UNIVERSITÀ POLITECNICA DELLE MARCHE

DIPARTIMENTO DI ECONOMIA

DETERMINANTS OF EXPORT DIVERSIFICATION:
AN EMPIRICAL INVESTIGATION

Aleksandra Parteka* and Massimo Tamberi**

QUADERNI DI RICERCA n.327

*Gdansk University of Technology, Faculty of Management and Economics, Gdansk, Poland; aparteka@zie.pg.gda.pl

**Università Politecnica delle Marche, Faculty of Economics, Ancona, Italy

UNIVERSITÀ POLITECNICA DELLE MARCHE

DIPARTIMENTO DI ECONOMIA

EXPORT DIVERSIFICATION AND DEVELOPMENT
- EMPIRICAL ASSESSMENT

Aleksandra Parteka and Massimo Tamberi

QUADERNI DI RICERCA N. 359

Giugno 2011

Scientific committee:

Balducci Renato

Gallegati Marco

Niccoli Alberto

Zazzaro Alberto

Editor:

Tamberi Massimo

Abstract

This paper assesses the role played by country specific factors as determinants of exports' diversification process. Using a panel data-set for 60 countries and twenty years (1985-2004) we confirm that even after clearing out differences in income per capita, cross section variability in the degree of exports' diversification is significant. In general, apart from per capita income, variables influencing the size of accessible markets (domestic and foreign) are the most relevant and robust determinants of the export diversification process. Diversification opportunities grow if countries are large, not located far from economic core areas and when barriers to trade are restricted.

Contact Details:

Aleksandra Parteka
Gdansk University of Technology, Faculty of Management and Economics, Gdansk, Poland;
aparteka@zie.pg.gda.pl

Massimo Tamberi
Università Politecnica delle Marche, Faculty of Economics, Ancona, Italy;
m.tamberi@univpm.it

1. INTRODUCTION

The theme of sectoral diversification along the path of development has recently been touched upon deeply in trade literature, especially empirical. The argument of how countries with different levels of per capita income manage to diversify their economic structures is important in terms of policy implications. On the one hand, the ongoing process of diversification confirms that structural change is an aspect of development. On the other hand, it is important because concentration of resources in few sectors (i.e., a high degree of overall specialisation[1]) may be risky in the case of idiosyncratic (sectoral) shocks (Kalemli-Ozcan et al. 2003) which can limit economic growth (De Ferranti et al., 2002).

On the empirical side, since there are no neat theoretical indications on the nature of GDP *per capita* - specialisation relationship and its determinants, most studies propose restricted empirical estimations of the 'specialisation curve' (showing a link between a measure of diversification and the level of development in a sample of countries).

A common result is the decrease in the degree of overall specialisation in the initial phase of economic development (thus increasing diversification). However, there is less agreement on the trend at higher stages of development. Imbs and Wacziarg (2003), Cadot et al. (2011), as well as Koren and Tenreyro (2007) found a U-shaped industrial specialisation pattern, with some signs of re-specialisation at higher levels of development. A non-linear but monotonically decreasing trend has been found by de Benedictis et al. (2009) and Parteka (2010), using relative measures of specialization[2]. A helpful summary of the debate is presented in Carrer et al. (2009). Moreover, while most of the works find a non-linear relationship between diversification and development level, there are also others directly testing a linear one (e.g. : Easterly et al., 2009).

Literature so far has several limits that challenge future work on this topic. Previous works on diversification along the development process are not directly comparable because they differ in several aspects. First of all, apart from analyzing different periods and different cross sections of countries, they follow diverse methodological approaches, consisting in the use of non-parametric, semi-parametric and full parametric estimations. Secondly, only some of them, such as de Benedictis et al. (2009) and Parteka (2010), make use of country fixed effects in flexible estimations (using GAM – General Additive Models) which prove to be

---

[1] From now on, we use the term 'degree of overall specialisation' as an opposite of 'diversification' so that low export specialisation means high export diversification and vice versa.

[2] Similar results can be found in Minondo's (2011) non parametric and parametric results.

crucial for the conclusions on the eventual existence of a U-shaped pattern of diversification. Thirdly, there are different specific indexes of specialisation (relative vs. absolute) and the U-shaped pattern tends to be 'found' with the latter ones without comparing the degree of specialisation of a given country to the rest of the country group.[3]

Most importantly, from the point of view of this paper and to the best of our knowledge, there is no study that presents in a systematic way an empirical assessment of overall specialisation (diversification) with respect to GDP *per capita* levels along with the examination of factors influencing the process of diversification. Usually, only *per capita* income and sometimes country specific fixed effects are the sole explanatory variables taken into consideration to explain the diversification process. Only a few recent contributions propose limited alternative sets of explanatory variables, sometimes in order to check specific, limited hypothesis (as 'product centrality' in Minondo, 2011, and 'trade facilitation' in Dennis and Shepherd, 2011).

Finally, authors usually (apart from an attempt by Starosta de Waldemar, 2010) do not take into consideration possible problems of endogeneity in the development-specialization relationship. While almost 'automatically' it has been assumed that specialization is a function of development, one could also consider the opposite causal relationship (when for example productivity levels depend on trade diversification, as in Hammouda et. al, 2010; also the literature related to the impact of exports on productivity should be considered in this perspective: Wagner , 2007).

Consequently, keeping in mind the findings of existing studies, we aim at filling, at least partially, some of the gaps in the literature. Consequently, we move along four main research lines. Firstly, starting from non-parametric and semi-parametric results (Imbs and Wacziarg, 2003; de Benedicts et al., 2009; Parteka, 2010) we will directly compare, in a parametric estimations context, different functional shapes, in order to find out which one better fits the data. Then, we will present the results with and without country fixed effects, determining what proportion of sectoral diversification depends on country specific characteristics (cross-section dimension). Thirdly, in order to approach the problem of diversification determinants in more detail, we will make use of a large set of explanatory

---

[3] See Parteka (2010) for a simultaneous comparison of the results obtained with the two groups of indices. She argues that in a semi-parametric framework with fixed effects (accounting for cross country heterogeneity), employment and export specialization patterns show very similar trends, and there is a clear tendency towards despecialization in the initial levels of economic growth, not sensitive to the way of specialization measurement.

variables to show which kind of factors, besides the level of development, can explain export diversification. Following economic theory, we will consider: variables relative to the size of countries (because of the possible presence of increasing returns effects, i.e., small countries should be more specialised than large ones); measures linked to human capital and technological progress (diversification dependents on the forces driving economic growth); measures linked to the geographical characteristics of countries (especially those which can influence transport costs and thus the ability to trade intensively); measures of institutional quality (focusing on factors that can possibly influence a sectoral structural change).

Finally, we will compare the results from different methods of estimation. In particular, we will introduce IV estimations taking into account possible endogeneity effects between diversification and economic development.

In short, the general objective is to uncover the forces which lie behind the link between overall specialisation patterns and development. We will use a panel data-set (60 countries, 1985-2004) and employ different synthetic indices of specialisation (obtained with disaggregated manufacturing exports data: SITC Rev.2, 3 digit and ISIC Rev.2, 3 digit - as a robustness check).

The rest of the paper is organized as follows: in Section 2 we sketch the theoretical background for our analysis; then – in Section 3 – we describe the data, the composition of our panel and specialization measurement issues. The core of the paper is presented in Section 4 which is entirely dedicated to the exploration of factors determining the degree of heterogeneity in export manufacturing structures, along with some robustness checks. Finally, Section 5 concludes. Our results suggest that manufacturing export specialisation is decreasing in country dimension. Additionally, countries located far from the economic core of the world and those for which barriers to trade are large tend to have less diversified manufacturing exports. We argue that our results confirm the role played by the economies of scale in the diversification process: when access to a big market (domestic and foreign) is possible, the diversification process is facilitated. Important policy implications, concerning the role of free trade in limiting risk exposure, stem from our analysis.

2. THEORETICAL BACKGROUND

Diversification originates from a structural change which is a multifaceted issue concerning a deep transformation of economies along their development paths, with strong

interconnections and mutual dependence among its multiple sides. For example, Matsuyama (2005:1) recognizes that it is a "complex, intertwined phenomenon" in which the income growth process and the various aspects of structural change, like "sector composition …, organization of the industry, financial system, income and wealth distribution, demography, political institutions, and even the society's value system", mutually affect and complement each other (also Kuznets on this: 1972, 1973). In short, structural change involves changes in: products, size and location of firms, labour force skills, legal and social innovations, etc. - all necessary and interconnected aspects of economic growth[4].

In our paper we focus on a specific, sectoral side of structural change (the identification of structural change with the sole sectoral transformation is largely accepted: UN-WESS 2006). This phenomenon depends on the impact of different technological advancements among different industries, resulting from the invention of new goods and cost-reducing innovations and from the heterogeneous response of the demand structure to increases in income due to the different income elasticities of various goods.

In particular, we do not consider trade specialisation patterns from a qualitative point of view (e.g., a given country exports certain kind of products more than others)[5]. We are rather interested in describing and explaining the changes in the quantitative distribution of trade activity across the manufacturing sector. If a country has a highly concentrated sector structure, that means low diversification of exports and a high degree of overall export specialisation (and major risk). Theoretical background can be derived from several strands of literature (roughly presented below).

Firstly, in modern growth models structural transformation of the economy enters as a fundamental input to the growth process (Barro and Sala-I-Martin, 1990; Grossman and Helpman, 1991). The usual symbolic representation of the final goods production function in this strand of literature is the following:

$$Y_i = A \cdot L_i^{1-\alpha} \cdot \sum_{j=1}^{N} (X_{ij})^{\alpha} \tag{1}$$

where Y is a final product of type i, L is the labor input, j refers to differentiated intermediate inputs and $0<\alpha<1$. The key variable X can be interpreted in two different ways which give

---

[4] For the sake of brevity, we have not cited many other aspects of structural change, e.g., we have not considered the ideological aspect. To have an idea, it is sufficient to think of the deep debate around the notion of life itself, generated by the introduction of modern bio-technologies (along with many legal problems).

[5] For an example of a qualitative assessment of export specialisation at international level, see Hausmann et al. (2007).

origin to two perspectives of looking at the structural transformation process. We can interpret X in two ways: firstly as the quality-adjusted quantity of the jth type of intermediate goods i.e.:

$$X_{ij} = \sum_{k=0}^{kj}(q^k \cdot x_{ijk})$$ (2)

where q^k is a quality indicator and aggregate X results as a weighted sum, so that equation (2) defines the 'quality content' of intermediate goods. This interpretation is typical for the Schumpeterian models of 'creative destruction' (new qualities intermediate inputs replace the old ones: Aghion and Howitt, 1992).[6] We could also think of X as the quantity (not qualitatively adjusted as before) of the jth type of intermediate goods. In this case, we are in the dominion of the so-called models with an expanding (intermediate) product variety, with no substitution among inputs (Grossman and Helpman, 1991: 43-83).

