Bank i Kredyt 44 (5), 2013, 435-466

BANK I KREDYT
BANK & CREDIT

www.bankikredyt.nbp.pl
www.bankandcredit.nbp.pl

The evolving structure of Polish exports (1994–2010) – diversification of products and trade partners

Aleksandra Parteka*

Submitted: 6 September 2012. Accepted: 21 April 2013.

Abstract

This paper presents empirical evidence on the diversification process concerning Polish exports (1994–2010), compared to European and global samples of countries. It analyzes both the commodity structure of Polish trade and the geographical diversification of Poland's trading partners. The analysis draws on highly disaggregated data on exports (HS 6 digit) and combines descriptive analysis with non-parametric, semi-parametric and parametric estimation models. The results suggest that Poland (exporting 84% of all goods present in the sample) can be placed among countries with well-diversified export products. In terms of geographical diversification, Poland exploits approximately one-fifth of its theoretical overall market reach potential (the best score among new member states) and the diversification of its partner countries increased in the period analyzed. The Polish export portfolio, in terms of the variety of both its products and receiving markets, is more diversified than what is typical for countries at approximately the same stage of economic development.

Keywords: diversification, trade, export

JEL: F14, F41, O11

* Gdansk University of Technology, Faculty of Management and Economics; e-mail: aparteka@zie.pg.gda.pl.

1. Introduction

The main aim of this paper is to present empirical evidence on two aspects of the diversification process concerning Polish exports. It focuses on: (i) changes in the commodity structure of Polish trade (diversification of export products) and (ii) diversification of trade partners (geographical export diversification).

The subject of export diversity is crucial from the point of view of the risk that a country faces in an open economy setting. Not only does export diversification create greater opportunities in regional and global markets and can be a factor of economic growth (Herzer, Danzinger 2006; Hesse 2009), but it is also among the key aspects of risk reduction, especially in the case of less developed countries exposed to external idiosyncratic shocks (Shaw, Newfarmer, Walkenhorst 2009). Product diversification is understood here as the antonym of specialization, which, despite being potentially beneficial in terms of productivity gains and the exploitation of economies of scale, implies a concentration of exports in few product lines. Consequently, countries with highly diversified economic structures are less exposed to shocks and price swings in foreign markets; while economies with less heterogeneous sets of export products are more vulnerable to terms-of--trade shocks (Ghosh, Ostry 1994).[1] Similarly, an excessive geographical concentration of exports (thus a low diversity of trading partners) results in potentially dangerous dependence on few receiving markets and increased income volatility (Bacchetta et al. 2009).

In such a context, Poland is a very interesting case to examine. Since the early 1990s, the country has been notably modifying its trade structure, firstly because of a geographical reorientation of trade due to the collapse of the Soviet Union and the general economic transformation in Central and Eastern Europe (as documented in Zaghini 2005), and later as a result of progressing integration with EU markets. In 2011, 77.8% of Polish exports were directed to the EU and 69.4% of Polish imports came from EU countries (data from Eurostat; EU refers to the EU27). Along with this geographical reorientation of trade, the product diversity of Polish exports has been changing too. Rapid economic growth has gone hand in hand with increasing export capacity and Poland can be placed among the EU countries with well diversified export portfolios covering 84% of all the goods that could theoretically be traded.[2]

The theme of Polish export diversity is also an interesting research subject in the context of integration processes in Europe. Poland has already reached a high degree of economic integration with the EU but still (2013) remains outside the euro area, and in the light of the recent/present problems manifested in the euro zone the eventual prospect of euro adoption in Poland seems to have been postponed. Nevertheless, it is important to provide up-to-date evidence on changes in the structural composition of the Polish economic structure (here, exports) affecting its resistance to external shocks. Most of the research concerning Poland focuses on its structural similarity with the eurozone as one of the crucial components of real convergence and business cycle synchronization (see, among others: Adamowicz et al. 2009; Bruzda 2011; Konopczak 2009;

[1] Di Giovanni and Levchenko (2010) argue that export variety matters – in their view a country's export structure is more risky when its exports are highly undiversified, or when it exports in riskier sectors. They construct a sector--level measure of the riskiness of a country's pattern of export specialization, according to which among 130 countries examined for the 1990s, Poland has a middle-low risk content of exports (being concentrated in relatively less volatile/less risky sectors: "Wearing apparel, except footwear" and "Food products").

[2] Data based on HS0 6 digit products present in our sample (4,963 export lines), average 1994−2010.

Konopczak, Marczewski 2011; Koźluk 2005). At the same time, empirical studies on the evolving diversity of Polish exports in recent years, concerning both export products and export partners, in particular analyzing highly disaggregated data, are missing.[3] Consequently, this paper aims to fill this research gap. Its empirical analysis is based on a very detailed product-level database (HS0 6 digit: almost 5,000 product lines) containing statistics on Polish exports to all the world economies. The trends concerning Poland are observed across the years 1994–2010 (due to data availability) and, wherever possible, are assessed in a comparative setting with respect to overall global trends and to the EU countries in particular.

The rest of the paper is organized as follows. Section 2 presents literature related to the subject of trade diversification analysis. Section 3 outlines the analytical approach adopted in this paper, presenting the data and methodology used. Section 4, the core of the paper, presents the empirical results, focusing on patterns of diversification in the Polish export structure assessed in a comparative setting. Finally, the fifth Section summarizes and concludes. The main findings show that Polish exports are characterized by a relatively high degree of diversification – both in terms of product heterogeneity and the geographical diversity of the country's trading partners.

