0.14	0.26	0.86	$RelN_exp_{it}$	LI	168	0.22	0.16	0.04	0.71
	LMI	325	0.72	0.12	0.30	0.95		LMI	326	0.35	0.23	0.04	0.95
	UMI	605	0.81	0.12	0.37	0.98		UMI	608	0.52	0.27	0.03	0.97
	HI-nonOECD	193	0.77	0.16	0.29	0.97		HI-nonOECD	194	0.47	0.26	0.04	0.94
	HI-OECD	571	0.91	0.06	0.50	0.99		HI-OECD	571	0.83	0.15	0.08	0.99
$RelT_imp_{it}$	LI	167	2.08	0.52	1.18	3.31	$RelT_exp_{it}$	LI	172	4.58	0.83	2.30	6.22
	LMI	330	1.65	0.44	0.95	3.38		LMI	327	3.57	1.06	1.42	6.13
	UMI	606	1.17	0.45	0.46	3.19		UMI	602	2.72	1.14	0.86	6.24
	HI-nonOECD	193	1.40	0.59	0.57	2.83		HI-nonOECD	195	2.49	0.92	1.11	6.06
	HI-OECD	571	0.58	0.25	0.23	1.94		HI-OECD	571	1.47	0.81	0.44	4.64
$RelGini_imp_{it}$	LI	173	0.83	0.05	0.69	0.92	$RelGini_exp_{it}$	LI	167	0.98	0.02	0.93	1.00
	LMI	325	0.77	0.06	0.62	0.92		LMI	331	0.95	0.05	0.75	1.00
	UMI	607	0.68	0.08	0.48	0.92		UMI	607	0.91	0.08	0.64	1.00
	HI-nonOECD	191	0.73	0.10	0.50	0.92		HI-nonOECD	191	0.92	0.07	0.73	1.00
	HI-OECD	571	0.52	0.09	0.33	0.77		HI-OECD	571	0.76	0.12	0.49	0.99
DI_imp_{it}	LI	175	1.41	0.12	1.09	1.63	DI_exp_{it}	LI	165	1.85	0.08	1.56	1.95
	LMI	324	1.28	0.13	0.99	1.60		LMI	328	1.75	0.15	1.28	1.94
	UMI	607	1.09	0.18	0.71	1.58		UMI	608	1.59	0.22	1.02	1.94
	HI-nonOECD	190	1.17	0.20	0.75	1.63		HI-nonOECD	195	1.59	0.21	1.18	1.92
	HI-OECD	571	0.76	0.15	0.49	1.38		HI-OECD	571	1.22	0.26	0.73	1.92

Note: country groups according to the World Bank (2011) classification: LI- low income, LMI - lower middle income, UMI - upper middle income, HI- high income (OECD and non-OECD members). Sample without outliers defined as observations below 1st or above 99th percentile.

Source: own elaboration

Table 4. Observations potentially driving upward part of U-shaped lowess curve of relative diversification

Country name	Income group	No of obs.	GDPpc [PPP, const.2005 int.USD]*	Share of fuel exports [%]*	Population [1000]*
Ireland	HI-OECD	1	41025	0.8	4420
Hong Kong SAR	HI-nonOECD	2	41147	4.4	7055
United States	HI-OECD	8	42415	4.1	300040
Norway	HI-OECD	14	45460	60.9	4556
Macao SAR	HI-nonOECD	2	45901	0.0	512
Singapore	HI-nonOECD	8	46574	12.9	4509
United Arab Emirates	HI-nonOECD	4	48322	21.1	2566
Kuwait	HI-nonOECD	1	49411	96.5	2317
Brunei Darussalam	HI-nonOECD	4	50294	96.6	293
<i>For comparison:</i>					
Rest of the sample**		1784	13100	12.6	48934

Note: *average values across country-year observations, identified as potentially driving Upward part of U-curve (marked with triangles in Figure 2 and being above GDPpc of 40,000 [PPP, const.2005 int.USD]). **outliers excluded. Income groups according to the World Bank (2011) classification. Source: own elaboration based on trade data obtained from UNComtrade (2011); GDP per capita, share of fuel exports (as % of all merchandise exports) and population data obtained from WB WDI (2011).

Table 5. Regression estimation results – relative diversification and income per capita

<i>Dependent variable: inverse index of relative imports diversification</i>						
	(1)	(1)	(1)	(1)	(1)	(1)
	$\ln_RelT_imp_{it}$	$\ln_RelT_imp_{it}$	$\ln_RelT_imp_{it}$	$\ln_RelT_imp_{it}$	$\ln_RelT_imp_{it}$	$\ln_RelT_imp_{it}$
	FE	FE	FE	FE	FE	FE
$\ln GDP_{pc}$	-0.238***	-0.238***	-0.238***	-0.238***	-0.238***	-0.238***
[PPP, const.2005 int.USD]	[-4.78]	[-4.78]	[-4.78]	[-4.78]	[-4.78]	[-4.78]
No of observations	1820	1820	1820	1820	1820	1820
adj R2	0.47	0.47	0.47	0.47	0.47	0.47
year effect	no	no	no	no	no	no
country effect	yes	yes	yes	yes	yes	yes
No. of countries	149	149	149	149	149	149
years	1988-2010	1988-2010	1988-2010	1988-2010	1988-2010	1988-2010
<i>Dependent variable: inverse index of relative exports diversification</i>						
	(1)	(2)	(3)	(4)	(5)	(6)
	$\ln_RelT_exp_{it}$	$\ln_RelG_exp_{it}$	$\ln_DI_exp_{it}$	$\ln_RelT_exp_{it}$	$\ln_RelG_exp_{it}$	$\ln_DI_exp_{it}$
	FE	FE	FE	FE	FE	FE
$\ln GDP_{pc}$	-0.236***	-0.047***	-0.119***	-0.152**	-0.029*	-0.061***
[PPP, const.2005 int.USD]	[-5.56]	[-3.75]	[-6.66]	[-2.57]	[-1.80]	[-2.82]
No of observations	1822	1825	1827	1822	1825	1827
adj R2	0.41	0.3	0.38	0.39	0.25	0.34
year effect	no	no	no	yes	yes	yes
country effect	yes	yes	yes	yes	yes	yes
No. of countries	150	151	151	150	151	151
years	1988-2010	1988-2010	1988-2010	1988-2010	1988-2010	1988-2010

Note. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. Sample without outliers defined as observations below 1st or above 99th percentile. In FE estimates only countries with multiple observations have been considered. Robust t-statistics in parenthesis under coefficients. Constant included-not reported.