The so-called 'economic dualism' literature is another relevant place to look at. Among many existing contributions, we can identify a few common lines: the economy usually consists of two sectors (traditional and modern), and this kind of dualism is a consequence of: differences in production functions (technology) and/or consumer preferences (elasticities) between goods matched with functional linkages between sectors (Matsuyama 1991, 1992; Temple and Woessman, 2006).[7] However, there are no indications, relevant to our purposes, on the evolution of diversification along the process of growth. Nevertheless, an important feature of 'economic dualism' approach is that it allows sectoral changes to happen not only in intermediate goods' markets (like in the aforementioned growth theory) but also in the final goods markets. Moreover, supply and demand forces are equally important. From the demand perspective, as countries develop, patterns of consumption preferences adjust to higher income levels (Engel type effects[8]): increasing output *per capita* means modifications in the structure of the economy through a shift towards goods with higher demand elasticity. This mechanism, in turn, influences sectoral

---

[6] Qualitative aspects of economic structure are also underlined in other supply-side contribution (Lucas, 1988) and in the Keynesian demand-side literature (see Thirlwall, 1979; McCombie and Thirlwall, 2004).

[7] Usually, authors consider several other characteristics such as: frictions in the economy (their strength explains dualism persistence) and the possibility that dualism emerges as an endogenous process (for example due to the presence of externalities in the advanced sector).

[8] The so-called Engel's Law states that goods in demand have different income elasticities, thus along the process of economic growth which implies growing income per head. Structural demand shifts may also cause structural transformation.

productivities which change relative prices and, again, the structural composition of the economy.

Finally, in 'new economic geography' models (Fujita, Krugman, Venables, 2001), a 'love for variety' (Dixit-Stiglitz, 1977) is introduced in consumers' utility functions. In this literature, both static and dynamic models (incorporating growth equations) stress that economic activity and trade tend to expand through a rise in the number of differentiated goods available to the population.

As far as the justification of a low degree of diversification at the initial stage of growth is concerned, Acemoglu and Zilibotti (1997) provide a theoretical framework which emphasises limited diversification opportunities at lower levels of development because of the scarcity of capital and indivisibility of investment projects. Growing GDP *per capita* is usually linked with dynamic changes regarding the quality of institutions, human capital or widely understood conditions for 'doing business' which altogether favour a more dynamic and heterogeneous economic structure. Development goes hand-in-hand with better diversification climate, and that is also why more diversified (i.e., less specialised) structures of economic activity can run parallelly with higher levels of *per capita* output.

In the following parts of the paper we implicitly incorporate these theoretical arguments into the empirical framework, assessing factors of the export diversification process. In particular, our empirical contribution assesses (some of) the determinants of sectoral transformation resulting in a bigger variety of exported goods without distinguishing whether they are for intermediate or final use.

3. DATA AND MEASUREMENT ISSUES

3.1. Panel composition and the data

As stated in the introduction, we measure specialisation in terms of internationally exchanged goods using manufacturing export data. Sectoral export statistics come from the UN Comtrade database (available through WITS[9]) and are classified following two typologies of disaggregation schemes - SITC Rev.2, 3 digit in the major analysis and ISIC Rev.2, 3 digit - used at a later stage as a robustness check.[10] [List of sectors in Table R1 and Table R2 in the Appendix].

---

[9] World Integrated Trade Solutions, available from www. wits.worldbank.org

[10] In the main analysis, we use export specialisation measures calculated with SITC Rev.2 3 digit manufacturing data because changes in specialisation patterns are more likely to take place among product groups within the

Data availability was the only criteria of choice for the dataset, and in the end our analysis covers manufacturing exports from 60 countries and the time horizon of 20 years (1985-2004).[11] The complete list of countries can be found in Table A1 in the appendix.

We limit our analysis to manufacturing data: they are less sensible to natural and very specific conditions than other sectors (for example, primary sectors), while at the same time they cover a large part of countries' exports. Manufacturing is defined as sectors grouped under headings 5 to 8: in 2004 manufacturing accounted for 82% of the total exports reported by 60 countries present in our sample. After the elimination of items with a very pronounced presence of missing values[12], we kept 149 SITC industries which gave us 176714 sectoral observations in the SITC dataset.[13] With these disaggregated sectoral data we compute our specialisation indexes.

As far as additional statistics are concerned, GDP *per capita* (in 2000 int. US$), population size and the degree of openness come from PWT 6.2 (Heston *et al.* 2006). Human capital and technological variables come mainly from UNESCO. The Fraser Institute and the World Bank are our primary sources of institutional variables: references can be found in

---

same aggregated sector than between aggregate sectors. Greater detail is desirable here. Thus we give priority to SITC database, giving us more information than the data classified into 28 ISIC broad sectors. Concordance tables between the two classifications are available from WITS. We use SITC revision 2 instead of the more recent revision 3 or 4 because the older revision gives us the possibility to extend the time span of our analysis back to the 1980's (while for example SITC revision 3 has been used from 1988 onwards), and many countries have never switched their statistics to newer revisions.

[11] Countries with a very pronounced presence of missing cells have been excluded from the analysis. In a few papers (Cadot et al., 2011; Easterly et al. 2009), the number of nonzero sectoral export flows is used as a direct evidence of diversification (zero sectoral export flows as evidence of concentration); Tamberi (2010) reveals a growth trend, at world level, of zero sectoral export flows, interpreting this as a statistical bias, possibly influencing specialization indexes. Along these lines, we preferred to work with a (quasi) balanced dataset because in Comtrade set of export statistics missing cells are not casually distributed but rather concentrated in older periods and poorer countries. Thus the inclusion of more countries with a considerable number of sectoral missing observations would have biased our results. We could have included more recent export data but our main sources of additional variables report complete cross country statistics up to 2004. It would have been possible to include years prior to 1985, but it would have meant the exclusion of China from our analysis (complete Chinese export statistics are available only since mid-1980s.). Given China's importance in the world's economy and trade, we have decided not to do so.

[12] Namely: 688 (Uranium depleted in u235&thorium) and 675 (Hoop & strip of iron/steel, hot-roll).

[13] Randomly distributed missing values (6% of total) have been filled in through interpolation/extrapolation techniques. In order to have a balanced panel we had to replace exports with 0 in 0.4% of the cells.

Gwartney and Lawson (2007) for the former source and in Kaufmann *et al.* (2008) for the latter. Geographical characteristics are for the most part based on Gallup *et al.* (1999). We also use distances from CEPII. In addition, we employ micro data from European Values Study Group and Word Values Survey (2006) for the construction of institutional quality indicators.

3.2 Measurement of the degree of exports diversification

In order to measure the degree of diversification of countries' manufacturing exports, we use relative Theil index (another measure: relative Gini index will be used as a robustness check).[14] Given *m* industries (sectors) present in *n* countries, denote X_{ij} as a value of exports from sector *j* of country *i*. The share of exports of products from sector *j*=1, 2,…*m* in total exports of country *i*=1,2,…*n* as*:*

$$s_{ij} = X_{ij} / \sum_j X_{ij} \tag{3}$$

Analogically, we define the typical share of industry *j* in total 'world'[15] exports as:

$$w_j = \sum_i X_{ij} / \sum_i \sum_j X_{ij} \tag{4}$$

Then, we calculate relative Theil entrophy index defined as (Cowell 1995:49):

$$\mathrm{Re}lTheil_i = \sum_{j=1}^{m} \left(s_{ij} \cdot \ln \frac{s_{ij}}{w_j} \right) \tag{5}$$

The lower bound of Theil indices is 0 while the upper limit is equal to ln*(m),* where *m* is the number of sectors (industries). The index is positively related to the degree of overall specialisation – the bigger its value the higher the specialisation. Thus – as we use specialisation and diversification terms as antonyms - high values of *RelTheil* are associated with less diversified export structures than the overall benchmark trend in the sample of countries.

As dependent variable of the specialisation curve estimations we have a series of overall specialisation measure (*RelTheil_SITC*) composed of 1200 pooled observations (*n*=60 and *t*=20). Summary statistics are presented in Table 1: variability 'between' is much higher than 'within' which means that we can observe a considerable dispersion of the degree of

---

[14] Measures of specialisation are adopted from commonly used indices of inequality and concentration (see Iapadre 2001). The class of most popular relative indices includes: relative Gini index, relative Theil index, dissimilarity index, and Krugman specialisation index. We have chosen Theil measure given its desirable properties of independence of scale and 'population' size.
[15] Note that 'world' here is treated conventionally because it consists of those *m*=60 countries which are included in our analysis and not *all* world economies. As a result, we do not use the benchmark value w_i as the 'real' industry share in total world exports but rather as the share referring to its part consisting of *m* economies. However, we cover a very large part of total world exports: the countries included in our sample in 2004 amounted to 84% of the total world trade value and 90% of world manufacturing trade (total values refer to 160 countries for which the disaggregated data are available from the UN Comtrade database).

diversification between each country in our sample but not a big variability of sample observations about their separate means (*i.e.,* dispersion around a country's medium degree of export specialisation registered between 1985 and 2004 is lower than cross country variability). As a consequence we expect cross section variability to play a major role in explaining international diversification patterns.

Table 1. Summary statistics for Relative Theil index obtained with SITC data

		Mean	Std. Dev.	Min	Max	Obs
$RelTheil\_SITC_{it}$	Overall	1.28	0.88	0.10	4.01	N =1200
	Between		0.86	0.12	3.29	n = 60
	Within		0.23	0.41	2.74	T = 20

Source: own elaboration

In order to give a sample of between country dispersion of diversification patterns we examined countries that are characterized by the lowest/the highest values of the index.[16] The differences across countries are very big: for example, taking into account the relative Theil measure, in 2004 the most specialized country – Trinidad and Tobago - had 28 times more concentrated export structure than the country with the most diversified export structure - Germany. Another observation is that rankings do not change a lot. Many countries which were classified as those with the most diversified (thus the least specialised) export structures in 1985, are still more or less in the same place in the ranking after 20 years. The same is also true for the most specialised (the least diversified) economies. For instance, correlation coefficient between the series of $RelTheil\_SITC_{1985}$ and $RelTheil\_SITC_{2004}$ is equal to 0.86. This was in part expected since we are dealing with the structural characteristics of economies, needing time to adjust, but it may also be a signal that manufacturing export specialisation is a persistent phenomenon, largely dependent on country specific characteristics slowly changing through time or features which are virtually time invariant (like geographical conditions).

4. DETERMINANTS OF EXPORT DIVERSIFICATION PROCESS

4.1. Econometric model

The starting point of our analysis is what different authors (de Benedictis et al, 2009; Parteka, 2010) have confirmed, that is, low levels of GDP *per capita* tend to be associated

---

[16] Due to space limitations only crucial findings have been presented.. However, detailed statistics concerning individual countries are available from the authors upon request.

with a low degree of relative economic structures' heterogeneity (thus high overall specialisation). Hence, the basic model has the following form:

$$DIV_{it} = \alpha + f(GDPpc_{it}) + \varepsilon_{it} \tag{6}$$

where *DIV* denotes the degree of manufacturing export diversification (measured in benchmark estimations by *RelTheil_SITC*), *GDPpc* is the level of development measured by real income per capita, *i* refers to countries and *t* to time. Finally, *f(.)* is a link function that can take on several forms - as nonlinearity (argued by the supporters of the U-shaped pattern of diversification along the path of growth: Imbs and Wacziarg, 2003 or Cadot et al., 2011) can be present in the data, we consider several functional forms associated with GDP per capita (linear, quadratic, logarithmic).