2. Related literature on trade diversification

From the theoretical point of view, references to trade diversity can be found mainly in the so-called "new trade theory" literature (Krugman 1979a; 1979b; 1980; 1981; Helpman, Krugman 1985). This uses monopolistic competition models (assuming a market with a large number of firms, each producing a unique variety of a differentiated product) and is built on the "love of variety" concept: consumer utility (modelled according to Dixit-Stiglitz framework) increases with the number of available product varieties. The number of goods depends on the size of the market and trade acts as a natural way of market extension. It has been shown that gains from trade can occur even if it takes place between similar countries (Krugman 1979a), which is contrary to the "traditional" view explaining trade as a result of differences in technology (Ricardo) or factor endowments (Heckser-Ohlin) across countries. Newer models of trade have relaxed some of the assumptions of Krugman's workhorse view. In particular, the seminal model by Melitz (2003) introduced heterogeneity in firm productivity, starting "new-new trade theory". Melitz shows that even though exposure to trade forces less productive firms to exit the market, an increase in the number of new foreign exporters results in a net welfare gain in terms of greater product variety available to the consumer.

Increase in product diversity is also a key notion in endogenous growth models, especially those on expanding product variety (e.g. Grossman, Helpman 1991 – Chapters 3, 8, 9). The focus here is on the production process, where a rising variety of differentiated inputs results in an increase in output (modelled through the CES production function), in the same way as a rise in the variety of final goods increases total consumer utility in "love-of-variety" models of trade according to Krugman framework.

[3] Brodzicki (2009) provides an indicator of the relative product variety in Polish manufacturing industry based on product data in NACE rev1 groups, but his analysis is limited to the period 1999–2006 and concerns only Polish manufacturing exports to the EU15. The geographical aspect of export diversification is not examined.

However, given the applied aspect of the present research, henceforth we will centre our attention on empirical evidence on the changing composition of trade structures. The international empirical literature has focused mainly on the following three aspects: (i) the evolution of diversification along the path of economic growth; (ii) margins of trade expansion/diversification; (iii) the determinants of the diversification process.

Turning to the first aspect, empirical studies on the relationship between diversification of economic structures and income *per capita* levels begin with the seminal contribution by Imbs and Wacziarg (2003). They found, using sectoral employment and value added data, that countries appear to follow "stages of diversification": an initial economic growth process can be associated with a progressive diversification of economic structures, while at higher stages of growth countries re-concentrate. Given the use of inverse measures of diversification (based on inequality/concentration indices) such a pattern is illustrated by a U-shaped curve. Cadot et al. (2011) found a similar result for exports. This result holds when conventional measures of concentration such as the Herfindahl index, Theil index or Gini index are used as inverse measures of economic structure diversity. A diversification pattern, without any evident reverse of the trend at higher levels of economic development has been found in alternative diversification/income *per capita* studies, adopting relative measures. These compare each country's product diversity with respect to a common benchmark in the whole sample (de Benedictis, Gallegati, Tamberi 2009; Parteka 2010). In a wide panel data setting (156 countries, 1988–2006) and with the use of very detailed export statistics (almost 5,000 product lines) they reveal a hump-shaped pattern (diversification followed by product re-concentration) with a turning point at approximately 25,000 USD (2005 PPP). Parteka (2013), providing evidence based on import and export data (the same level of disaggregation as in Cadot et al. 2011), confirms such a pattern in a European sample: in line with the "stages of diversification" approach of Imbs and Wacziarg (2003), EU27 countries are characterized by a high degree of trade diversity and there is a possibility of re-concentration at higher stages of development (observable in non-parametric estimates). These results are mainly based on the estimation of non-parametric lowess curves (see Section 4 for a description of the relative methodology). Misztal (2011) performs a similar analysis linking export concentration measures with economic growth in European Union countries (1995–2009). He reveals a "W" shape of the lowess curve linking the degree of export concentration and GDP *per capita*. However, this somewhat surprising result can be driven by the presence of outliers (Malta, Cyprus, Luxembourg) in a relatively small overall sample used to plot non-parametric curves (27 observations). For comparison, to obtain lowess plots, Cadot et al. (2011) employ approximately 2,500 observations in an international sample while Parteka (2013) uses 484 observations in an EU27 sample.

Even though a focus on product structure has dominated the literature on diversification and the economic growth nexus (see Cadot et al. 2012 for a thorough survey), the geographical aspect of trade diversification has also been analyzed. Bacchetta et al. (2009) reveal that export diversification increases with level of development, not only in terms of product heterogeneity but also with respect to geographical diversification. Additionally they find (in a sample of 180 countries, 1960–2004) that the increase in the number of trading partners was more significant than the increase in the number of products exported, and such geographical diversification helps to buffer external shocks and reduces the transmission of external volatility to the exporting economy. Reis and Farole (2012, p. 43) confirm that the ability of low-income countries to reach foreign markets (measured by an index of export market penetration – see Section 3.2) is far below that of high-income countries.[4]

A different aspect of the geographical diversity of trade, concerning imports, is examined by Cadot et al. (2010). They focus on the geographical concentration of OECD imports across countries of origin (i.e. diversity of suppliers), and find that a diversification period was followed by re-concentration after the year 2000 (mainly caused by the increasing importance of Chinese products in OECD imports). Evidence for a progressive diversification of import sources along the path of economic growth can be found in an international panel data study (1962–2000, 160 importers) by Jaimovich (2012).

Matching the product-related and geographical aspects of trade diversification, another strand of the literature explicitly explores the role played by intensive and extensive margins of trade. The former concerns existing trade flows (changes in diversification among the set of products that are commonly traded over the period of analysis) while the latter refers to new export flows (new exports of existing products to new markets, new exports of new products to existing markets and new exports of new products to new markets). Even though empirical evidence suggests that export growth is driven mainly by activity at the intensive margin (Evenett, Venables 2002; Brenton, Newfarmer 2007),[5] the extensive margin remains important in the case of developing countries. These tend to expand exports by adding new product lines to their export portfolios, so decreasing export concentration for lower income countries results mainly from a rise in the number of goods exported (Cadot et al. 2011).