Source: own elaboration based on trade data obtained from UNComtrade, 2011 (HS0, 4963 product lines); GDP per capita – data obtained from WB WDI (2011).

Table 6. Correlation coefficients between import-export similarity and dissimilarity indices vs. relative trade diversification (Theil) and income per capita

		<i>Import-export similarity</i>			<i>Import-export dissimilarity</i>	
		RR	J	SM	MK	B
<i>Inverse measures of relative product diversification</i>	$RelT_exp$	-0.87	-0.87	-0.78	0.78	0.81
	$RelT_imp$	-0.84	-0.82	-0.65	0.70	0.74
<i>Development level</i>	GDP_{pc}	0.59	0.58	0.58	-0.69	-0.68

Note: RR- Russel-Rao index, J-Jaccard index, SM-Sokal and Michener index, MK- Michaely-Krugman Index, B - Battarcharyya Index, $RelT_exp$ ($RelT_imp$) - Theil index of exported (imported) products' relative specialization; 1866 obs. (sample without outliers defined as observations below 1st or above 99th percentile of GDP_{pc})

Source: own elaboration based on trade data from UNComtrade (HS0, 4963 product lines); GDP per capita from WB WDI.

ADDITIONAL MATERIAL

Appendix A

Table A1. List of countries in the dataset

Country code	Country Name	Income group*	GDP per capita [PPP, const.2005 int.USD]**	Years covered by the analysis	
				First	Last
BDI	Burundi	LI	403.02	1993	2010
BEN	Benin	LI	1321.28	1999	2006
BFA	BurkinaFaso	LI	967.256	1995	2010
BGD	Bangladesh	LI	929.165	1989	2006
CAF	CentralAfricanRepublic	LI	738.693	1993	2006
COM	Comoros	LI	1110.42	1995	1995
ETH	Ethiopia	LI	715.259	1995	2010
GIN	Guinea	LI	899.902	1995	2006
GMB	Gambia,The	LI	1112.19	1995	2001
KEN	Kenya	LI	1342.44	1992	2010
KGZ	KyrgyzRepublic	LI	1665.18	1995	2010
KHM	Cambodia	LI	1712.71	2004	2010
MDG	Madagascar	LI	892.184	1990	2010
MLI	Mali	LI	906.449	2004	2010
MOZ	Mozambique	LI	625.294	2001	2006
MWI	Malawi	LI	656.596	1990	2010
NER	Niger	LI	622.345	1995	2010
NPL	Nepal	LI	942.249	1994	2010
RWA	Rwanda	LI	911.698	2005	2009
TCD	Chad	LI	775.475	1995	1995
TGO	Togo	LI	834.185	1994	2005
TZA	Tanzania	LI	994.958	1995	2010
UGA	Uganda	LI	879.5	1994	2010
ZWE	Zimbabwe	LI	169.291	1995	2010
ARM	Armenia	LMI	3749.78	1997	2010
BLZ	Belize	LMI	5380.63	1992	2010
BOL	Bolivia	LMI	3668.07	1992	2010
BTN	Bhutan	LMI	3501.01	1993	2010
CIV	Coted'Ivoire	LMI	1691.27	1995	2009
CMR	Cameroon	LMI	1877.97	1995	2006
COG	Congo,Rep.	LMI	3243.59	1993	1995
CPV	CapeVerde	LMI	2637.41	1997	2010
DJI	Djibouti	LMI	2087.17	2009	2009
EGY	Egypt,ArabRep.	LMI	4658.63	1994	2010
FJI	Fiji	LMI	4220.25	2002	2009
GEO	Georgia	LMI	3425.32	1999	2010
GHA	Ghana	LMI	1310.47	2005	2010
GTM	Guatemala	LMI	3984.07	1993	2010
GUY	Guyana	LMI	2447.61	1991	2010
HND	Honduras	LMI	3092.84	1994	2009
IDN	Indonesia	LMI	2475.73	1989	2001
IND	India	LMI	1785.88	1988	2010
KIR	Kiribati	LMI	2119.82	1995	2005
LKA	SriLanka	LMI	3101.67	1990	2010
MAR	Morocco	LMI	3145.68	1993	2006
MDA	Moldova	LMI	2173.6	1994	2010