We argue that country specific effects can be relevant in the diversification process. Thus, in the first instance, the model (6) can be enriched by the sole inclusion of country fixed effects:

$$DIV_{it} = \alpha + f(GDPpc_{it}) + D_i + \varepsilon_{it} \tag{7}$$

In order to specify precisely what kind of variables determine diversification, we will gradually extend eq. (7) towards the full version of the model:

$$DIV_{it} = \alpha + f(GDPpc_{it}) + \sum_{k=1}^{K} \beta_k X_{k,it} + \sum_{l=1}^{L} \delta_l X_{l,i} + D_t + \varepsilon_{it} \tag{8}$$

where *X* is a set of explanatory variables, other than income per capita, which can determine the process of diversification. We consider both time invariant characteristics (such as geographical measures of location or some dummies defining the status of countries) denoted as X_k (*k=1,...,K*) and variables that vary through time (such as GDP) denoted as X_l (*l=1,...,L*). Time effects are captured by time dummies D_t.

The choice of explanatory variables is guided by the economic theory. We proxy country size in both geo-demographical and economic terms, measuring the former by population size (*POP*) while approximating the latter with total *GDP*.[17] Theoretical explanations on the link between the degree of overall specialisation and country size can be found in New Trade Theory (Dixit and Norman, 1980; Helpman and Krugman, 1985) arguing that market size directly affects the degree of product differentiation. According to the view presented in monopolistic competition models, bigger countries can produce wider range of

---

[17] We have also tried to use land area as a proxy of country size and the results are very similar, even though it is not a very robust variable and its economic interpretation is troublesome.

products (thus they are less specialised).[18] Human capital and technology characteristics, rooted in endogenous growth theory (Aghion and Howitt, 1998) can affect general conditions for product differentiation. Institutional setting is not only an important factor of growth (Rodrik *et al,.* 2004) but it also influences diversification opportunities. New Economic Geography models (Amiti and Venables, 2002; Venables and Limao, 2002) suggest that among the important factors that influence the economic structure of a country we may find the proximity to world markets and other geographical characteristics. Geography influences trade costs and may affect the ability to operate intensively in the international market (Frankel and Romer, 1999). Finally, trade liberalisation can act as market extension (Krugman and Venables, 1990; Dennis and Shepherd, 2011), and potential gains from trade may cause major product diversification (Costas *et al,.* 2008).

In all, we have considered 33 alternative explanatory variables possibly influencing the diversification process and referring to: country size, human capital, technology and R&D, institutions, geographical characteristics, trade and endowments. [See additional Table R3 in the appendix for full description of variables].

4.2 Results

4.2.1 First stage

In line with our empirical strategy exposed in the introduction, we first examine the link between a measure of specialization with the development level only and, eventually, country fixed effects.

GDP per capita (rescaled and expressed in thousands of US$, 2000) is introduced in various functional forms – linear (Model 1) and quadratic (Model 2). Finally, we show log-log model (Model 3) with both *RelTheil* and *GDPpc* expressed in natural logs so that the estimated coefficient can be conveniently interpreted as elasticity[19].

Ordinary Least Squares estimates shown in Table 2 should only be treated as a starting point for the analysis. Table 3 presents analogous results obtained with LSDV estimation (thus correcting the model for the inclusion of country specific effects). In Table 4, we demonstrate the results of the IV FE estimation, correcting the basic ones for the inclusion of country fixed effects, but also taking into account potential endogeneity between the degree of

---

[18] Hummels and Klenow (2005) empirically estimate the link between economy size (measured by total income) and the overall degree of specialisation.

[19] The log-log model is the linearization of a general fractional model: $DIV_{it} = a/GDPpc_{it}^{\,b} + \varepsilon_{it}$

diversification and income per capita. The general result is that independently on the estimation procedure and functional form of the model, development level is always significantly and negatively related to the measure of specialization, being an opposite of diversification. Hence, as countries develop, export specialization decreases (diversification of exports increases). Quadratic formulation turns out to be significant, too. So some reversal of the trend is plausible, but the log-log model seems to fit better the data.

However, first of all we address the issue of country specific characteristics, possibly influencing the diversification process. From estimated coefficients relating to the development level in a country fixed effects framework (results of LSDV estimator presented in Table 3), we can confirm that the development level is negatively related to the degree of specialization. A rise in income per capita by 1% is associated with a decline in specialization measure by approximately 0.5% (thus export diversification process takes place along the path of growth). So, even when we take into account cross-country heterogeneity in the panel, diversification takes place in the course of economic development, but it is rather slow.[20] Importantly, in the case of LSDV estimates, test of joint significance of country specific effects confirms that their inclusion in the model is correct. Hence, the diversification process of countries depends not only on their development level, but there are other characteristics of countries that matter and should be taken into account in the analysis.[21]

Another problem that needs to be accounted for is the eventual endogeneity in the model. Indeed, it is possible that there is reverse causation among our main variables of interest, level of development and diversification: a wider set of exported goods could enhance the possibility of development, enlarging processes of knowledge accumulation.[22]

---

[20] In order to illustrate this phenomenon we can look at a rapidly developing country like China which in 20 years between 1985 and 2004 moved from the level of GDP per capita of US$ 1134 in 1985 to US$ 5333 in the year 2004 (both values in constant prices, 2000). At the same time, China's relative diversification of manufacturing exports (measured with RelGini_SITC) increased by 70%.

[21] Additionally, in line with de Benedictis et al.(2009) and Parteka (2010), we have considered semi parametric GAM plots which, thanks to the inclusion of country specific effects, can be interpreted as typical 'specialisation curves' along the path of economic development. Estimated lines are monotonically decreasing and confirm progressing diversification as income per capita grows (results available upon request from the authors). [See Additional Figure R1 and Table R4].

[22] Such linkages are closely related to those widely discussed in the productivity-export literature based on micro-data analysis (a survey in Wagner, 2007).

Table 2. Results of first stage estimation (diversification and development level) - OLS

	Dep.variable: *DIV*		
	RelTheil_SITC	*RelTheil_SITC*	*ln_RelTheil_SITC*
	(1)	(2)	(3)
	OLS	OLS	OLS
GDPpc (in 1000 US$)	-0.047***	-0.076***	
	[-18.820]	[-7.731]	
squared GDPpc		0.001***	
		[3.069]	
ln_GDPpc			-0.474***
			[-21.486]
cons	1.895***	2.025***	4.285***
	[47.957]	[35.037]	[21.147]
N	1200	1200	1200
R2	0.23	0.23	0.28
p	0.000	0.000	0.000

Note: t-statistics in parenthesis; *, ** and *** denote significance at 10%, 5% and 1% level, respectively.

Table 3. Results of first stage estimation (diversification and development level with country specific fixed effects) - LSDV

	Dep.variable: *DIV*		
	RelTheil_SITC	*RelTheil_SITC*	*ln_RelTheil_SITC*
	(1)	(2)	(3)
	LSDV	LSDV	LSDV
GDPpc (in 1000 US$)	-0.016***	-0.108***	
	[-5.471]	[-10.660]	
squared GDPpc		0.002***	
		[9.430]	
ln_GDPpc			-0.495***
			[-17.441]
Cons	1.098***	1.796***	4.479***
	[18.049]	[19.020]	[16.961]
N	1200	1200	1200
R2	0.93	0.94	0.97
P	0.000	0.000	0.000
F_joint	201.24	216.43	404.67
P_joint	0.000	0.000	0.000

Note: t-statistics in parenthesis; *, ** and *** denote significance at 10%, 5% and 1% level, respectively.
F_joint, p_joint – refer to test of joint significance of country specific effects

Table 4 demonstrates results of the models analogous to those described above, but obtained with instrumental variables (IV) FE estimator. As two-way relations between

diversification and economic development are plausible (Hesse, 2008)[23], our measure of economic development is instrumented (by its own lags). Consequently, at the same time we take into account both the importance of country fixed effects, as well as potential endogeneity issues. Obtained coefficients associated with GDP per capita (again in various functional forms) are very similar to LSDV estimates - the diversification process takes place (slowly) as GDP per capita grows and, as before, importance of country specific characteristics is confirmed by testing their joint significance. As far as endogeneity of development in the model is concerned, the Davidson-MacKinnon (1993, p.236) test of exogeneity with *GDPpc,* considered to be potentially endogeneous, is applied. In none of the models the null hypothesis of exogeneity (i.e., that OLS would be a consistent estimator of the equation) cannot be rejected: p-value of the test never exceeds the standard 5% threshold. In other formulations, endogeneity turns out to be a potential issue. Thus, we shall formally proceed with the IV estimator, treating GDP per capita as potential source of endogeneity in the model. However, it should be noted that the coefficients associated with income per capita estimated with the IV FE method are practically identical to those resulting from the LSDV estimator.

Table 4. Results of first stage estimation (diversification and development level with country specific effects and corrected endogeneity) - IV FE

	Dep.variable: *DIV*		
	RelTheil_SITC	*RelTheil_SITC*	*ln_RelTheil_SITC*
	(1)	(2)	(3)
	IV FE	IV FE	IV FE
GDPpc (in 1000 US$)	-0.016***	-0.103***	
	[-4.953]	[-9.403]	
squared GDPpc		0.002***	
		[8.294]	
ln_GDPpc			-0.485***
			[-15.950]
Cons	1.479***	2.090***	4.383***
	[34.969]	[24.827]	[15.756]
N	1140	1140	1140
R2	0.22	0.22	0.28
p_fe	0.000	0.000	0.000
F_exo	4.7	4.96	10.77
p_exo	0.03	0.007	0.001

Note: development level instrumented by its lag; t-statistics in parenthesis; *, ** and *** denote significance at 10%, 5% and 1% level, respectively.
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)

---

[23] Among recent contributions, Starosta de Waldemar (2011) uses GMM estimations to correct for this bias.

First step results permit us to conclude that since cross-country variability is fully captured by fixed effects, income per capita explains essentially the time evolution of the diversification process. Consequently, our interpretation that diversification is a characteristic linked to the development process seems suitable. What remains is to uncover concrete country specific factors determining diversification opportunities captured in the FE framework.

4.2.2 Second stage - explaining country fixed effects as determinants of diversification process

The result of first stage estimations is indeed clear, but the use of closer undefined fixed effects is in some way not satisfying, since they collect many features of a country that remain unknown. We have adopted a gradual approach to choose, out of 33 alternative variables[24], only those to be put in the final model. We first checked for potential collinearity between variables deciding on the ones that could not be put in the model contemporarily (e.g. GDP and POP)[25]. Then, we followed a step-by-step procedure of variable selection, eliminating insignificant variables from the most general form of the model, gradually restricting the set of right hand side variables. Resulting final estimations contain only significant ones.

In Table 5 we show these final results of multivariate second stage estimation, aiming at revealing the importance of 'tangible' characteristics, incorporated previously in country fixed effects, in the diversification process. While many of our variables are time invariant, like geographical ones, some of them can change in time, more or less slowly (such as GDP); as a consequence, their role in explaining diversification may not be limited to cross-sectional dimension. In order to check for common business cycle effects, we include a set of time dummies in each model.

---

[24] We have considered: GDP per capita; *measures of size*: total population and GDP; *human capital measures*: Enrolment in secondary education as % of population, Enrolment in tertiary education as % of population, Illiteracy rate (% population aged 15-24); *R&D measures*: Spending on R&D as % of GDP, Number of researchers per mln citizens; *institutional indices* concerning: Government size, Legal structure and security of property rights, Sound money, Freedom to trade internationally, Regulation of credit, labor and business, Summary Economic Freedom Index, Voice and accountability, Political stability and absence of violence/terrorism, Government effectiveness, Regulatory quality, Rule of Law, Control of corruption, three alternative summary indices of governance, Trust; *geographical indicators*: % of land within 100 km of the ocean coastline or navigable river, % of land area in tropical zones, Distance (in km) from the closest major market (New York, Rotterdam or Tokyo); *trade measures*: Spatial correlation of export structures between neighbouring countries, Exports plus Imports as % of GDP, RTA dummy, % of manufacturing exports; *measures of endowments*: share of petrol and petroleum products in total exports (%), Value added in agriculture as % of GDP, Share of employment in agriculture in total employment.
[25] We performed formal collinearity diagnostic tests by computing the condition number.