Finally, several factors have been examined as potential drivers of diversification, and similar features seem to promote both product-related and geographical diversification of exports. Gravity-type variables, such as market size and distance from the main markets, have been found to be among the most relevant features affecting the product diversity of export structures (Amurgo-Pacheco, Pierola 2008; Parteka, Tamberi 2013). In the same way, the probability of exporting to a given destination, influencing new markets margin of trade (thus the geographical diversification of exports), decreases with distance and increases with market size (Evenett, Venables 2002). Additionally, Shepherd (2010) shows the importance of export costs, tariffs and international transport costs as determinants of geographical export diversification, estimating that a 10% reduction in any of these factors in the case of developing countries can be related to as much as a 5–6% increase in the number of destination markets for their exports.

As far as the literature referring to the specific case of Poland is concerned, most recent studies on changes in its economic structure have focused on its similarity with the euro zone, assessed in the context of monetary integration and the necessary structural and business cycle synchronization.

Studies analyzing the period before the financial crisis (Fidrmuc, Korhonen 2006; Adamowicz et al. 2009; Skrzypczyński 2009; Konopczak 2009) show that of Poland's degree of business cycle synchrony with the euro area was one of the highest among the CEECs. However, in the period of the crisis, which, apart from the EU15 economies, also strongly hit countries from the region, the reaction of the Polish economy was surprisingly good. Konopczak and Marczewski (2011) analyze the reasons for such a relatively good performance of the Polish economy during the crisis and

[4] In fact, no country is able to export all its exports to all the countries that import them. Germany serves as a benchmark, exploiting in 2008 around 50% of all potential product-market relationships; in comparison, Albania exploited only 2% of its potential (Reis, Farole 2012, p. 43).

[5] In contrast, Hummels and Klenow (2005) find a greater importance of the extensive rather than the intensive margin of trade when correcting measures of the margins for differences in the economic importance of exported goods.

the cyclical divergence of Poland from the euro area. They argue that with respect to other CEECs Poland was characterized by a lower responsiveness of output to foreign shocks and a lower share of those shocks in structural impulses, combined with a dichotomy in internal activity resulting from differential structural characteristics. In particular, the latter factor may suggest that the relatively high degree of synchronization of the Polish economy with the euro area might have diminished.

Hence, although they are important from the policy point of view, they are not strictly linked to the subject of the present research. Along these lines, de Benedictis and Tajoli (2007) look at the similarity of trade structures toward the EU market between four CEECs (including Poland) and the EU15 between 1989 and 2001. Concerning self-similarity, they find that Poland started moving away from its initial (1989) export specialization pattern, but at the same time it continued to converge towards the EU trade structure, especially after 1994.

Turning back to the literature on trade diversity, the Polish case has not been thoroughly analyzed and existing evidence is somewhat limited in terms of country focus, product/partner details and the time dimension. The closest study to our research topic performed by other authors is the one by Shepotylo (2012), who includes Poland in his sample of countries from Eastern Europe (EE) and the Commonwealth of Independent States (CIS), comparing actual and predicted (by gravity equations) levels of product and geographical diversification in the years 2001–2007. Poland appears to have more diversified exports (in both dimensions) than the predicted levels and, along with the Czech Republic, emerges as having reached a degree of export diversification comparable with the export patterns of the EU15. However, the analysis by Shepotylo (2012) is based on sectoral data (10 SITC 1-digit sectors), which does not allow for a detailed study of changes in export diversity manifested at a much more disaggregated level. Evidence on export variety in 14 Eastern European countries based on more detailed trade data (1,473 commodities) but for an earlier period (1993–2000) provided by Funke and Ruheweld (2005) confirms that the diversity of products exported by Poland is among the highest in the group analyzed. Poland is also included in Cadot et al. (2011) panel in the aforementioned study on export diversification performed with highly disaggregated product-level data, but no explicit evidence concerning the country is provided. Given their estimated turning point (25,000 USD in 2005 PPP) we can only deduce that Poland should be located on the decreasing part of the U curve, corresponding to a progressive product diversification of exports along the path of economic growth. This would be in line with the findings of Fertó and Soós (2008), whose analysis of specialization patterns in Eastern European countries based on the Balassa index of RCA (revealed comparative advantage) confirms that, as with most CEECs, Poland's trade specialization dropped radically between 1995 and 2002. Brodzicki (2009) also finds that the product variety (measured as simple product counts) of Polish exports to the EU15 increased in the years 1999–2003.

3. Empirical setting – data and methodology

3.1. Data

Our analysis draws on two separate databases: the first, used to measure the product diversity of exports, is three-dimensional (reporter, product, time) and contains information on export flows from as many countries as possible (including Poland) to the world.[6] The second database has four dimensions (reporter, partner, product, time) and is employed to analyze the geographical aspect of the diversity of Polish exports. In both cases, we use annual trade data from UNComtrade (accessed through on-line software WITS – world integrated trade solutions; https://wits.worldbank.org/WITS), expressed in current 1,000 USD and classified according to the HS (harmonized system). In order to avoid problems with lower quality direct export statistics, we employ mirrored data. We use the deepest level of product detail possible for international comparisons: the six-digit HS0 code system (subheadings).

Since the introduction of the HS system (1988) there have been three major revisions: HS1 – 1996, HS2 – 2002 and HS3 – 2007. In order to obtain consistent long time data series we use automatic conversion of various revisions into the basic HS0 one (based on product-level conversion tables from WITS). Disaggregation series deeper than six digits are not harmonized across countries. Rough trade data reported by UNComtrade does not contain the information on every product line for every country-year pair, so in line with Cadot et al. (2011) we fill in the database, adding missing product lines and assign them zero trade values.

Product detail is impressive as, for instance, we are able to distinguish between "Coffee, not roasted, decaffeinated" (code 090111) and "Coffee, not roasted, not decaffeinated" (code 090112). Theoretically, the HS0 classification consists of 5016 product lines but we exclude from the sample 53 HS0 codes which correspond to "silent" (never traded in the period analyzed) product lines and in the end we observe 4,963 product lines (merchandise goods).