MNG	Mongolia	LMI	2884.68	2003	2007
NGA	Nigeria	LMI	1725.66	1999	2010
NIC	Nicaragua	LMI	2149.23	1993	2010
PAK	Pakistan	LMI	2203.22	2003	2010
PHL	Philippines	LMI	2707.4	2000	2001
PNG	PapuaNewGuinea	LMI	1858.78	2001	2004
PRY	Paraguay	LMI	4090.58	1989	2010
SDN	Sudan	LMI	1532.39	1995	2009
SEN	Senegal	LMI	1584.04	1996	2010
SLV	ElSalvador	LMI	5425.54	1994	2010
STP	SaoTomeandPrincipe	LMI	1018.62	1999	2000
SWZ	Swaziland	LMI	4091.36	2000	2001
SYR	SyrianArabRepublic	LMI	4199.88	2001	2008
TKM	Turkmenistan	LMI	2151.75	1999	2000
TMP	Timor-Leste	LMI	692.021	2004	2005
TON	Tonga	LMI	4150.35	2008	2010
UKR	Ukraine	LMI	4077.3	2001	2001
VNM	Vietnam	LMI	2157.9	2000	2009
WSM	Samoa	LMI	3601.49	2001	2010
YEM	Yemen,Rep.	LMI	2251.41	2006	2009
ZMB	Zambia	LMI	1148.38	1995	2010
ALB	Albania	UMI	5413.32	1996	2010
ARG	Argentina	UMI	10838.6	1993	2010
ATG	AntiguaandBarbuda	UMI	13932.7	1999	1999
AZE	Azerbaijan	UMI	5446.31	1999	2010
BGR	Bulgaria	UMI	8857.79	1996	2010
BIH	BosniaandHerzegovina	UMI	6638.91	2003	2010
BLR	Belarus	UMI	8244.79	1998	2010
BRA	Brazil	UMI	8065.28	1989	2010
BWA	Botswana	UMI	9624.23	2000	2001
CHL	Chile	UMI	10505.7	1990	2010
CHN	China	UMI	3359.38	1992	2010
COL	Colombia	UMI	7052.69	1991	2010
CRI	CostaRica	UMI	8610.85	1994	2010
CUB	Cuba	UMI	9370.14	2001	2006
DMA	Dominica	UMI	7278.84	1993	2010
DOM	DominicanRepublic	UMI	7367.72	2001	2010
DZA	Algeria	UMI	6442.37	1992	2010
ECU	Ecuador	UMI	6065.84	1991	2010
GAB	Gabon	UMI	13917.1	1993	2006
GRD	Grenada	UMI	5841	1993	2001
IRN	Iran,IslamicRep.	UMI	8140.25	1997	2006
JAM	Jamaica	UMI	6777.94	1991	2010
JOR	Jordan	UMI	4141.21	1994	2010
KAZ	Kazakhstan	UMI	7563.11	1998	2009
KNA	St.KittsandNevis	UMI	11084.1	1993	2001
LBN	Lebanon	UMI	9868.63	1997	2010
LCA	St.Lucia	UMI	7823.73	1992	2001
LTU	Lithuania	UMI	12000.1	1994	2010
LVA	Latvia	UMI	10659.3	1994	2010
MDV	Maldives	UMI	3729.26	1995	2008
MEX	Mexico	UMI	11286.3	1990	2010
MKD	Macedonia,FYR	UMI	7185.25	1994	2008
MUS	Mauritius	UMI	9374.96	1993	2010

MYS	Malaysia	UMI	9673.83	1989	2010
NAM	Namibia	UMI	5053.94	2000	2008
PAN	Panama	UMI	8685.63	1995	2008
PER	Peru	UMI	6126.72	1992	2010
ROM	Romania	UMI	8257.19	1989	2010
RUS	RussianFederation	UMI	10663.5	1997	2009
SUR	Suriname	UMI	4876.47	1994	2010
SYC	Seychelles	UMI	16674.2	1994	2002
THA	Thailand	UMI	5655.72	1988	2010
TUN	Tunisia	UMI	5753.52	1991	2010
TUR	Turkey	UMI	9886.54	1989	2010
URY	Uruguay	UMI	9731.02	1994	2009
VCT	St.VincentandtheGrenadines	UMI	6019.17	1993	2010
VEN	Venezuela,RB	UMI	10001.5	1994	2006
ZAF	SouthAfrica	UMI	8177.64	1992	2010
ARE	UnitedArabEmirates	HI-nonOECD	46482.5	1991	2008
BHR	Bahrain	HI-nonOECD	26071.2	2002	2010
BHS	Bahamas,The	HI-nonOECD	24975	1997	2009
BMU	Bermuda	HI-nonOECD	35062.4	1995	1995
BRB	Barbados	HI-nonOECD	17973.7	2000	2006
BRN	BruneiDarussalam	HI-nonOECD	50294.4	1992	2004
CYP	Cyprus	HI-nonOECD	21986.6	1989	2010
HKG	HongKongSAR,China	HI-nonOECD	32559	1993	2010
HRV	Croatia	HI-nonOECD	13256.2	1992	2010
KWT	Kuwait	HI-nonOECD	49411.3	2006	2006
MAC	MacaoSAR,China	HI-nonOECD	32125.3	1991	2010
MLT	Malta	HI-nonOECD	19966.7	1994	2010
OMN	Oman	HI-nonOECD	18298.1	1989	2007
QAT	Qatar	HI-nonOECD	67156.3	2000	2006
SAU	SaudiArabia	HI-nonOECD	20044.4	1991	2010
SGP	Singapore	HI-nonOECD	37397.2	1989	2010
TTO	TrinidadandTobago	HI-nonOECD	16309	1991	2009
AUS	Australia	HI-OECD	28625.7	1988	2010
AUT	Austria	HI-OECD	31864.6	1994	2010
BEL	Belgium	HI-OECD	31735.3	1999	2010
CAN	Canada	HI-OECD	31088.3	1989	2010
CHE	Switzerland	HI-OECD	34228.6	1988	2010
CZE	CzechRepublic	HI-OECD	18435.4	1993	2010
DEU	Germany	HI-OECD	29069.6	1988	2010
DNK	Denmark	HI-OECD	30139.5	1989	2010
ESP	Spain	HI-OECD	24020.7	1989	2010
EST	Estonia	HI-OECD	13504.8	1995	2010
FIN	Finland	HI-OECD	26349.9	1988	2010
FRA	France	HI-OECD	28154.1	1994	2010
GBR	UnitedKingdom	HI-OECD	29747.5	1993	2010
GRC	Greece	HI-OECD	21101.5	1988	2010
HUN	Hungary	HI-OECD	13499.7	1992	2007
IRL	Ireland	HI-OECD	30952.2	1992	2010
ISL	Iceland	HI-OECD	29545.1	1988	2010
ISR	Israel	HI-OECD	23044	1995	2010
ITA	Italy	HI-OECD	27161.6	1994	2010
JPN	Japan	HI-OECD	28353.2	1988	2010
KOR	Korea,Rep.	HI-OECD	17867.3	1988	2009
LUX	Luxembourg	HI-OECD	66777.1	1999	2010