Drawing on the first step results, we proceed with a formulation in which all variables are expressed in logs, and we basically rely on the IV estimator (2SLS). The use of the instrumental variable estimation (with *GDPpc* being a potential source of endogeneity) is justified by two alternative formal tests of exogeneity: Durbin-Watson chi^2 statistic and Wu-Hausman *F* test.

We present estimations with (Models 1-5 estimated with the IV method) and without (Models 6 and 7 estimated with OLS) direct introduction of GDP *per capita* as one of the explanatory variables. We gradually show estimations with measures of size, location and trade that turned out to be robust. Knowing that *RelTheil* is inversely related to the degree of diversification, if estimated coefficients associated with some explanatory variables are negative, then they are among positive determinants of the diversification process.

Table 5. Second step estimation results (multivariate estimations – determinants of export diversification), IV and OLS

	Dep.variable: *DIV (ln_RelTheil_SITC)*						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	IV-2SLS	IV-2SLS	IV-2SLS	IV-2SLS	IV-2SLS	OLS	OLS
ln of GDPpc	-0.352***	-0.183***	-0.117***	-0.103***	-0.337***		
	[-19.210]	[-9.756]	[-5.666]	[-4.986]	[-15.072]		
ln of GDP	-0.231***	-0.174***	-0.186***	-0.178***		-0.179***	
	[-26.941]	[-21.287]	[-21.887]	[-20.652]		[-20.580]	
ln of POP					-0.216***		-0.150***
					[-23.604]		[-16.896]
ln of MarketDist		0.312***	0.263***	0.251***	0.227***	0.287***	0.359***
		[17.826]	[15.559]	[14.882]	[13.810]	[18.351]	[23.228]
ln of FreeTrade			-0.976***	-0.858***	-0.840***	-0.994***	-1.236***
			[-10.910]	[-9.352]	[-9.543]	[-12.269]	[-14.139]
RTA				-0.171***	-0.158***	-0.189***	-0.261***
				[-4.871]	[-4.676]	[-5.417]	[-7.234]
Cons	8.652***	3.320***	5.293***	4.885***	5.067***	4.192***	0.980***
	[35.521]	[9.013]	[14.004]	[12.754]	[14.159]	[12.161]	[3.738]
N	1140	1140	1064	1064	1064	1120	1120
R2	0.56	0.65	0.7	0.71	0.73	0.7	0.67
F_exo_wu	8.271	5.523	17.694	17.832	23.838		
p_exo_wu	0.004	0.019	0.000	0.000	0.000		
chi2_exo_durb	8.372	5.609	17.799	17.953	23.864		
p_exo_durb	0.004	0.018	0.000	0.000	0.000		
time dummies	yes	yes	yes	yes	yes	yes	yes

Note: GDPpc instrumented by its lags, t-statistics in parenthesis; *, ** and *** denote significance at 10%, 5% and 1% level, respectively.
GDPpc – real per capita income from PWT 6.2, *GDP* – real Gross Domestic Product from PWT 6.2, *POP* – population from PWT 6.2, *MarketDist*- distance (in km) from the closest major market (New York, Rotterdam or Tokyo) from Gallup *et al.* (1999), *FreeTrade* - Freedom to trade internationally (size of restraints that affect international exchange: tariffs, quotas, hidden administrative restraints, size of trade sector relative to expected, black market exchange rates and international capital market controls) - index from the Fraser Institute (a higher score indicates a greater degree of economic freedom in trade), *RTA*=1 if active member of selected Regional Trade Agreement: EU (from the moment of accession), EEA, EFTA, NAFTA, MERCOSUR, ASEAN (own elaboration, based on WTO).
F_exo_wu and p_exo_wu – refer to Wu-Hausman test of exogeneity (IV)
chi2_exo_durb and p_exo_durb – refer to Durbin test of exogeneity (IV)
Source: own elaboration

First stage results are confirmed as in all the models presented in Table 5 the coefficient associated with per capita income is statistically significant – it appears to be negatively related to the Theil measure of specialization. Thus, as per capita income grows, the degree of specialization decreases and the diversification of exports proceeds. However, the strength of this relationship depends on the inclusion of additional variables. In particular, when we correct for the location characteristics of the countries (in particular, distance from major world markets) and GDP, then the role played by the development level in the diversification process diminishes (Models 2-4).

Eventually, apart from the development level measured by GDP per capita, three main factors appear to determine the degree of manufacturing exports diversification: country size (measured in the form of *GDP:* models 1-4 and 6, or alternatively, *POP:* model 5 and 7), easy access to main world markets (measured by *MarketDist*), participation in trade agreement and institutional framework, at least in the form caught by *FreeTrade* (a variable describing freedom to trade internationally: size of restraints that affect international exchange: tariffs, quotas, hidden administrative restraints, size of trade sector relative to expected, black market exchange rates and international capital market controls) and RTA agreements. Other variables from our set lost their significance when put with others in the estimation.

Results reported in Table 5 indicate that, ceteris paribus, an increase in country size by 1% can be associated with an increase in the degree of exports diversification by approximately 0.2%. An increase in the distance from major markets by 1% can be associated with a decrease in the degree of exports diversification by approximately 0.2-0.3%. Moreover, a better quality of institutions related to trade activity, affects the diversification process positively. [26] Finally, participation in Regional Trade Agreements also fosters diversification. Hence, factors influencing access to big markets (domestic or foreign – through trade openness) are crucial. The goodness of fit of all models is fairly good and with just these few variables we were able to explain up to 73% of the variability in the diversification patterns. Hence, we have managed to uncover a large part of the factors hidden behind country fixed effects from first stage estimations.

We can conclude that the exports diversification process depends on two main causes. Firstly, in terms of the development of countries, it probably depends on the initially sketched theoretical model that is linked to the supply side (technological process of product innovation) and to the demand side (consumers' love for variety). Secondly, diversification

---

[26] A higher score of the index *FreeTrade* indicates a greater degree of economic freedom in trade.

depends on the size of accessible markets (both domestic and foreign – possibly to exploit in the case where distance and barriers to trade are not too great), and we interpret this as an indication of the role of scale economies in trade specialization (and in production), i.e., in the presence of scale economies, diversification is possible only in large markets. Note that these observations hold true even if we exclude from the model direct income per capita measure (Models 6 and 7).

4.2.3 Robustness checks

First of all, we checked whether our results were sensitive to the choice of a specific measure of specialization/diversification. We computed an alternative one: the relative Gini index. It is defined as in Amiti (1999): first rank in ascending order sectoral Balassa indexes (BI_{ij}), calculated for each country and sector as a ratio of s_{ij} (eq.3) to w_j (eq.4) where, $i=1,...,n$ refers to countries and $j=1,...,m$ to sectors. Then, for each country i we represent the cumulative of the denominator of BI ($p_j = \sum_{k=1}^{j} w_k ,\vee j=1,...,m$) on the horizontal axis and the cumulative of the numerator of BI ($q_{ij} = \sum_{k=1}^{j} s_{kj} ,\vee j=1,...,m$) on the vertical axis. The Gini index can be calculated as twice the area between the 'Lorenz curve' and the 45 degree line which is associated with a case when country i has the same pattern of revealed comparative advantage as the benchmark group of countries. In order to compute the relative Gini index we use the approximate statistical formula (Piccolo 1998: 140):

$$\mathrm{Re}\,lGini_i = \sum_{j=0}^{m-1}(p_j - q_{ij})/\sum_{j=1}^{m-1}(p_j) \qquad (6)$$

The relative Gini index may vary from 0 (when $q_{ij} = p_j$ for all j) to 1 (when $q_{ij} = 1$ and $p_j < 1$ for $j=m$). A higher *RelGini* corresponds to a higher degree of overall specialization (thus lower diversification). [See additional Table R5.1 and Table R5.2].

Then, we considered a change in the disaggregation scheme. We recalculated specialisation measures using export statistics (for the same set of countries and the same time span) classified according to ISIC rev. 2, 3 digit schemes (28 manufacturing sectors).[27] Comparing the results obtained from the data disaggregated either into broad sectors (ISIC dataset) or into more detailed groups of products (SITC dataset) permits us to check the

---

[27] Out of 33600 sectoral observations, 2.6% of missing values have been filled in through interpolation/extrapolation.

relevance of the disaggregation level for the assessment of export specialisation patterns and their determinants. [See additional Table R6.1 and Table R6.2]). [See additional Table R7.1 and Table R7.2]

Moreover, we also considered modifications in the estimated multivariate model (second stage), concerning linear and quadratic form of *GDPpc* in US$1000 (instead of the log-log model presented in the main text) and other explanatory variables rescaled accordingly (*GDP* in 10^{12}US$, *POP* in 10^6, *MarketDist* in 1000km [See additional Table R8].

Summary statistics of the alternative measures (*RelGini_SITC, RelTheil_ISIC, RelGini_ISIC)* used in the robustness checks section are presented in Table A2 in the Appendix. As in the case of our benchmark measure (*RelTheil_SITC*), cross country variability exceeds the variability around the national mean which again suggests that country specific effects play a role in the diversification process. The measurement of overall manufacturing export specialisation seems to be robust to changes in the disaggregation scheme. The correlations between four alternative indices of overall manufacturing export diversification are very high (Table A3 in the appendix).

We repeated the whole two stage procedure, confirming our basic results. First stage estimations are not sensitive to the changes described above[28]- export diversification takes place along the path of growth, and country fixed effects are relevant. Modified second stage models again allow us to conclude that variables describing country size, distance from the major markets and institutional conditions favouring free trade activity are statistically significant, and their signs remain the same as in the benchmark estimates. Values of the estimated coefficients do not change in sign or significance[29]. Despite the changes in the estimated model, the goodness of fit of second stage multivariate estimations is still pretty high (up to 0.75) which confirms that those few variables explain a large part of variability in international export diversification patterns.

---

[28] Only when *RelGini_SITC* and *RelTheil_ISIC* is used, results of exogeneity tests change when GDPpc is introduced in the model in linear or quadratic form.

[29] Only when *RelTheil_ISIC* is used as diversification measure, *GDPpc* looses its significance in models controlled for *GDP*, *MarketDist*, *FreeTrade* (and *RTA*).

5. CONCLUSIONS

The aim of this study was to fill the gap existing in the empirical literature on diversification - development nexus. Our main purpose was to uncover what country specific factors determine the diversification process in the course of a country's economic development. This argument is of crucial importance due to the risk which countries face when their trade structure is not heterogeneous enough to bear idiosyncratic shocks. Thanks to the inclusion of the 20-year data on 60 world countries in our sample, we were able to trace the trends regarding the degree of diversification of export manufacturing in a vast group of economies at very different stages of economic development.