Using the first dataset we calculate synthetic measures of product diversity (see Section 3.2) for Poland and, for comparison, other countries. The final set of 163 countries is based on the joint availability of disaggregated trade statistics and real income *per capita* data needed for the analysis. We use GDP *per capita* expressed in PPP terms in constant 2005 international USD from the World Bank's World Development Indicators 2011, in a few cases completed with data from the Penn World Table 7.0. In the end, we make use of an unbalanced panel, summarized in Table 1 in the Appendix, with 1,905 observations on 163 countries,[7] including EU27 economies (listed in Table 2). The actual number of observations used in the regressions is slightly lower due to the exclusion of extreme values and outliers. On a variable-by-variable basis we remove observations below the 1st or above the 99th percentile. In the case of regression analysis, only observations falling into the window between the 1st and 99th percentiles (for all variables) are retained in the sample. We have also eliminated from the final sample those country-year pairs for which more than 15% of the total trade would not be taken into account due to imperfect matches of products between the various revisions of the HS.

The second dataset, which is four-dimensional, is used to measure the degree of geographical diversification of Polish exports. We first download the data on bilateral exports (with Poland as reporter and all available foreign countries as partners), also classified by product (retaining the same set of 4,963 HS0 product lines as before). This is a huge set of statistics containing 84,371 observations (4,963 products observed across 17 years) and 239 separate variables on the value of exports from Poland to each of 239 potential partner countries. With these data we calculate

[6] Similar export data, concerning EU countries and additionally enriched by import statistics, has been used in Parteka (2013).

[7] The full list of countries is available upon request.

synthetic measures of the diversity of Poland's trade partners (defined in Section 3.2). Given computational limits, we are not able to construct analogous measures for all other countries. However, in order to show trends concerning the geographical diversification of Polish exports in a comparative setting, we rely on the inbuilt WITS tool, which permits us to retrieve indices of export market penetration (defined in Section 3.2), also calculated with HS0 6 digit mirrored export data, for all 163 countries in the broad sample.

In the analysis we take into account a change in the methodology of data collection concerning Polish trade flows. This took place in 2004, resulting in a break in the series 2003–2004 in the second (four-dimensional) dataset.

"Since 1992 till the end of April 2004 Single Administrative Document – customs declaration SAD was the only source of data on foreign trade statistics. Since 1 May 2004, as a result of Polish accession to the European Union, foreign trade statistics is based on two sources of information: INTRASTAT declaration – only for recording of arrivals and dispatches in the framework of intra-UE trade; Single Administrative Document – only for recording of Polish trade turnover with non-member countries (co called third countries) in the framework of EXTRASTAT system." (GUS 2005, p. 4). In very disaggregated bilateral trade data, this change results in the appearance of very small trade values (which previously may not have been registered) from 2004 onwards. This might affect the simple count of partner countries by product but an appropriate use of thresholds and weighting schemes can help. UNComtrade statistics and trade indices reported in WITS do not account for the 2003–2004 bias in Polish data. Despite having contacted official representatives from the statistical office we have not managed to receive any satisfactory reply to our enquiry. We have verified that the change in data collection methodology is not a significant problem in the case of analysis based on data aggregated by partners or by product groups.

The break proves to be important especially in the case of bilateral trade flows used for the calculation of indices of geographical diversification based on the product count, but can be corrected by the appropriate use of thresholds and weighting schemes.[8]

The complete disaggregated trade statistics needed for both parts of the analysis (on product and geographical diversification) are available for Poland for the years 1994–2010, which determines the time span of our analysis. The international sample of countries is observed across the years 1988–2010 (unbalanced panel).

3.2. Synthetic measures of export diversification

Product diversity

The simplest way to measure the degree of product differentiation of a country's exports is to count the number of active product lines *(N_prod)* in its export portfolio and observe whether it grows (a sign of progressive product diversification) or collapses (ongoing commodity specialization/concentration) over time. Consequently,

$$N\_prod_{it} = \sum_{k=1}^{n} \mu_{ikt} \tag{1}$$

[8] We would like to thank an anonymous referee for drawing our attention to data issues and pointing out this problem.

where i stands for the reporting country, k for the product and t for time, while $\mu = \{0, 1\}$ serves as a binary product identification dummy, denoting an active or inactive product line ($\mu_{ikt} = 1 \Leftrightarrow x_{ikt} \neq 0$, with x referring to the export value).

However, the measure defined in (1) gives the same importance to all exported goods independently of the actual value of trade. Consequently, other synthetic measures are usually used – typically ones based on classical concentration/inequality indices (e.g. the Herfindahl index, Theil index, or Gini index), which reflect how different the distribution of product shares is from a uniform distribution (i.e. when each product has an equal share in the total economy of a given country).[9] In other words, the values of such indices are positively related to the degree of product concentration and inversely related to the degree of diversification of goods traded by the country under analysis.

The Herfindahl-Hirschman index is the sum of the squared shares of each product in a country's total exports, and in the normalized version (which accounts for the number of products that could actually be exported) is expressed as:

$$HH\_prod_{it} = \frac{\sum_{k=1}^{n} \left(\frac{x_{ikt}}{X_{it}} \right)^2 - \frac{1}{n}}{1 - \frac{1}{n}} \tag{2}$$

where x_{ik} is export value of product $k = 1,\ldots, n$ in i-th country (so that n refers to the number of potential export lines) at time t and X_t denotes country i total exports. In our case, $n = 4{,}963$. A country with a perfectly diversified export portfolio will have *HH_prod* close to zero, whereas a country which exports only one product will have a value of 1 (extremely concentrated product structure).

Alternatively, we can employ the absolute Theil index of product concentration, which is defined as:

$$Theil\_prod_{it} = \frac{1}{n} \sum_{k=1}^{n} \left(\frac{x_{ikt}}{\overline{\overline{x_{it}}}} \cdot \ln \frac{x_{ikt}}{\overline{\overline{x_{it}}}} \right) \tag{3}$$

where i refers to countries, $k = 1,\ldots, n$ to products, x stands for exports, and $\overline{x_{it}} = \frac{\sum_{k=1}^{n} x_{ikt}}{n}$ is the average (across products) export value.