NLD	Netherlands	HI-OECD	31695.1	1992	2006
NOR	Norway	HI-OECD	43558.9	1993	2010
NZL	NewZealand	HI-OECD	22381.9	1989	2010
POL	Poland	HI-OECD	12676.2	1994	2010
PRT	Portugal	HI-OECD	19329.4	1988	2010
SVK	SlovakRepublic	HI-OECD	14937.6	1994	2010
SVN	Slovenia	HI-OECD	21107.5	1994	2010
SWE	Sweden	HI-OECD	29122.1	1992	2010
USA	UnitedStates	HI-OECD	38398.1	1991	2010

Note:

Countries ordered alphabetically within income groups

*country groups according to the World Bank's (2011) classification: LI- low income, LMI - lower middle income, UMI - upper middle income, HI- high income (OECD and non-OECD members)

**year average

Table A2. Composition of country-year data-set matching trade diversification indices with data on income per capita

		Dataset based on HS0 trade data (4963 product lines)					
		All countries	Countries divided by income group*				
		LI	LMI	UMI	HI-nonOECD	HI-OECD	
Total number of country-year obs.		1905**	178	333	609	195	590
Time span		1988-2010	1989-2010	1988-2010	1988-2010	1989-2010	1988-2010
Number of countries		163	24	43	48	17	31
Number of year obs. per country	Mean	11	7	8	12	19	11
	Min	1	1	1	1	1	12
	Max	23	20	19	22	22	23
Number of country obs. per year	Mean	82	8	14	26	8	25
	Min	11	1	1	1	9	3
	Max	130	18	27	44	31	14

Note:*country groups according to the World Bank (2011) classification: LI- low income, LMI - lower middle income, UMI - upper middle income, HI- high income (OECD and non-OECD members)

**less when considering the number of new lines (due to 3-year or 5-year moving window- see explanation in text)

Source: own elaboration

Appendix B - robustness checks

Table B1. Correlation coefficients between benchmark diversification indices and indices calculated with exclusion of oil/petrol-related products

	$RelT_imp_{it}$	$RelG_imp_{it}$	DI_imp_{it}	$RelT_imp_{it_noPetrol}$	$RelG_imp_{it_noPetrol}$	$DI_imp_{it_noPetrol}$
$RelT_imp_{it}$	1					
$RelG_imp_{it}$	0.9059	1				
DI_imp_{it}	0.9427	0.9791	1			
$RelT_imp_{it_noPetrol}$	0.9923	0.9039	0.9343	1		
$RelG_imp_{it_noPetrol}$	0.9083	0.9927	0.9717	0.9143	1	
$DI_imp_{it_noPetrol}$	0.94	0.9747	0.9909	0.9448	0.9811	1

	$RelT_exp_{it}$	$RelG_exp_{it}$	DI_exp_{it}	$RelT_exp_{it_noPetrol}$	$RelG_exp_{it_noPetrol}$	$DI_exp_{it_noPetrol}$
$RelT_exp_{it}$	1					
$RelG_exp_{it}$	0.7973	1				
DI_exp_{it}	0.89	0.9595	1			
$RelT_exp_{it_noPetrol}$	0.9563	0.8316	0.9059	1		
$RelG_exp_{it_noPetrol}$	0.7999	0.9952	0.9537	0.8334	1	
$DI_exp_{it_noPetrol}$	0.8775	0.9631	0.9915	0.9118	0.9655	1

Note: Indices with suffix *_noPetrol* are calculated excluding from the product sample the following HS0 product lines: 27900, 271011, 271119, 271129, 271210, 271311, 271312, 271320, 271390.