Indeed, we confirm that poor countries tend to have highly homogeneous (specialised, poorly diversified) export structures. In line with the findings of the other authors who have also applied relative measures of diversification (referring the export structure of a country to trends in the rest of the group), we confirm that economic development is accompanied by the tendency towards manufacturing exports despecialization. However, the crucial question was to find out what additional determinants linked to a country's geographical, institutional or economic conditions, other than its level of development, are responsible for an export diversification process.

We adopted a gradual analysis, firstly by assessing the general importance of country specific effects in the diversification process and, secondly, by revealing what characteristics of a country are hidden behind these closer undefined country fixed effects. Importantly, our econometric strategy permitted us to correct for potential endogeneity issues in diversification-development estimations.

It turns out that among the many factors which theoretically could affect sectoral composition of trade structures, the most significant and robust ones are those describing: country size, its location (in particular the distance from major world markets) and trade conditions (freedom to trade internationally and participation in trade agreements). Unsurprisingly, manufacturing export specialisation is decreasing in country dimension - small countries tend to be more specialised (and are thus more exposed to risk). Additionally, being far from major markets weakens their ability to extend their market size. Thus countries located far from the economic core of the world tend to have less diversified manufacturing exports. The same difficulty occurs if there are trade barriers.

We argue that our results confirm the role played by the economies of scale: when access to a big market (domestic and foreign) is possible, the diversification process is facilitated. An important policy implication follows: in order to limit a country's exposure to risk, we should

allow it to exploit international markets through open trade. It is a robust result, not sensitive to changes in the disaggregation level of exports, way of export specialisation measurement or estimation strategy.

BIBLIOGRAPHY AND DATA SOURCES

Acemoglu D. and F. Zilibotti (1997). Was Prometheus unbound by chance? Risk, diversification and growth. *Journal of Political Economy,* 105(4): 709-751.

Aghion P. and P. Howitt (1992). A model of growth through creative destruction. *Econometrica,* 60(2): 323-351.

Aghion and Howitt (1998). *Endogenous Growth Theory*. The MIT Press. Cambridge(MA).

Amiti M. (1999). Specialisation patterns in Europe. *Review of World Economics,* 135(4): 573-593.

Amiti M. and A.J. Venables (2002). The geography of intra-industry trade. [in:] P. J. Lloyd and H.-H. Lee (eds.) *Frontiers of Research in intra-industry trade.* Palgrave Macmillan, Basingstoke.

Barro R.J. and X. Sala-I-Martin (1990). *Economic Growth*. McGraw-Hill

Cadot O., C. Carrère and V. Strauss-Kahn (2011). Export diversification: what's behind the hump?. *The Review of Economics and Statistics*, 93(2): 590-605.

Carrère C., O. Cadot, V. Strauss-Kahn (2009). Trade diversification, income, and growth: what do we know?, CERDI, Etudes and Documents E 2009.31

CEPII - Centre d'Etudes Prospectives et d'Informations Internationales. Bilateral distance data (dist_cepii) available at: www.cepii.fr

Costas A., S. Demidova, P. Klenow and A. Rodriguez-Clare (2008). Endogenous Variety and the Gains from Trade. NBER Working Paper No.13933; published in *American Economic Review Papers and Proceedings (*2008), 98(2): 444-450.

Cowell F. (1995). *Measuring inequality*. Prentice Hall/Harvester Wheatsheaf, London (2nd edition).

Davidson R. and J.G. MacKinnon(1992). *Estimation and Inference in Econometrics*, Oxford University Press, New York.

De Benedictis L., Gallegati M. and M. Tamberi (2009). Overall specialisation and development: countries diversify!. *The Review of World Economics*, 145(1): 37-55.

De Ferranti D., G. Perry, D. Lederman, W. Maloney (2002). *From Natural resources to the knowledge economy: trade and job quality*, the World Bank

Dennis A. and Shepherd B. (2011), Trade facilitation and export diversification, The World Economy, 34: 101–122

Dixit A. and V.D. Norman (1980). *Theory of International Trade: a dual, general equilibrium approach*. Cambridge University Press, Cambridge.

Dixit A. and Stglitz J. (1977), Monopolistic competition and optimum product diversity, American Economic Review, 67 (3): 297-308

Easterly W., A. Reshef, J. Schwenkenberg (2009). The power of exports. World Bank, Policy Research Working Paper 5081, October

European Values Study Group and Word Values Survey Association (2006). European and World Values Surveys four-wave integrated data file, 1981-2004, v.20060423. Data available at www.europeanvalues.nl

Frankel J. and Romer D. (1999), Does trade causes growth?, *American Economic Review*. 89(3): 379-399

Fujita M., Krugman P., Venables A.J. (2001), *The spatial economy*, the MIT press

Gallup J.L., J. Sachs and A.D. Melinger (1999). Geography and Economic Development. *International Regional Science Review,* 22 (2): 179-232; database available at www.cid.harvard.edu/ciddata

Grossman G.M. and E. Helpman (1991). *Innovation and Growth in the Global Economy*. The MIT Press

Gwartney, James and Robert Lawson (2007). *Economic Freedom of the World: 2007 Annual Report.* Vancouver: The Fraser Institute. Data retrieved from www.freetheworld.com

Hammouda H.B., S.N. Karingi, A.E. Njuguna, M.S. Jallab (2010). Growth, productivity and diversification in Africa. *Journal of Productivity Analysis*, 33(2): 125-146.

Hausmann R., J. Hwang and D. Rodrik (2007). What you export matters. *Journal of Economic Growth,* 12(1): 1-25.

Helpman E. and P. Krugman (1985). *Market structure and foreign trade.* The MIT Press, Cambridge (MA).

Heston A., R. Summers and B. Aten (2006). *Penn World Table* Version 6.2, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, September 2006 version.

Hesse H. (2008). Export diversification and growth. Commission on Growth and Development Working Paper No. 21, The International Bank for Reconstruction and Development / The World Bank

Hummels D. and P.J. Klenow (2005). The variety and quality of a nation's exports. *The American Economic Review* 95(3): 704-723.

Iapadre, L. (2001). Measuring international specialisation. *International Advances in Economic Research*, 7: 173-83.

Imbs J. and R. Wacziarg (2003). Stages of Diversification. *American Economic Review,* 93(1): 63-86

Kalemli-Ozcan S., Sørensen B.E. and O. Yosha (2003). Risk sharing and industrial specialisation: regional and international evidence. *American Economic Review,* 93(3): 903-918.

Kaufmann D. , A. Kraay and M. Mastruzzi (2008). Governance matters VII: Aggregate and individual governance indicators, 1996-2007. World Bank Research Working Paper No.4654. Data downloadable at: www.worldbank.org/wbi/governance

Koren M. and S. Tenreyro (2007). Volatility and Development. *Quarterly Journal of Economics,* 122(1): 243-287.

Krugman P. and Venables A.J. (1990). Integration and the competitiveness of peripheral industry. [in:] Bliss C. and J. Braga de Macedo (eds.). *Unity with diversity in the European economy: the Community's southern frontier.* Cambridge University Press, Cambridge.

Kuznets S. (1972). Innovations and adjustment in economic growth, *Swedish Journal of Economics*, 74 (4): 431-451.

Kuznets S. (1973). Modern economic growth: findings and reflections. *American Economic Review*, LXIII (3): 247-258.

Lucas R. (1988), On the mechanics of economic development, *Journal of Monetary Economics*, 22: 3-42

Matsuyama K. (1991). Increasing returns, industrialization, and indeterminacy of equilibrium. *Quarterly Journal of Economics*, 106: 617-650

Matsuyama K. (1992). Agriculture productivity, comparative advantage, and economic growth. *Journal of Economic Theory*, 59: 317-334

Matsuyama K. (2005). Structural Change. [in:] Blume L. and S. Durlauf (eds.). *The New Palgrave Dictionary of Economics*, 2nd Edition, Macmillan (downloadable from http://faculty.wcas.northwestern.edu/~kmatsu/)

McCombie J.S.L and A.P. Thirlwall [ed.] (2004). *Essays on balance of payments constrained growth: theory and evidence.* Routledge, Taylor and Francis Group.

Minondo A. (2011 - forthcoming). Does comparative advantage explain countries' diversification level? *Review of World Economics*, Doi: 10.1007/s10290-011-0097-7

Parteka A. (2010). Employment and export specialisation along the development path: some robust evidence. *Review of World Economics (Weltwirtschaftliches Archiv)*, 145(4): 615-640

Piccolo D. (1998). *Statistica.* Il Mulino.

Rodrik D., A. Subramanian and F. Trebbi (2004). Institutions rule: the primacy of institutions over geography and integration in economic growth. *Journal of Economic Growth,* 9 (13): 131-165.

Starosta de Waldemar F. (2010). How costly is rent-seeking to diversification: an empirical approach, CES - Centre de Economie de la Sorbonne - working paper n. 2010.0

Tamberi M. (2010). Be Careful! A Short Note on a Possible Bias in (Trade) Structural Change Analysis. Università Politecnica delle Marche, Dipartimento di Economia, Quaderni di Dipartimento No. 341 (downloadable at http://193.205.135.133/dea/quaderni/pdf/341.pdf)

Temple J., Woessman L. (2006). Dualism and cross-country growth regression, *Journal of Economic Growth,* 11(3): 187-228

Thirlwall A.P. (1979). The balance of payments constraint as an explanation of international growth rate differences. *BNL Quarterly Review,* 32: 45–53.

United Nations - WESS (2006). *World Economic and Social Survey* – chapter II: Structural Change and Economic Growth. (http://www.un.org/esa/policy/wess/wess2006files/chap2.pdf)

United Nations Data Retrieval System (UNdata): www. data.un.org

United Nations Statistical Division (UNSD), *Commodity Trade Database (COMTRADE)*, available through World Bank's World Integrated Trade Solution (WITS): www.wits.worldbank.org

Venables A. and Limão (2002). Geographical disadvantage: a Hecksher-Ohlin-von Thünen model of international specialisation. *Journal of International Economics,* 58(2): 239-263

Wagner J (2007). Export and productivity: a survey of the evidence from firm-level data, *The World Economy* 30(1): 60-82

Wörtz J. (2005). Dynamics of trade specialisation in developed and less developed countries. *Emerging Markets Finance and Trade,* 41(3): 92-111.