The lower bound of Theil indices is 0 (and corresponds to maximum commodity diversification) while the upper limit is equal to $\ln(n)$, signalling maximum product concentration (minimum

[9] Alternatively, so-called "relative" measures of specialization (e.g. the dissimilarity index, relative Gini index, or relative Theil index) can be used. These are constructed in a way that allows measurement of a country's degree of economic activity dispersion across various sectors (products) with respect to a common benchmark (usually, other countries in the sample). For an application of this kind of measure to the study of trade diversity, see: de Benedictis, Gallegati, Tamberi (2009), or Parteka (2010).

diversity) of the export portfolio. Note that the indices defined in (2) and (3) serve as inverse measures of export diversification. Summary statistics of all the indices from this section are included in Table 3.

Geographical diversity

In the first instance, as a measure of geographical export diversification we count, for each product line, the number of partner countries (*N_geo*) reported by country *i* (here *i* = *PL*):

$$N\_geo_{PLkt} = \sum_{p=1}^{P} \lambda_{PLkpt} \tag{4}$$

where *PL* stands for Poland as reporting country, *p* for partner country, *k* for product and *t* for time, while $\lambda = \{0, 1\}$ serves as a binary partner-product-year identification dummy, denoting (separately for each good *k* and time period *t*) an active or inactive partner country.

A rise in the average (across products) value of *N_geo* would be a sign of a progressive geographical diversification of exports.

Alternatively, for each product line and time period, we can construct a Herfindahl-Hirschman index of geographical export concentration (similar to that defined in eq. (2)), which is the sum of squared shares of exports to a particular partner country *p* out of total Polish exports of good *k*.[10] In the normalized version (to account for the number of partners where goods could potentially be exported) it is expressed as:

$$HH\_geo_{PLkt} = \frac{\sum_{p=1}^{P_k} \left(\frac{x_{PLkpt}}{X_{PLkt}} \right)^2 - \frac{1}{P_k}}{1 - \frac{1}{P_k}} \tag{5}$$

where x_{PLkpt} is the Polish export value of product $k = 1,\ldots, n$ to partner country $p = 1,\ldots, P_k$ (so that P_k refers to the number of potential partner countries and is product specific[11]) at time *t*, and X_{PLkt} denotes the total Polish exports of good *k* at time *t*. In addition, here we adopt a cut-off point for x_{PLkpt} of 10,000 USD.

In the case of exports perfectly diversified across partner countries we would obtain a geographical Herfindahl-Hirschman index close to zero, whereas a value equal to 1 would indicate an extremely concentrated structure of the receiving markets (exports of good *k* limited to one external market only).

In order to correct the bias resulting from the change in data collection methodology (see remark on p. 422) and following Reis and Farole (2012, p. 60), we set a threshold for minimum export values considered significant when calculating (4) and (5). In the first instance, instead of any export above 0 USD, only exports above 10,000 USD are considered ($\lambda_{PLkpt} = 1 \Leftrightarrow x_{PLkpt} > 10{,}000$), with x_{PLkpt} referring to the export value of product *k* from *PL* to country *p*. Alternatively, we use another cut-off, based on the exclusion of observations with a negligible share of product-partner lines out

[10] A similar measure, for imports, is adopted by Cadot et al. (2010).

[11] In this definition our set of potential partner countries (receiving markets) for Polish exports, P_k, is time invariant but product specific; for each product line *k* we calculate the average number of countries to which good *k* was exported from Poland over the years 1994–2010.

of total exports of a particular good ($\lambda_{PLkpt} = 1 \Leftrightarrow s_{PLkpt} > 0.5\%$). For example, imagine a product k for which Polish exports to partner p are of negligible value. Before 2004 they might not have been registered, while after the change in the methodology of data collection, they are suddenly visible in the statistics. Consequently, if any exports > 0 were taken into account, such a product line k exported to partner p would be seen as a "new" export line (even though it is not new) and a rapid rise in *N_geo* between 2003 and 2004 would constitute a pure statistical bias. Additionally, in order to give less importance to export lines of negligible value, when reporting the average values of N_geo_{PLkt} and HH_geo_{PLkt} over all products, we adopt a weighting scheme (with weights corresponding to the share of each product line out of total Polish exports).

Finally, as a last measure of geographical diversification, we use the index of export market penetration (*IEMP*) from WITS, which can be interpreted as an indicator of market reach success (Reis, Farole 2012, p. 43). *IEMP* looks at a country's total number of exports and the number of markets that each of these products reaches. Then, the number of countries in the rest of the world that import each of the products (which the country of interest exports) is counted. Matching these two pieces of information, we get the maximum potential number of export relationships that a country can establish given its export portfolio at present. The actual number of export relationships is then divided by the potential number to assess how many export opportunities a country is exploiting. The formal definition is given by the following formula, based on Brenton and Newfamer (2009, p. 123):

$$IEMP_{it} = \frac{\sum_{k \in K_{kit}} \sum_{p} \lambda_{ikpt}}{\sum_{k \in K_{kit}} \sum_{p} Z_{kpt}}$$

where i refers to the reporter country (exporter), k to the product, p to the partner country (importer) and t to time. K_{kit} is the set of products k in which positive exports from country i are observed at time t; $\lambda_{ikpt} = 1 \Leftrightarrow x_{ikpt} > 0$, otherwise $\lambda_{ikpt} = 0$; $Z_{kpt} = 1 \Leftrightarrow m_{kpt} > 0$, otherwise $Z_{kpt} = 0$; with m_{kp} denoting imports of product k by importer p.[12]

Consequently, *IEMP* is the share of the actual number of export relationships forged by country i out of the maximum possible number of export relationships it can form given the number of its exports. Reliable values of *IEMP* for Poland are reported from 2003 onwards.[13]

Summary statistics of the indices from this section are reported in Table 4.