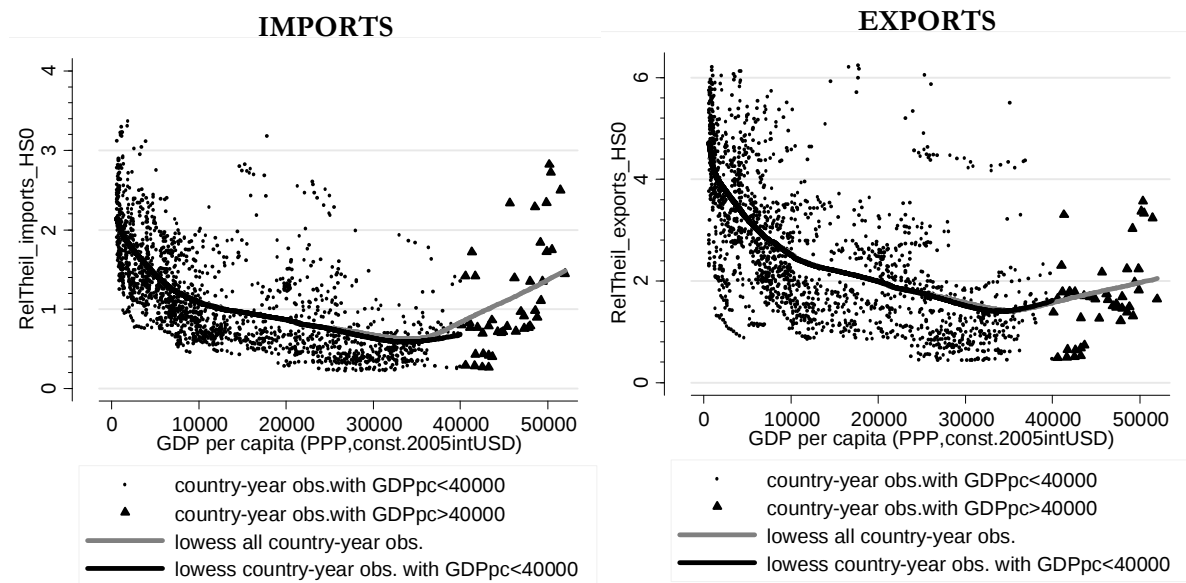


Figure B1. Robustness check – lowess, alternative value of span parameter

Source: own elaboration based on trade data obtained from UNComtrade, 2011 (HS0, 4963 product lines) and GDP per capita obtained from WB WDI (2011).

Note: lowess - span=0.5, sample without outliers defined as observations below 1st or above 99th percentile (number of obs.=1828)

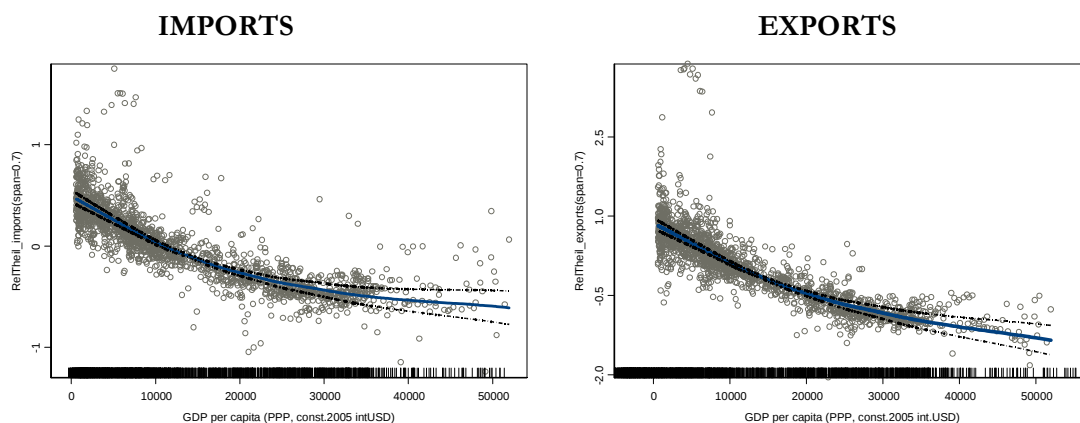


Figure B2. Robustness check – semiparametric curves, alternative value of span parameter

Source: own elaboration based on trade data from UNComtrade, 2011 (HS0, 4963 product lines) and GDP per capita from WB WDI (2011).

Note: yaxis: inverse diversification measure (Relative Theil index)- partial residuals of GAM model (eq. 5), span=0.7; sample without outliers defined as observations below 1st or above 99th percentile (number of obs.=1828). Due to FE inclusion only countries with multiple observations have been considered..

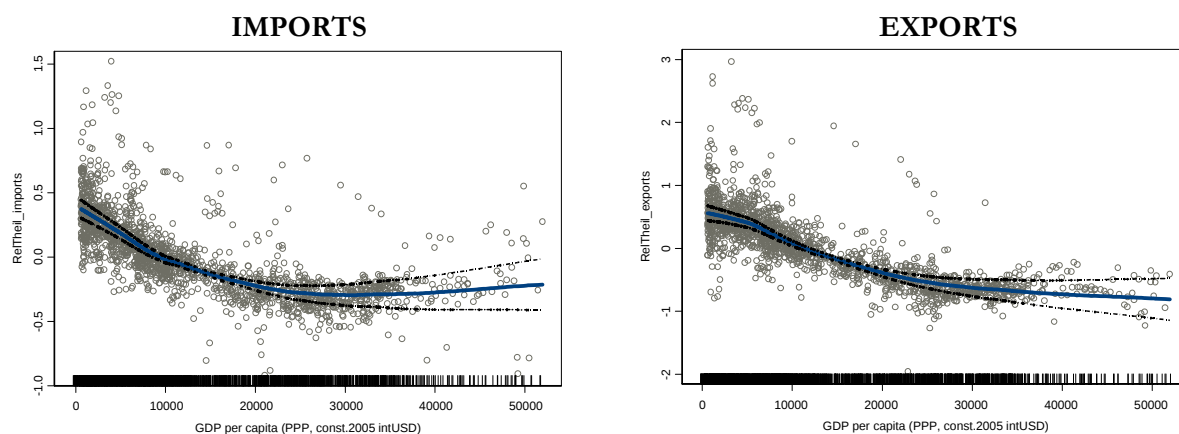


Figure B3. Robustness check – semiparametric curves controlling for country fixed effects and time fixed effects

Source: own elaboration based on trade data from UNComtrade, 2011 (HS0, 4963 product lines) and GDP per capita from WB WDI (2011).

Note: yaxis: inverse diversification measure (Relative Theil index)- partial residuals of GAM model (eq. 7), span=0.5; sample without outliers defined as observations below 1st or above 99th percentile (number of obs.=1828). Due to FE inclusion only countries with multiple observations have been considered.