APPENDIX

Table A1: List of countries and adopted abbreviations

Algeria DZA	El Salvador SLV	Kenya KEN	Philippines PHL
Argentina ARG	Finland FIN	Korea, Rep. KOR	Poland POL
Australia AUS	France FRA	Macao MAC	Portugal PRT
Austria AUT	Germany GER	Madagascar MDG	Saudi Arabia SAU
Barbados BRB	Greece GRC	Malawi MWI	Singapore SGP
BolithroughBOL	Hong Kong, China HKG	Malaysia MYS	Spain ESP
Brazil BRA	Iceland ISL	Mauritius MUS	St. Lucia LCA
Canada CAN	India IND	Mexico MEX	Sweden SWE
Chile CHL	Indonesia IDN	Morocco MAR	Switzerland CHE
China CHN	Ireland IRL	Netherlands NLD	Thailand THA
Colombia COL	Israel ISR	New Zealand NZL	Trinidad&Tobago TTO
Cyprus CYP	Italy ITA	Norway NOR	Tunisia TUN
Denmark DNK	Jamaica JAM	Oman OMN	Turkey TUR
Ecuador ECU	Japan JPN	Pakistan PAK	United Kingdom GBR
Egypt, Arab Rep. EGY	Jordan JOR	Peru PER	United States USA

Table A2. Summary statistics for alternative diversification measures (robustness check)

		Mean	Std. Dev.	Min	Max	Obs
RelGini_SITCit	Overall	0.69	0.19	0.24	0.97	N =1200
	Between		0.19	0.27	0.95	n = 60
	Within		0.045	0.42	1.05	T = 20
RelTheil_ISICit	Overall	0.85	0.65	0.04	3.28	N =1200
	Between		0.63	0.05	2.95	n = 60
	Within		0.19	0.05	2.10	T = 20
RelGini_ISICit	Overall	0.58	0.21	0.14	0.95	N =1200
	Between		0.20	0.17	0.93	n = 60
	Within		0.05	0.37	0.89	T = 20

SITC: Rev.2, 3 digit, 1985-2004, 60 countries, 149 manuf. industries
ISIC: Rev.2, 3 digit, 1985-2004, 60 countries, 28 manuf. sectors

Table A3. Correlation coefficients between alternative manufacturing specialization indices

	RelTheil_SITC	*RelGini_SITC*	*RelTheil_ISIC*	*RelGini_ISIC*
RelTheil_SITC	1			
RelGini_SITC	0.90	1		
RelTheil_ISIC	0.85	0.81	1	
RelGini_ISIC	0.89	0.93	0.92	1

Table R1. List of manufacturing sectors *SITC* Rev.2, 3-digit (headings 5-8)

code	productname	code	Productname
511	Hydrocarbons nes,& their halogen.&	711	Steam & other vapour generating boi
512	Alcohols,phenols,phenol-alcohols,&	712	Steam & other vapour power units,st
513	Carboxylic acids,& their anhydrides	713	Internal combustion piston engines
514	Nitrogen-function compounds	714	Engines & motors,non-electric
515	Organo-inorganic and heterocyclic c	716	Rotating electric plant and parts
516	Other organic chemicals	718	Other power generating machinery an
522	Inorganic chemical elements,oxides	721	Agricultural machinery and parts
523	Other inorganic chemicals	722	Tractors fitted or not with power t
524	Radio-active and associated materia	723	Civil engineering & contractors pla
531	Synth.org.dyestuffs,etc.nat.indigo	724	Textile & leather machinery and par
532	Dyeing & tanning extracts;synth.tan	725	Paper & pulp mill mach.,mach for ma
533	Pigments,paints,varnishes & related	726	Printing & bookbinding mach.and par
541	Medicinal and pharmaceutical produc	727	Food processing machines and parts
551	Essential oils,perfume and flavour	728	Mach.& equipment specialized for pa
553	Perfumery,cosmetics and toilet prep	736	Mach.tools for working metal or met
554	Soap,cleansing and polishing prepar	737	Metal working machinery and parts
562	Fertilizers,manufactured	741	Heating & cooling equipment and par
572	Explosives and pyrotechnic products	742	Pumps for liquids,liq.elevators and
582	Condensation,polycondensation & pol	743	Pumps & compressors,fans & blowers,
583	Polymerization and copolymerization	744	Mechanical handling equip.and parts
584	Regenerated cellulose;cellulose nit	745	Other non-electrical mach.tools,app
585	Other artificial resins and plastic	749	Non-electric parts and accessories
591	Disinfectants,insecticides,fungicid	751	Office machines
592	Starches,inulin & wheat gluten;albu	752	Automatic data processing machines
598	Miscellaneous chemical products,n.e	759	Parts of and accessories suitable f
611	Leather	761	Television receivers
612	Manufactures of leather/of composit	762	Radio-broadcast receivers
613	Furskins,tanned/dressed,pieces/cutt	763	Gramophones,dictating,sound recorde
621	Materials of rubber(e.g.,pastes,pla	764	Telecommunications equipment and pa
625	Rubber tyres,tyre cases,etc.for whe	771	Electric power machinery and parts
628	Articles of rubber,n.e.s.	772	Elect.app.such as switches,relays,f
633	Cork manufactures	773	Equipment for distributing electric
634	Veneers,plywood,improved or reconst	774	Electric apparatus for medical purp
635	Wood manufactures,n.e.s.	775	Household type,elect.& non-electric
641	Paper and paperboard	776	Thermionic,cold & photo-cathode val
642	Paper and paperboard,cut to size or	778	Electrical machinery and apparatus,
651	Textile yarn	781	Passenger motor cars,for transport
652	Cotton fabrics,woven	782	Motor vehicles for transport of goo
653	Fabrics,woven,of man-made fibres	783	Road motor vehicles,n.e.s.
654	Textil.fabrics,woven,oth.than cotto	784	Parts & accessories of 722--,781--,
655	Knitted or crocheted fabrics	785	Motorcycles,motor scooters,invalid
656	Tulle,lace,embroidery,ribbons,& oth	786	Trailers & other vehicles,not motor
657	Special textile fabrics and related	791	Railway vehicles & associated equip
658	Made-up articles,wholly/chiefly of	792	Aircraft & associated equipment and
659	Floor coverings,etc.	793	Ships,boats and floating structures
661	Lime,cement,and fabricated construc	812	Sanitary,plumbing,heating,lighting
662	Clay construct.materials & refracto	821	Furniture and parts thereof
663	Mineral manufactures,n.e.s	831	Travel goods,handbags,brief-cases,p
664	Glass	842	Outer garments,men's,of textile fab
665	Glassware	843	Outer garments,women's,of textile f

666	Pottery	844	Under garments of textile fabrics
667	Pearls,precious& semi-prec.stones,u	845	Outer garments and other articles,k
671	Pig iron,spiegeleisen,sponge iron,i	846	Under garments,knitted or crocheted
672	Ingots and other primary forms,of i	847	Clothing accessories of textile fab
673	Iron and steel bars,rods,angles,sha	848	Art.of apparel & clothing accessori
674	Universals,plates and sheets,of iro	851	Footwear
676	Rails and railway track constructio	871	Optical instruments and apparatus
677	Iron/steel wire,wheth/not coated,bu	872	Medical instruments and appliances
678	Tubes,pipes and fittings,of iron or	873	Meters and counters,n.e.s.
679	Iron & steel castings,forgings & st	874	Measuring,checking,analysing instru
681	Silver,platinum & oth.metals of the	881	Photographic apparatus and equipmen
682	Copper	882	Photographic & cinematographic supp
683	Nickel	883	Cinematograph film,exposed-develope
684	Aluminium	884	Optical goods,n.e.s.
685	Lead	885	Watches and clocks
686	Zinc	892	Printed matter
687	Tin	893	Articles of materials described in
689	Miscell.non-ferrous base metals emp	894	Baby carriages,toys,games and sport
691	Structures & parts of struc.;iron,s	895	Office and stationery supplies,n.e.
692	Metal containers for storage and tr	896	Works of art,collectors pieces & an
693	Wire products and fencing grills	897	Jewellery,goldsmiths and other art.
694	Nails,screws,nuts,bolts etc.of iron	898	Musical instruments,parts and acces
695	Tools for use in hand or in machine	899	Other miscellaneous manufactured ar
696	Cutlery		
697	Household equipment of base metal,n		
699	Manufactures of base metal,n.e.s.		

Table R2. ISIC Rev.2, 3-digit manufacturing sectors' codes and names

311 Food products	354 Misc. petroleum and coal products
313 Beverages	355 Rubber products
314 Tabacco	356 Plastic products
321 Textiles	361 Pottery, china, earthenware
322 Wearing apparel, except footwear	362 Glass and products
323 Leather products	369 Other non-metallic mineral products
324 Footwear, except rubber or plastic	371 Iron and steel
331 Wood products, except furniture	372 Non-ferrous metals
332 Furniture, except metal	381 Fabricated metal products
341 Paper and products	382 Machinery, except electrical
342 Printing and publishing	383 Machinery, electric
351 Industrial chemicals	384 Transport equipment
352 Other chemicals	385 Professional & scientific equipment
353 Petroleum refineries	390 Other manufactured products

Table R3. Explanatory variables' description and data sources

Variables' group	Variable name	Description	Source	Time Span
EconDEV	GDPpc	Gross Domestic Product per capita (const US$ 2000)	PWT 6.2	1985-2004 (annual)
SIZE *Country size*	POP	Total population	PWT 6.2	1985-2004 (annual)
	GDP	Gross Domestic Product (const US$ 2000) GDP = GDPpc*POP	PWT 6.2	1985-2004 (annual)
TECH_HC *Technology Human Capital*	EnrSec_pop	Enrolment in secondary education as % of population	Enrolment from UNdata (primary source: Unesco), population from PWT 6.2	1985-2003 (annual)
	EnrTer_pop	Enrolment in tertiary education as % of population	Enrolment from UNdata (primary source: Unesco), population from PWT 6.2	1985-2003 (annual)
	Illit_Rate	Illiteracy rate (% population aged 15-24)	UNdata (primary source: Unesco)	1985-2004 (annual)
	RDspending	Spending on R&D as % of GDP	Unesco	1997-2004[2]
	Researchers	Number of researchers per mln citizens	Unesco	1997-2004[2]
INST *Institutions*	Gov_size	Government size (government consumption spending, transfers and subsidies, government enterprises and investment, tax rates) Values out of 10 (10 is the highest possible score and zero is the lowest). A higher score indicates a greater degree of economic freedom in the area described by the index.	Fraser Institute (primary sources: World Bank, IMF	1985-2004[1]
	Leg_PropRights	Legal structure and security of property rights (judicial independence, impartial courts, protection of property rights, rule of law, integrity of legal system, legal enforcement of contracts, regulatory restrictions on the sale of real property) Values out of 10 (10 is the highest possible score and zero is the lowest). A higher score indicates a greater degree of economic freedom in the area described by the index.	Fraser Institute (primary sources: World Economic Forum: Global Competitiveness Report, World Bank: Governance Indicators, International Country Risk Guide)	1985-2004[1]
	SoundMoney	Sound money (consistency of monetary policy with long term price stability: money growth, inflation, ease with which other currencies can be used via domestic and foreign bank accounts: freedom to own foreign currency bank account) Values out of 10 (10 is the highest possible score and zero is the lowest). A higher score indicates a greater degree of economic freedom in the area described by the index.	Fraser Institute (primary sources: IMF and World Bank)	1985-2004[1]