4. Results

4.1. Product diversity of Polish exports

Table 5 presents a set of alternative indicators of exported product diversity, with values for Poland, and – for comparison – for all the countries in our international sample (divided into two income

[12] Unfortunately, as *IEMP* is retrieved from the WITS system, we do not have the possibility of adopting the cut-off point for exports used for its calculation, as in (4) and (5).

[13] Due to the break in the series of disaggregated product-partners 2003–2004 (not corrected in the WITS trade indicators), values of *IEMP* for Poland before 2003 cannot be directly compared with those after 2003.

groups) and for European countries only (EU27 listed in Table 2). It can be seen that the degree of product variety in the Polish export basket is relatively high – this is reflected in a higher number of products exported (4,176 out of the 4,963 in our set of HS0 product lines) than the international and European averages (2,701 and 3,864, respectively). Average *N_prod* for Poland is high even when compared to the typical value for developed countries (3,691). Similarly, the values of product concentration indices (*HH_prod* and *Theil_prod*) typical for Poland are lower than the benchmark ones. Altogether, these are signs that, from the point of view of the commodity structure, Polish exports are highly diversified.

In order to situate Poland more precisely with respect to other countries, in Figure 1 and Figure 2 we show a comparison between the product diversity of Polish exports and of single countries from the EU sample, referring to the years 1994 and 2010 (the first and last years for which we have the data for Poland). Indices are expressed with respect to the EU27 average in order to account for any possible variation in the general structure of European trade.

Both in terms of the number of products exported (Figure 1) and the synthetic measure of product concentration (Figure 2), Poland can be placed among the EU countries with well diversified export baskets (high *N_geo* and low *Theil_prod*). There is a huge variation across the EU countries. In 2010, Malta (exporting only 1,314 out of the 4,963 products in our HS0 set which could theoretically be traded) had the least diversified range of exports (*N_prod* [EU27 = 100%] = 42%). In comparison, the number of active lines in the most diversified country – Germany (with 4,259 active export lines) – was 14% higher than the EU27 average. The cross-country disparity in terms of the Theil index (Figure 2) is even more pronounced.

Concerning the change which took place in the period analyzed, between 1994 and 2010 the differences between the less and more diversified EU countries decreased. However, Poland moved even further up the ranking towards those EU countries which are characterized by a highly diversified structure of products exported. In 2010 it exported 4,071 goods and its *N_geo* indicator was 9% above the EU average (7% above the EU average in 1994).

The relative rise in the product diversity of Polish exports (with respect to the EU27 average) is visible in Figure 3, where we plot the Polish *N_prod* and *Theil_prod* against time. A major change concerns the concentration index (it dropped from 95% of the EU27 average in 1994 to 81% in 2010; values of *Theil_prod* below 100% confirm that Polish exports are less concentrated/ more diversified than the average in the EU27 sample). At the same time, the set of products exported (on average 7–10% more heterogeneous from Poland than the EU27 average) remained relatively more stable with respect to the typical European basket. This suggests that the process of diversification of Polish exports mainly concerned activity across products (a more even spread across already active export lines) rather than the addition of new active product lines (low activity at the extensive margin).

Export diversity has been found to be strongly linked to income *per capita* levels (Cadot et al. 2011; Cadot et al. 2012) Consequently, following the recent interest in the non-linear relationship between the diversification process and economic development, in Figure 4 we show a lowess representation of non-parametric curves with income *per capita* as a single covariate. These are obtained from all the country-year observations in our panel (163 countries including Poland: a sample broader than the European one is used in order to assure maximum variability of income *per capita*) and correspond to the following model:

$$Y\_prod_{it} = \alpha + s(GDPpc_{it}) + \varepsilon_{it} \tag{6}$$

where $Y\_prod = \{N\_prod, HH\_prod, Theil\_prod\}$ denotes one of the synthetic measures of product diversification defined in Section 3.2, *i* refers to countries and *t* to the time period. *GDPpc* is a proxy of the development level (income *per capita* in real terms), while *s*(.) is an unspecified smooth function estimated through use of the lowess smoother (Cleveland 1979) and represented graphically.

This method is particularly useful when the standard linear model should not be applied as a result of a built-in non-linearity in the data (in the present context, such non-linearity in the diversification-development relationship has been revealed by Imbs and Wacziarg 2003 and Cadot et al. 2011). Lowess stands for "locally weighted scatterplot smoother", as it fits local polynomial regressions and joins them together in order to obtain the non-parametric curve revealed by the data.

For each data observation only the number of nearest neighbours, defined by span, is used. The weight function gives the greatest weight to observations that are closest to the observation currently examined (x_o); in practice the tricube weight function is used. Then, a polynomial regression using weighted least squares is employed to calculate the fitted value for x_o and plot it on the scatterplot. The fitted values are then connected and result in a non-parametric lowess curve, approximating the relationship between the two variables of interest. For formal details concerning the lowess procedure, see Cleveland (1979).

The plots in Figure 4 indeed reveal a non-linear trend of export diversification along the path of economic growth in the sample of international economies[14] – up to income *per capita* levels of approximately 33,000 USD (PPP 2005) the number of products exported tends to rise (plot A), while measures of product concentration (*HH_prod* and *Theil_prod*) decrease (plot B and plot C). At higher stages of economic development a re-concentration track can be observed, which is in line with the findings of Klinger and Lederman (2006) and Cadot et al. (2011), who explain it in the light of movement across diversification cones, as discussed by Schott (2004).

Focusing on Poland (marked with red dots in Figure 4) and its relative position *vis-à-vis* the other countries from the international sample, it can be seen that its level of export product diversity is indeed high. Poland is situated above the lowess curve obtained with *N_prod* (plot A) and below the curves revealed with *HH_prod* and *Theil_prod* (plots B and C, respectively), which means that its export portfolio is relatively more diversified than that typical for countries at approximately the same stage of economic development.