Table B2. Robustness check: regression estimation results – relative *imports* diversification versus income per capita and additional country specific factors

<i>Dependent variable: inverse index of exports diversification (ln_RelTheil_imp)</i>							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Country sample#	All	No microstates	No Petrol abundant	All+controls	All+controls	All+controls	All+controls
<i>LnGDPpc</i>	-0.358***	-0.390***	-0.372***	-0.198***	-0.393***	-0.213***	-0.411***
[PPP, const.2005 int.USD]	[0.009]	[0.008]	[0.009]	[0.009]	[0.007]	[0.011]	[0.008]
<i>LnGDP</i>				-0.152***		-0.116***	
[PPP, const.2005 int.USD]				[0.004]		[0.005]	
<i>LnPOP</i>					-0.149***		-0.161***
[1000]					[0.004]		[0.005]
<i>LnDISTANCE</i>						0.093***	
[km]						[0.011]	
<i>LnPETROL</i>							0.007***
[%total exports]							[0.001]
No of observations	1828	1627	1706	1821	1825	1613	1825
adj R2	0.47	0.6	0.52	0.71	0.69	0.73	0.70
year dummy	yes	yes	yes	yes	yes	yes	yes
No. of countries	157	133	146	155	156	128	156
years	1988-2010	1988-2010	1988-2010	1988-2010	1988-2010	1988-2010	1988-2010

Note. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. OLS log-log estimates, #sample without outliers defined as observations below 1st or above 99th percentile. Model (2) – without countries with population below 1mln (microstates). Model (3) – without countries with fuel exports above 50% of total exports. Robust standard errors in parenthesis under coefficients.

Source: own elaboration based on trade data from UNComtrade, 2011 (HS0, 4963 product lines); GDP per capita, GDP, population and fuel exports from WB WDI (2011), Distance (in km from the closest major market: New York, Rotterdam or Tokyo) from Gallup et al.(1999).

Table B3. Robustness check: regression estimation results – relative *exports* diversification versus income per capita and additional country specific factors

<i>Dependent variable: inverse index of exports diversification (ln_RelTheil_exp)</i>							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Country sample#	All	No microstates	No Petrol abundant	All+controls	All+controls	All+controls	All+controls
<i>LnGDPpc</i>	-0.338***	-0.371***	-0.357***	-0.141***	-0.389***	-0.093***	-0.396***
[PPP, const.2005 int.USD]	[0.009]	[0.009]	[0.010]	[0.007]	[0.006]	[0.007]	[0.007]
<i>LnGDP</i>				-0.192***		-0.159***	
[PPP, const.2005 int.USD]				[0.004]		[0.004]	
<i>LnPOP</i>					-0.194***		-0.198***
[1000]					[0.004]		[0.004]
<i>LnDISTANCE</i>						0.203***	
[km]						[0.008]	
<i>LnPETROL</i>							0.003**
[%total exports]							[0.001]
No of observations	1828	1632	1705	1820	1825	1618	1825
adj R2	0.41	0.53	0.43	0.76	0.76	0.83	0.76
year dummy	yes	yes	yes	yes	yes	yes	yes
No. of countries	155	132	145	155	155	127	155
years	1988-2010	1988-2010	1988-2010	1988-2010	1988-2010	1988-2010	1988-2010

Note. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. OLS log-log estimates, #sample without outliers defined as observations below 1st or above 99th percentile. Model (2) – without countries with population below 1mln (microstates). Model (3) – without countries with petrol exports above 50% of total exports. Robust standard errors in parenthesis under coefficients.

Source: own elaboration based on trade data from UNComtrade, 2011 (HS0, 4963 product lines); GDP per capita, GDP, population from WB WDI (2011), Distance (in km from the closest major market: New York, Rotterdam or Tokyo) from Gallup et al.(1999).

Table B4. Robustness check – regressions with modified dependent variable: indices calculated excluding petrol related products

	<i>Dependent variable: inverse index of relative imports diversification</i>					
	(1) $\ln_RelIT_imp_{it}^{noPetrol}$ FE	(2) $\ln_RelG_imp_{it}^{noPetrol}$ FE	(3) $\ln_DI_imp_{it}^{noPetrol}$ FE	(4) $\ln_RelIT_imp_{it}^{noPetrol}$ FE	(5) $\ln_RelG_imp_{it}^{noPetrol}$ FE	(6) $\ln_DI_imp_{it}^{noPetrol}$ FE
$\ln GDP_{pc}$	-0.229***	-0.023	-0.122***	-0.361***	-0.092***	-0.162***
[PPP, const.2005 int.USD]	[-4.40]	[-1.12]	[-5.89]	[-4.36]	[-2.81]	[-4.73]
No of observations	1828	1828	1828	1828	1828	1828
adj R2	0.49	0.46	0.52	0.49	0.47	0.52
year effect	no	no	no	yes	yes	yes
country effect	yes	yes	yes	yes	yes	yes
No. of countries	157	156	156	157	156	156
years	1988-2010	1988-2010	1988-2010	1988-2010	1988-2010	1988-2010
	<i>Dependent variable: inverse index of relative exports diversification</i>					
	(1) $\ln_RelIT_exp_{it}^{noPetrol}$ FE	(2) $\ln_RelG_exp_{it}^{noPetrol}$ FE	(3) $\ln_DI_exp_{it}^{noPetrol}$ FE	(4) $\ln_RelIT_exp_{it}^{noPetrol}$ FE	(5) $\ln_RelG_exp_{it}^{noPetrol}$ FE	(6) $\ln_DI_exp_{it}^{noPetrol}$ FE
$\ln GDP_{pc}$	-0.260***	-0.052***	-0.123***	-0.212***	-0.042**	-0.083***
[PPP, const.2005 int.USD]	[-6.02]	[-3.96]	[-6.48]	[-3.32]	[-2.27]	[-3.56]
No of observations	1828	1831	1833	1828	1831	1833
adj R2	0.41	0.3	0.38	0.39	0.29	0.37
year effect	no	no	no	yes	yes	yes
country effect	yes	yes	yes	yes	yes	yes
No. of countries	157	157	157	157	157	157
years	1988-2010	1988-2010	1988-2010	1988-2010	1988-2010	1988-2010

Note. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. Outliers (defined as observations below 1st or above 99th percentile) excluded. Robust t-statistics in parenthesis under coefficients. Constant included-not reported.