	FreeTrade	Freedom to trade internationally (size of restraints that affect international exchange: tariffs, quotas, hidden administrative restraints, size of trade sector relative to expected, black market exchange rates and international capital market controls) Values out of 10 (10 is the highest possible score and zero is the lowest). A higher score indicates a greater degree of economic freedom in the area described by the index.	Fraser Institute (primary sources: World Economic Forum: International Country Risk Guide, Global Competitiveness Report and World Bank: Doing Business data)	1985-2004[1]
	RegCreditBusiness	Regulation of credit, labor and business (freedom in the domestic credit market i.e. private banking sector, presence of foreign banks and controls on interest rates; free labor market regulations i.e. presence of minimum wages, centralised wage setting, union contracts, conscription; freedom of business activities i.e. if regulatory constraints and bureaucratic procedures limit competition and the operation of markets). Values out of 10 (10 is the highest possible score and zero is the lowest). A higher score indicates a greater degree of economic freedom in the area described by the index.	Fraser Institute (primary sources: IMF, World Bank, World Economic Forum: Global Competitiveness Report)	1985-2004[1]
	EconFreedomIndex	Summary Economic Freedom Index; constructed as a chained linked (based on the 2000) summary index incorporating 23 components grouped into five areas: 1 – size of government, 2 – legal structure and security of property rights, 3 – access to sound money, 4 – freedom to trade internationally, 5 - Regulation of credit, labor and business Values out of 10 (10 is the highest possible score and zero is the lowest). A higher score indicates a greater overall degree of economic freedom.	Fraser Institute	1985-2004[1]
	VoiceAccountability	Voice and accountability Values 1.5 to 6.5[3] (higher values correspond to better governance outcomes)	World Bank Governance Indicators	1996, 1998, 2000, 2002, 2003, 2004
	PoliticalStability	Political stability and absence of violence/terrorism Values 1.5 to 6.5[3] (higher values correspond to better governance outcomes)	World Bank Governance Indicators	1996, 1998, 2000, 2002, 2003, 2004
	GovEffectiveness	Government effectiveness Values 1.5 to 6.5[3] (higher values correspond to better governance outcomes)	World Bank Governance Indicators	1996, 1998, 2000, 2002, 2003, 2004
	RegulatoryQuality	Regulatory quality Values 1.5 to 6.5[3] (higher values correspond to better governance outcomes)	World Bank Governance Indicators	1996, 1998, 2000, 2002, 2003, 2004
	RuleOfLaw	Rule of Law Values 1.5 to 6.5[3] (higher values correspond to better governance outcomes)	World Bank Governance Indicators	1996, 1998, 2000, 2002, 2003, 2004

	ControlOfCorruption	Control of corruption Values 1.5 to 6.5[3] (higher values correspond to better governance outcomes)	World Bank Governance Indicators	1996, 1998, 2000, 2002, 2003, 2004
	Gov6index	Summary index of governance (mean of: *VoiceAccountability, PoliticalStability, GovEffectiveness, RegulatoryQuality, RuleOfLaw, ControlOfCorruption*) Values 1.5 to 6.5[3] (higher values correspond to better governance outcomes)	Own (based on partial indices from World Bank)	1996, 1998, 2000, 2002, 2003, 2004
	Gov4index	Summary index of governance (mean of: *GovEffectiveness, RegulatoryQuality, RuleOfLaw, ControlOfCorruption*) Values 1.5 to 6.5[3] (higher values correspond to better governance outcomes)	Own (based on partial indices from World Bank)	1996, 1998, 2000, 2002, 2003, 2004
	Gov3index	Summary index of governance (mean of: *RegulatoryQuality, RuleOfLaw, ControlOfCorruption*) Values 1.5 to 6.5[3] (higher values correspond to better governance outcomes)	Own (based on partial indices from World Bank)	1996, 1998, 2000, 2002, 2003, 2004
	Trust	% of persons in a given country answering "most people can be trusted" out of total answers to the question "Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?" (question A165)	European Values Study Group and Word Values Survey Association (2006)	1985-2004[2]
GEO *Geography*	CoastRiv	% of land within 100 km of the ocean coastline or navigable river	Gallup *et al.* (1999)	1985-2004 (Constant through time)
	Tropics	% of land area in tropical zones[5]	Gallup *et al.* (1999)	1985-2004 (Constant through time)
	MarketDist	Distance (in km) from the closest major market (New York, Rotterdam or Tokyo)	Gallup *et al.* (1999)	1985-2004 (Constant through time)
SPATIAL_CORR	SpatialCorrDist	Spatial correlation of export structures between neighbouring countries (see explanations in text)	Own (based on distances from CEPII)	1985-2004 (annual)
TRADE	Open	Exports plus Imports as % of GDP	PWT 6.2	1985-2004 (annual)
	RTADummy	=1 if active member of selected Regional Trade Agreement: EU (from the moment of accession), EEA, EFTA, NAFTA, MERCOSUR, ASEAN	Own (based on WTO)	1985-2004 (annual)
	ManufExports	% of manufacturing exports (*SITC* Rev.2 products under headings 5-8) in total GDP	Exports from UNComtrade, GDP from PWT 6.2	1985-2004 (annual)
ENDOW *Endowments*	Petrol	*Petrol* – share of petrol and petroleum products (sectors grouped under SITC Rev.2 code 33) in total exports (%).	Own, based on UNComtrade data	1985-2004 (annual)
	AgricVA	Value added in agriculture as % of GDP	Word Bank: WDI 2007	1985-2004 (annual)
	AgricEmpl	Share of employment in agriculture in total employment	Word Bank: WDI 2007	1985-2004 (annual)

[1] Between 1985 and 2000 only 5-year values are available

[2] Only selected years

[3] originally, governance statistics from the World Bank were measured in units ranging from about -2.5 to 2.5, but in order to be able to calculate their logs we have reported them to positive values ('new value' of GOV = original GOV +4)

[5] In case of several countries, the percentage of land in tropical zones is equal to zero, but in order to be able to calculate the logarithm of a *Tropic* variable needed for estimations, we have replaced 0 values with 10^{-10}.

Table R4. GAM estimations (with FE) results, 60 countries, 1985-2004

	Manufacturing exports SITC Rev.2, 3 digit (149 sectors) 1985-2004, 60 countries	
	RelTheil_*SITC*	RelGini_*SITC*
N	1200	1200
Df	2.5	2.5
NPar F	36.26	49.64
P<F	0.0000	0.0000
Country fixed effects	yes	yes

Note: non parametric span =0.5

Figure R1. GAM estimations (with FE) results, 60 countries, 1985-2004
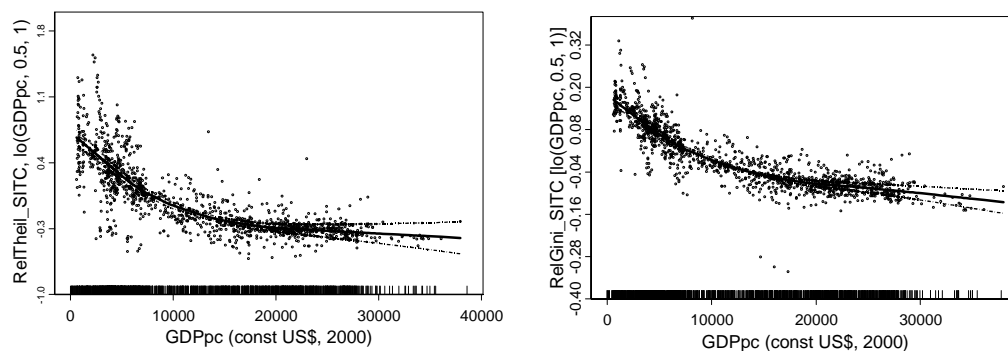
Note: Separate plots correspond to GAM estimations in the table. Horizontal axis: GDP *per capita* (const int.US$), vertical axis: relative export specialization (partial residuals). Variables enter in levels. Diversification measures (*RelTheil*, *RelGini*) calculated with 149 *SITC* Rev.2 ,3 digit sectors headings 5 to 8.

Table R5.1. Results of first stage estimation (diversification and development level with country specific effects and corrected endogeneity) - IV FE. Robustness check: alternative measure of diversification (*RelGini_SITC*)

	Dep.variable: *DIV*		
	RelGini_SITC	*RelGini_SITC*	*ln_RelGini_SITC*
	(1)	(2)	(3)
	IV FE	IV FE	IV FE
GDPpc (in 1000 US$)	-0.005***	-0.023***	
	[-7.634]	[-10.846]	
squared GDPpc		0.000***	
		[8.973]	
ln_GDPpc			-0.188***
			[-14.684]
cons	0.747***	0.874***	1.295***
	[91.574]	[54.179]	[11.055]
N	1140	1140	1140
R2	0.34	0.33	0.29
p_fe	0.000	0.000	0.000
F_exo	**2.33**	**2.16**	5.5
p_exo	**0.127**	**0.116**	0.019

Note: development level instrumented by its lag; t-statistics in parenthesis; *, ** and *** denote significance at 10%, 5% and 1% level, espectively.

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) Source: own elaboration

Table R5.2 Second step estimation results (multivariate estimations – determinants of export diversification), IV and OLS. Robustness check: alternative measure of diversification (*RelGini_SITC*)

	Dep.variable: *DIV (ln_RelGini_SITC)*						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	IV-2SLS	IV-2SLS	IV-2SLS	IV-2SLS	IV-2SLS	OLS	OLS
ln of GDPpc	-0.143***	-0.076***	-0.050***	-0.046***	-0.136***		
	[-20.257]	[-10.598]	[-6.316]	[-5.782]	[-15.844]		
ln of GDP	-0.089***	-0.066***	-0.071***	-0.069***		-0.069***	
	[-26.942]	[-21.267]	[-21.873]	[-20.790]		[-20.550]	
ln of POP					-0.083***		-0.057***
					[-23.701]		[-16.465]
ln of MarketDist		0.123***	0.107***	0.103***	0.094***	0.118***	0.146***
		[18.528]	[16.503]	[15.924]	[14.882]	[19.690]	[24.519]
ln of FreeTrade			-0.361***	-0.328***	-0.321***	-0.390***	-0.480***
			[-10.577]	[-9.309]	[-9.490]	[-12.499]	[-14.200]
RTA				-0.049***	-0.044***	-0.059***	-0.088***
				[-3.616]	[-3.373]	[-4.419]	[-6.274]
cons	2.993***	0.885***	1.610***	1.494***	1.560***	1.158***	-0.095
	[32.019]	[6.315]	[11.149]	[10.157]	[11.344]	[8.727]	[-0.936]
N	1140	1140	1064	1064	1064	1120	1120
R2	0.57	0.67	0.71	0.72	0.74	0.7	0.67
F_exo_wu	9.761	6.895	18.871	18.901	25.074		
p_exo_wu	0.002	0.009	0.000	0.000	0.000		
chi2_exo_durb	9.867	6.993	18.963	19.01	25.072		
p_exo_durb	0.002	0.008	0.000	0.000	0.000		
time dummies	yes	yes	yes	yes	yes	yes	yes

Note: GDPpc instrumented by its lags, t-statistics in parenthesis; *, ** and *** denote significance at 10%, 5% and 1% level, respectively.

GDPpc – real income per capita from PWT 6.2, *GDP* – real Gross Domestic Product from PWT 6.2, *POP* – population from PWT 6.2, *MarketDist*- distance (in km) from the closest major market (New York, Rotterdam or Tokyo) from Gallup *et al.* (1999), *FreeTrade* - Freedom to trade internationally (size of restraints that affect international exchange: tariffs, quotas, hidden administrative restraints, size of trade sector relative to expected, black market exchange rates and international capital market controls) - index from the Fraser Institute (a higher score indicates a greater degree of economic freedom in trade), *RTA*=1 if active member of selected Regional Trade Agreement: EU (from the moment of accession), EEA, EFTA, NAFTA, MERCOSUR, ASEAN (own elaboration, based on WTO).