Undoubtedly, product diversity depends on country size, which has been found to be one of the main determinants of export diversification processes (Amurgo-Pacheco, Pierola 2008; Parteka, Tamberi 2013). Bigger countries (and Poland can be considered as such, compared to many small EU economies) have greater opportunities to produce (and export) a greater variety of goods, mainly due to their more heterogeneous sets of available resources. This is confirmed both in our international set of countries and in the European subsample; in Table 6 we report the correlation

[14] We have checked the robustness of this finding using alternative values of the span parameter.

coefficients between measures of export product diversity and GDP and population, used as proxies for country size. These are all significant and have the expected signs (positive when *N_prod* is used, negative when *HH_prod* or *Theil_prod* are employed). Concerning the strength of the relationship, the correlations are higher in the case of the EU27 sample (composed of much more similar economies, at least in terms of stage of development or location). In the international set of countries, factors other than size (such as distance from main markets, trade costs) are likely to play a greater role in shaping diversification opportunities.[15]

Consequently, we extend eq. (6), which is a model with a single covariate, $Y = s(X) + \varepsilon$, towards the additive model with *p* covariates: $Y = \sum_1^p s_j(X_j) + \varepsilon$. In particular, we use the class of models proposed by Hastie and Tibshirani (1986) called generalized additive models (GAM).[16] In our case, we use a semi-parametric GAM to fit the model.[17] This is a mixture of non-linear (income *per capita*) and linear (GDP) components. It allows us to correct the shape of the lowess smoother obtained from estimation of model (6) with the importance of the additional covariate, the country size:

$$Y\_prod_{it} = \alpha + s(GDPpc_{it}) + \beta GDP_{it} + \varepsilon_{it} \tag{7}$$

where all the notation is as in eq. (6) and *GDP* denotes real GDP (in USD, 2005).

It is clearly visible that the GAM plots of the partial residuals (Figure 5) are smoother than the unconditional lowess curves (Figure 4), but the general conclusions hold: (i) the initial phase of economic development is associated with a tendency towards export diversification, and (ii) even after taking into account the country size, the Polish export portfolio is relatively more diversified than ones typical for countries at approximately the same levels of income *per capita*.

4.2. Geographical diversity of Polish exports

Figure 6 shows the evolution of the weighted averages (calculated over all 4,963 product lines in the sample) of our two basic indicators of the geographical diversification of Polish exports: the number of receiving markets (*N_geo*) and the Herfindahl-Hirschman index of geographical export concentration (*HH_geo*). Both of these indices are originally calculated separately for each product *k* (eq. (4) and eq. (5)), so in order to obtain weighted averages across export lines we account for the share of each product line out of the total of Polish exports in a particular year. In order to eliminate product-partner lines of negligible value, two alternative thresholds have been adopted (as described in Section 3.2). Comparing plot A and plot B, we note that a more restrictive cut-off point (only exports with *s_ikj* > 0.5%; plot B) results in a drop in *Avg_N_geo* but the trends over time are very similar whichever threshold is used.

[15] Parteka and Tamberi (2013, p. 15) find that, *ceteris paribus*, an increase in distance from major international markets of 1% can be associated with a decrease in the degree of export diversification of approximately 0.2–0.3%.

[16] A similar procedure, employing GAM in empirical studies on economic diversification is adopted by de Benedictis, Gallegati, Tamberi (2009) and Parteka (2010).

[17] We use the gam and gamplot modules in STATA12 (Royston, Ambler 2012). Alternatively, tools in *R* can be used (Wood 2006).

The export diversification trend is visible both in terms of a decreasing index of geographical export concentration and an increasing number of partner countries. The break in the series between 2003 and 2004 (see remark on p. 422) prevents us from directly comparing *Avg_N_geo* and *Avg_HH_geo* values before and after the accession to the EU. Instead, we will proceed by considering two separate sub-periods (1994–2003, 2004–2010). However, even when we focus only on the later period (2004–2010) a decline in the geographical concentration of exports (*Avg_HH_geo*) is noticeable. At the same time, the average number of countries to which Polish products were exported rose from 38 to 49 (values based on plot A).

Average values provide general information but there are undoubtedly huge cross-product differences in the number of destination countries for Polish exports. In order to illustrate changes in the disparity of trade partners across products exported from Poland, in Figure 7 we show four percentage histograms (the height of the bars is scaled so that the sum of their heights equals 100), corresponding to the boundary years of the two sub-periods (1994–2003, 2004–2010). Analysing the values of *N_geo* (cut-off A) for all 4,963 products, in 1994 they ranged from 0 (no receiving market – product line not active at all) to 52, and in 2003 from 0 to 63. In the second sub-period (2004–2010), after the change in data collection methodology, the number of destination countries for different product lines varied between 0 and 83 in 2004, and between 0 and 109 in 2010. In both sub-periods the distribution moved significantly towards the right, which illustrates a rise in the variety of markets receiving Polish products. The percentage of products not exported to any of the partner countries dropped significantly, especially in the first sub-period (activity at the extensive margin), while after 2004 a major change concerned an increase in the number of partner countries for already existing product lines (activity at the intensive margin).

In order to verify whether the trend of increasing geographical diversification of Polish exports is statistically meaningful, we employ a simple econometric model, regressing measures of partner country diversity on time:

$$Y\_geo_{PLkt} = \alpha + \beta t_{PLk} + D_k + \varepsilon_{PLkt} \tag{8}$$

where $Y\_geo = \{N\_geo, HH\_geo\}$, t refers to time, k to product line and ε is the standard error term. In order to account for product specificity, in a benchmark specification we adopt a fixed-effect model with product dummies D_k.

The estimation results are reported in Table 7. We use all 84,371 observations, but when *HH_geo* is undefined (no positive exports of a given product to any of the countries) then these observations are dropped from the model. The baseline FE estimates confirm the increasing trend of geographical diversification of Polish exports: the coefficient associated with *N_geo* is positive and statistically significant (model 1), while that referring to the concentration index *HH_geo* is negative and significant (model 2). Hence, in the years 1994–2010 the geographical diversity of partner countries receiving goods exported from Poland rose.