Source: own elaboration based on trade data from UNComtrade, 2011 (HS0, 4954 product lines); GDP per capita from WB WDI (2011).

Table B5. Robustness check – regressions with the whole set of available observations

<i>Dependent variable: inverse index of relative imports diversification</i>						
	(1)	(2)	(3)	(4)	(5)	(6)
	$\ln_RelIT_imp_{it}$	$\ln_RelG_imp_{it}$	$\ln_DI_imp_{it}$	$\ln_RelIT_imp_{it}$	$\ln_RelG_imp_{it}$	$\ln_DI_imp_{it}$
	FE	FE	FE	FE	FE	FE
$\ln GDP_{pc}$	-0.237***	-0.02	-0.119***	-0.393***	-0.082***	-0.157***
[PPP, const.2005 int.USD]	[-4.92]	[-0.97]	[-6.09]	[-5.33]	[-2.74]	[-5.26]
No of observations	1912	1912	1912	1912	1912	1912
adj R2	0.49	0.45	0.5	0.5	0.45	0.5
year effect	no	no	no	yes	yes	yes
country effect	yes	yes	yes	yes	yes	yes
No. of countries	160	158	158	160	158	158
years	1988-2010	1988-2010	1988-2010	1988-2010	1988-2010	1988-2010
<i>Dependent variable: inverse index of relative exports diversification</i>						
	(1)	(2)	(3)	(4)	(5)	(6)
	$\ln_RelIT_exp_{it}$	$\ln_RelG_exp_{it}$	$\ln_DI_exp_{it}$	$\ln_RelIT_exp_{it}$	$\ln_RelG_exp_{it}$	$\ln_DI_exp_{it}$
	FE	FE	FE	FE	FE	FE
$\ln GDP_{pc}$	-0.293***	-0.046***	-0.117***	-0.122***	-0.020**	-0.039***
[PPP, const.2005 int.USD]	[-8.78]	[-6.04]	[-9.84]	[-2.83]	[-2.46]	[-3.17]
No of observations	3484	3485	3487	3484	3485	3487
adj R2	0.38	0.27	0.35	0.36	0.24	0.31
year effect	no	no	no	yes	yes	yes
country effect	yes	yes	yes	yes	yes	yes
No. of countries	162	162	162	162	162	162
years	1988-2010	1988-2010	1988-2010	1988-2010	1988-2010	1988-2010

Note. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. Outliers (defined as observations below 1st or above 99th percentile) excluded. Robust t-statistics in parenthesis under coefficients. Constant included-not reported.

Source: own elaboration based on trade data from UNComtrade, 2011 (HS0, 4963 product lines); GDP per capita from WB WDI (2011).

Table B6. Robustness check – quadratic model

	<i>Dependent variable:</i>			
	<i>inverse index of export diversification</i>	<i>inverse index of import diversification</i>		
	(1)	(2)	(3)	(4)
	$RelIT_exp_{it}$	$RelIT_exp_{it}$	$RelIT_imp_{it}$	$RelIT_imp_{it}$
	OLS	FE	OLS	FE
(GDPpc) ²	3.154e-09***	2.064e-09***	1.754e-09***	1.280e-09***
	[23.76]	[6.10]	[18.22]	[5.21]
GDPpc	-1.914e-04***	-1.459e-04***	-9.774e-05***	-7.344e-05***
[PPP, const.2005 int.USD]	[-33.94]	[-7.80]	[-26.29]	[-6.06]
No of observations	3483	3484	1912	1912
adj R2	0.33	0.34	0.46	0.46
country effect	no	yes	no	yes
No. of countries/groups	162	162	160	160
years	1988-2010	1988-2010	1988-2010	1988-2010
Estimated turning point				
[PPP, const.2005 int.USD]	30335.9	35335.9	27855.6	28678.6
No. of obs. below turning point	3237	3398	1643	1681
% of obs. below turning point	93%	98%	86%	88%
No. of countries below turning point	155	160	153	153

Note. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. Sample without outliers defined as observations below 1st or above 99th percentile. In FE estimates only countries with multiple observations retained. Robust t-statistics in parenthesis under coefficients.

Source: own elaboration based on trade data from UNCOMTRADE, 2011 (HS0, 4963 product lines); GDP per capita from WB WDI (2011).

Table B7. Robustness check – IV regressions accounting for potential endogeneity of GDP per capita

	<i>Dependent variable: inverse index of relative imports diversification</i>			<i>Dependent variable: inverse index of relative exports diversification</i>		
	(1) <i>ln_RelIT_imp_{it}</i> IV FE	(2) <i>ln_RelG_imp_{it}</i> IV FE	(3) <i>ln_DI_imp_{it}</i> IV FE	(4) <i>ln_RelIT_exp_{it}</i> IV FE	(5) <i>ln_RelG_exp_{it}</i> IV FE	(6) <i>ln_DI_exp_{it}</i> IV FE
<i>lnGDPpc</i>	-0.149***	0.001	-0.097***	-0.183***	-0.038***	-0.106***
[PPP, const.2005 int.USD]	[-6.566]	[0.160]	[-10.206]	[-10.959]	[-8.510]	[-16.113]
No of observations	1560	1560	1560	1560	1560	1560
adj R2	0.42	0.37	0.43	0.39	0.27	0.35
p_FE	0.000	0.000	0.000	0.000	0.000	0.000
F_exo	58.75	40.51	45.89	2.3	0.45	0.11
p_exo	0.000	0.000	0.000	0.13	0.504	0.736
years	1988-2010	1988-2010	1988-2010	1988-2010	1988-2010	1988-2010

Note. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. IV FE – Instrumental Variables with fixed effects – *lnGDPpc* instrumented by its lag; t-statistics in parenthesis under coefficients; p_FE- refer to test of joint significance of country specific effects (FE); F_exo, p_exo – refer to Davidson-MacKinnon test of exogeneity (IV). Outliers (defined as observations below 1st or above 99th percentile) excluded. Constant included-not reported. Source: own elaboration based on trade data from UNComtrade, 2011 (HS0, 4963 product lines); GDP per capita from WB WDI (2011).