F_exo_wu and p_exo_wu – refer to Wu-Hausman test of exogeneity (IV)

chi2_exo_durb and p_exo_durb – refer to Durbin test of exogeneity (IV) Source: own elaboration

Table R6.1. Results of first stage estimation (diversification and development level with country specific effects and corrected endogeneity) - IV FE. Robustness check: change in the disaggregation scheme(ISIC rev. 2, 3 digit scheme - 28 manufacturing sectors)

	Dep.variable: *DIV*		
	RelTheil_ISIC	*RelTheil_ISIC*	*ln_RelTheil_ISIC*
	(1)	(2)	(3)
	IV FE	IV FE	IV FE
GDPpc (in 1000 US$)	-0.008***	-0.056***	
	[-3.120]	[-5.931]	
squared GDPpc		0.001***	
		[5.248]	
ln_GDPpc			-0.645***
			[-15.483]
Cons	0.953***	1.284***	5.358***
	[26.735]	[17.807]	[14.054]
N	1140	1140	1140
R2	0.17	0.15	0.23
p_fe	0.000	0.000	0.000
F_exo	1.65	1.12	8.51
p_exo	0.199	0.326	0.004

Note: development level instrumented by its lag; t-statistics in parenthesis; *, ** and *** denote significance at 10%, 5% and 1% level, respectively.
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) Source: own elaboration

Table R6.2 Second step estimation results (multivariate estimations – determinants of export diversification), IV and OLS. Robustness check: change in the disaggregation scheme(ISIC rev. 2, 3 digit scheme - 28 manufacturing sectors)

	Dep.variable: *DIV (ln_RelTheil_ISIC)*						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	IV-2SLS	IV-2SLS	IV-2SLS	IV-2SLS	IV-2SLS	OLS	OLS
ln of GDPpc	-0.361***	-0.133***	-0.03	-0.003	-0.225***		
	[-16.039]	[-5.940]	[-1.300]	[-0.113]	[-9.039]		
ln of GDP	-0.257***	-0.181***	-0.186***	-0.169***		-0.170***	
	[-24.456]	[-18.553]	[-19.370]	[-17.829]		[-18.064]	
ln of POP					-0.206***		-0.159***
					[-20.120]		[-17.090]
ln of MarketDist		0.421***	0.356***	0.333***	0.310***	0.334***	0.397***
		[20.175]	[18.676]	[17.900]	[16.939]	[19.687]	[24.547]
ln of FreeTrade			-1.479***	-1.248***	-1.231***	-1.190***	-1.458***
			[-14.655]	[-12.339]	[-12.538]	[-13.520]	[-15.918]
RTA				-0.335***	-0.323***	-0.332***	-0.393***
				[-8.650]	[-8.574]	[-8.747]	[-10.377]
cons	8.847***	1.650***	4.218***	3.420***	3.585***	3.537***	0.736***
	[29.568]	[3.756]	[9.890]	[8.097]	[8.983]	[9.453]	[2.679]
N	1140	1140	1064	1064	1064	1120	1120
R2	0.5	0.63	0.7	0.72	0.74	0.72	0.71
F_exo_wu	4.851	2.393	15.006	15.652	20.074		
p_exo_wu	0.028	0.122	0.000	0.000	0.000		
chi2_exo_durb	4.925	2.437	15.134	15.791	20.167		
p_exo_durb	0.026	0.118	0.000	0.000	0.000		
time dummies	yes	yes	yes	yes	yes	yes	yes

Note: GDPpc instrumented by its lags, t-statistics in parenthesis; *, ** and *** denote significance at 10%, 5% and 1% level, respectively.
GDPpc – real income per capita from PWT 6.2, *GDP* – real Gross Domestic Product from PWT 6.2, *POP* – population from PWT 6.2, *MarketDist*- distance (in km) from the closest major market (New York, Rotterdam or Tokyo) from Gallup *et al.* (1999), *FreeTrade* - Freedom to trade internationally (size of restraints that affect international exchange: tariffs, quotas, hidden administrative restraints, size of trade sector relative to expected, black market exchange rates and international capital market controls) - index from the Fraser Institute (a higher score indicates a greater degree of economic freedom in trade), *RTA*=1 if active member of selected Regional Trade Agreement: EU (from the moment of accession), EEA, EFTA, NAFTA, MERCOSUR, ASEAN (own elaboration, based on WTO).
F_exo_wu and p_exo_wu – refer to Wu-Hausman test of exogeneity (IV)
chi2_exo_durb and p_exo_durb – refer to Durbin test of exogeneity (IV)
Source: own elaboration

Table R7.1 Results of first stage estimation (diversification and development level with country specific effects and corrected endogeneity) - IV FE. Robustness check: change in the disaggregation scheme(ISIC rev. 2, 3 digit scheme - 28 manufacturing sectors) and in the diversification measure (*RelGini*)

	Dep.variable: *DIV*		
	RelGini_ISIC	*RelGini_ISIC*	*ln_RelGini_ISIC*
	(1)	(2)	(3)
	IV FE	IV FE	IV FE
GDPpc (in 1000 US$)	-0.007***	-0.029***	
	[-9.480]	[-12.260]	
squared GDPpc		0.001***	
		[9.865]	
ln_GDPpc			-0.316***
			[-17.988]
Cons	0.663***	0.822***	2.255***
	[70.931]	[44.766]	[14.039]
N	1140	1140	1140
R2	0.3	0.28	0.26
p_fe	0.000	0.000	0.000
F_exo	3.84	3.39	10.73
p_exo	0.05	0.034	0.001

Note: development level instrumented by its lag; t-statistics in parenthesis; *, ** and *** denote significance at 10%, 5% and 1% level, respectively.
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) Source: own elaboration

Table R7.2. Second step estimation results (multivariate estimations – determinants of export diversification), IV and OLS. Robustness check: change in the disaggregation scheme(ISIC rev. 2, 3 digit scheme - 28 manufacturing sectors) and in the diversification measure (*RelGini*)

	Dep.variable: *DIV (ln_RelGini_ISIC)*						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	IV-2SLS	IV-2SLS	IV-2SLS	IV-2SLS	IV-2SLS	OLS	OLS
ln of GDPpc	-0.178***	-0.071***	-0.032***	-0.022**	-0.122***		
	[-18.113]	[-7.458]	[-3.117]	[-2.161]	[-11.049]		
ln of GDP	-0.115***	-0.079***	-0.083***	-0.077***		-0.077***	
	[-24.910]	[-18.998]	[-19.640]	[-18.208]		[-18.323]	
ln of POP					-0.093***		-0.068***
					[-20.459]		[-16.180]
ln of MarketDist		0.198***	0.173***	0.164***	0.154***	0.171***	0.201***
		[22.385]	[20.619]	[19.883]	[18.957]	[22.605]	[27.441]
ln of FreeTrade			-0.558***	-0.474***	-0.467***	-0.490***	-0.602***
			[-12.596]	[-10.561]	[-10.699]	[-12.511]	[-14.558]
RTA				-0.122***	-0.116***	-0.126***	-0.155***
				[-7.080]	[-6.966]	[-7.447]	[-9.074]
cons	3.716***	0.326*	1.357***	1.067***	1.132***	0.973***	-0.355***
	[28.404]	[1.751]	[7.243]	[5.686]	[6.383]	[5.837]	[-2.858]
N	1140	1140	1064	1064	1064	1120	1120
R2	0.52	0.67	0.72	0.74	0.75	0.73	0.72
F_exo_wu	5.944	3.22	13.886	14.217	18.368		
p_exo_wu	0.015	0.073	0.000	0.000	0.000		
chi2_exo_durb	6.029	3.277	14.02	14.362	18.483		
p_exo_durb	0.014	0.07	0.000	0.000	0.000		
time dummies	yes	yes	yes	yes	yes	yes	yes

Note: GDPpc instrumented by its lags, t-statistics in parenthesis; *, ** and *** denote significance at 10%, 5% and 1% level, respectively.
GDPpc – real income per capita from PWT 6.2, *GDP* – real Gross Domestic Product from PWT 6.2, *POP* – population from PWT 6.2, *MarketDist*- distance (in km) from the closest major market (New York, Rotterdam or Tokyo) from Gallup *et al.* (1999), *FreeTrade* - Freedom to trade internationally (size of restraints that affect international exchange: tariffs, quotas, hidden administrative restraints, size of trade sector relative to expected, black market exchange rates and international capital market controls) - index from the Fraser Institute (a higher score indicates a greater degree of economic freedom in trade), *RTA*=1 if active member of selected Regional Trade Agreement: EU (from the moment of accession), EEA, EFTA, NAFTA, MERCOSUR, ASEAN (own elaboration, based on WTO).
F_exo_wu and p_exo_wu – refer to Wu-Hausman test of exogeneity (IV)
chi2_exo_durb and p_exo_durb – refer to Durbin test of exogeneity (IV) Source: own elaboration

Table R8. Second step estimation results (multivariate estimations – determinants of export diversification), IV and OLS. Robustness check: modification in the estimated second stage model and functional forms of the variables (no logs, quadratic form of GDPpc)

	Dep.variable: *DIV (RelTheil_SITC)*					
	(1)	(2)	(3)	(4)	(5)	(6)
	IV-2SLS	IV-2SLS	IV-2SLS	IV-2SLS	IV-2SLS	IV-2SLS
GDPpc in 1000 US$	-0.041***	-0.030***	-0.017***	-0.014***	-0.024***	-0.076***
	[-15.954]	[-10.845]	[-5.309]	[-4.303]	[-7.439]	[-7.512]
squared GDPpc in 1000 US$						0.002***
						[5.416]
GDP in 10^12 US$	-0.183***	-0.154***	-0.153***	-0.151***		
	[-10.016]	[-8.614]	[-8.971]	[-9.074]		
POP in 10^6					-1.032***	-1.157***
					[-9.856]	[-10.888]
MarketDist in 1000km		0.084***	0.074***	0.060***	0.062***	0.062***
		[9.102]	[8.354]	[6.788]	[7.078]	[7.139]
FreeTrade(index)			-0.201***	-0.161***	-0.159***	-0.147***
			[-9.265]	[-7.335]	[-7.283]	[-6.756]
RTA(dummy)				-0.345***	-0.353***	-0.363***
				[-6.946]	[-7.147]	[-7.413]
Cons	1.962***	1.475***	2.739***	2.594***	2.718***	2.829***
	[18.964]	[13.023]	[16.321]	[15.679]	[16.338]	[17.037]
N	1140	1140	1064	1064	1064	1064
R2	0.29	0.34	0.41	0.44	0.45	0.46
F_exo_wu	7.024	6.681	12.596	13.153	17.126	19.589
p_exo_wu	0.008	0.01	0.000	0.000	0.000	0.000
chi2_exo_durb	7.117	6.778	12.733	13.301	17.254	19.708
p_exo_durb	0.008	0.009	0.000	0.000	0.000	0.000
time dummies	yes	yes	yes	yes	yes	yes

Note: GDPpc instrumented by its lags, t-statistics in parenthesis; *, ** and *** denote significance at 10%, 5% and 1% level, respectively.
GDPpc – real income per capita from PWT 6.2, *GDP* – real Gross Domestic Product from PWT 6.2, *POP* – population from PWT 6.2, *MarketDist*- distance (in km) from the closest major market (New York, Rotterdam or Tokyo) from Gallup *et al.* (1999), *FreeTrade* - Freedom to trade internationally (size of restraints that affect international exchange: tariffs, quotas, hidden administrative restraints, size of trade sector relative to expected, black market exchange rates and international capital market controls) - index from the Fraser Institute (a higher score indicates a greater degree of economic freedom in trade), *RTA*=1 if active member of selected Regional Trade Agreement: EU (from the moment of accession), EEA, EFTA, NAFTA, MERCOSUR, ASEAN (own elaboration, based on WTO).
F_exo_wu and p_exo_wu – refer to Wu-Hausman test of exogeneity (IV)
chi2_exo_durb and p_exo_durb – refer to Durbin test of exogeneity (IV)
Source: own elaboration