Appendix C - Data Appendix

Trade data

The Harmonized System is an international nomenclature for classification of products, which allows participating countries to classify traded goods on a common basis for customs purposes. For the purpose of international comparisons, HS is a six-digit code system. Currently the HS comprises approximately 5,300 article/product descriptions, which appear as headings and subheadings, arranged in 99 chapters and grouped in 21 sections. The six digits can be divided into three parts. The first two digits identify the chapter the goods are classified in, the next two digits identify groupings within that chapter and the last two digits are even more specific. Up to the HS six digit level, all countries classify products in the same way. Deeper disaggregation series are not harmonised and thus are inappropriate for cross-country studies.

Our analysis covers the period from the first year when HS was launched (1988) until 2010. Within that period there were three major revisions following the first release, corresponding to the following data series: HS0 – 1988/92 (5016 products), HS1 – 1996 (5113 products), HS2 – 2002 (5224 products), HS3 – 2007 (5053 products). We combine data from these revisions, taking for each year the most accurate one (i.e. matching HS0 data for the years 1988-1995 with HS1 data for the years 1996-2001, and so on). In order to obtain a consistent long time data series we use automatic conversion into HS0 product division (based on conversion tables from WITS).

Rough trade data reported by COMTRADE do not contain the information on every product line for every country-year. This might be a problem if one considers potential mismatch between inactive trade lines (zero trade value) and the missing data on existing flows. To overcome this issue, following Cadot et al. (2011a) we fill in the database, adding missing product lines and assign them zero trade values (we thus assume that the unreported trade flow is zero flow). In the case of imports this shall not be a major weakness, as import statistics are believed to be much more complete and of a better quality than export data. However, we exclude from the sample 53 HS0 codes, which correspond to never-traded ('silent') product lines. Within the years 1988-2010 the imports of these 53 goods were never reported as positive neither by any of the countries, nor at the aggregate world level. In the end we are left with 4963 product lines and, with the use of these statistics we calculate product-wise diversification indices for each country and year.

The HS system is very detailed for some sectors (such as textiles) and less detailed for others (such as machinery). However, this is not necessarily a problem, because the share of every HS section in the total number of lines is largely correlated with its relative importance in world trade.

Our analysis covers 23 years during which undoubtedly new products have appeared on the market. Concordance tables between older and newer revisions take it into account in an indirect way – usually new products, which exist in newer revisions as separate codes are included as part of a more aggregate code in the older revision of trade data. For instance, the data on mobile phones, classified in HS2007 revision as 'Telephones for cellular networks/for other wireless networks, other than Line telephone sets with cordless handsets' (HS2007 code 851712) is included in HS1988 nomenclature as part of the product line 'Transmission apparatus for radio-telephony, radio-telegraphy, radio-broadcasting or television incorporating reception apparatus' (HS1988 code 852520). Consequently, the increasing importance of cellular phones in the world trade would be visible in the data classified according to older nomenclature (HS1988) as a rise in the value of product line 852520. It is indeed the case – between 1988 and 2011 its import value rose by 300 times.

Still, some problems may be caused by the fact that not every product present in the more recent HS revisions can be associated with basic HS0 1988/92 product code (this difficulty is present particularly in the case of the latest revision: HS3-2007). On country-year basis we have calculated the share of trade value (in total), which would be missed due to the imperfect matching between HS0 1988/92 and the newer revisions. The average non-counted imports share equals 4% of total country imports (s.d = 0.05, min=0, max=0.5). We exclude from the final sample these country-year pairs in which more than 15% of the total country trade value would not be accounted for by HS0 product classification (which corresponds to roughly 2% of all our country-year observations). Finally, we obtain a set of diversification indices and after matching them with the data on income per capita we are left with 1905 observations for 163 countries (unbalanced panel), with 11 average annual observations per country.

Additional data

GDP per capita data has been obtained from the World Bank WDI and is expressed in PPP terms (const. 2005 international \$). The data is not available for all country-year observations on trade which we have, thus in 15 observations we complement the WDI data on income per capita with those obtained from PWT (version 7.0, June 2011), also expressed in PPP(2005) terms.

Original citation:

Parteka, A. & Tamberi, M. (2013). Product diversification, relative specialisation and economic development: import-export analysis. GUT FME Working Paper Series A, No. 15/2013(15).Gdansk (Poland): Gdansk University of Technology, Faculty of Management and Economics.

All GUT Working Papers are downloadable at:

<http://www.zie.pg.gda.pl/web/english/working-papers>

GUT Working Papers are listed in Repec/Ideas

<http://ideas.repec.org/s/gdk/wpaper.html>



GUT FME Working Paper Series A jest objęty licencją Creative Commons Uznanie autorstwa-Użycie niekomercyjne-Bez utworów zależnych 3.0 Unported.



GUT FME Working Paper Series A is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License.

Gdańsk University of Technology, Faculty of Management and Economics

Narutowicza 11/12, (premises at ul. Traugutta 79)

80-233 Gdańsk, phone: 58 347-18-99 Fax 58 347-18-61

www.zie.pg.gda.pl