Table 7 also reports a number of robustness checks. First of all, we correct for the importance of the break in the series in 2003–2004 by introducing a dummy for the year 2004 (models 3 and 4). Then, as the number of partner countries (*N_geo*) is a count variable, we also employ a negative binomial estimation (model 5), but the fit of this model is much lower than the benchmark one. In the case of the model with *HH_geo* as dependent variable, we consider censoring (*HH_geo* is normalized and thus is bounded at 0 and 1) and perform a logistic transformation regression (column 6). These changes in the estimation strategy do not affect the main result: a statistically significant trend of geographical diversification of Polish export lines in the period 1994–2010.

As stated previously, we are not able to compute the product levels *N_geo* and *HH_geo* for all the other countries in the sample. However, in order to give an idea of the level of geographical diversity of Polish exports in a comparative setting, in Table 8 we show the values of *IEMP* (defined in eq. (6)) for all EU27 economies. In 2010, Poland exploited approximately one-fifth (22%) of its overall market reach potential. This may seem low, but Germany, the European and global leader in this respect (as documented in Reis and Farole (2012, pp. 43–44) exploits around half of its potential. Polish *IEMP* is the highest among all the new member states and is higher than that of richer (but smaller) EU15 countries.

In the next step, we verify whether in terms of geographical diversity Polish exports are above or below the value typical for countries at approximately the same stage of economic development. Figure 8 and Figure 9 illustrate the relationship between geographical market reach and levels of economic development, corresponding to the models adopted in Section 4.1. The lowess line plotted in Figure 8 represents the fit of a non-parametric model:

$$Y\_geo_{it} = \alpha + s(GDPpc_{it}) + \varepsilon_{it} \tag{9}$$

where *Y_geo* = *IEMP*.

Figure 9 is a GAM plot of the partial residuals obtained with a semi-parametric model with country size (*GDP*) as an additional covariate:

$$Y\_geo_{it} = \alpha + s(GDPpc_{it}) + \beta GDP_{it} + \varepsilon_{it} \tag{10}$$

The conclusions are similar to those referring to product diversity. The initial phase of the economic development process can be associated with a rise in the variety of receiving markets. Importantly, the degree of geographical diversification of Polish exports is higher than that typical for countries with similar income *per capita* levels, even after taking into account the importance of the country dimension.

5. Summary of the findings and conclusions

Following advances in trade theory, the empirical literature on trade diversification has been expanding rapidly in recent years. The theme of export diversity is particularly important from the point of view of risk reduction strategies: countries with highly diversified economic structures are less exposed to shocks and price swings in foreign markets.

This paper has focused on the specific case of Polish exports observed (due to data availability) in the period 1994–2010 and has compared them to European and global trends. We have analyzed two aspects of the trade diversification process: changes in the commodity structure of the Polish trade basket and the geographical diversity of the receiving markets for Polish exports.

In the first instance, we constructed synthetic measures of export product diversity based on concentration/inequality measures. Secondly, we used product-level statistics on exports from Poland to all its potential receiving markets in the world to measure the degree of geographical diversification and market reach success. Through the use of highly disaggregated data (almost 5,000 product lines) on Polish exports to all partner countries in the world, we have been able to contribute to the existing related literature, which is somewhat limited in terms of product/partner details and the time dimension.

The results suggest that Polish exports are already well diversified. In terms of its commodity structure, Poland can be placed among the EU countries with highly heterogeneous baskets of products exported (on average exporting 84% of all product lines present in the sample, while the EU27 average is equal to 78%). When employing a non-parametric model linking product diversification measures with income *per capita* data for 163 countries, Poland can be placed on the left-hand side of the U-shaped curve of trade diversification along the development process. Importantly, we have shown that the Polish export portfolio is relatively more diversified than this typical for economies at approximately the same stage of economic development. This result also holds when we adopt a semi-parametric methodology to correct the model for the importance of the country size.

Similarly, the geographical diversity of the partner countries receiving goods exported from Poland is relatively high and, additionally, it increased in the period analysed. At the moment of accession to the EU (2004), Poland exploited 17% of its overall foreign market penetration potential; in 2010 this indicator was equal to 22%, which placed Poland in first place among all the new member states. By means of a parametric regression model performed with product-level indicators of partner diversity, we demonstrated that the trend of increasing geographical diversification of Polish exports is statistically significant. Moreover, as with the case of commodity diversity, we demonstrated that the degree of geographical diversification of Polish exports is higher than that typical for economies at approximately the same stage of economic development.

The results reported are important from the policy point of view. The relatively high degree of export product diversity which is typical for Poland is a positive sign, especially in the light of export structure resistance to product-specific shocks. Similarly, the rising geographical diversification of trade partners implies a reduction in exposure to external country-specific shocks. On the other hand, there might also be a negative aspect of the high level of diversification of Polish exports (a low level of specialization could be a sign of a low level of maturity of Polish exports). This aspect could be further examined through exploration of the quality content of its exports.

References

Adamowicz E., Dudek S., Pachucki D., Walczyk K. (2009), Synchronizacja cyklu koniunkturalnego polskiej gospodarki z krajami strefy euro w kontekście struktury tych gospodarek, in: *Raport na temat pełnego uczestnictwa Rzeczypospolitej Polskiej w trzecim etapie Unii Gospodarczej i Walutowej. Projekty badawcze*, 1, Narodowy Bank Polski, Warszawa.

Amurgo-Pacheco A., Pierola M.D. (2008), *Patterns of export diversification in developing countries: intensive and extensive margins*, Policy Research Working Paper, 4473, The World Bank, Washington, DOI: 10.1596/1813-9450-4473.

Bacchetta M., Jansen M., Lennon C., Piermartini R. (2009), Exposure to external shocks and the geographical diversification of exports, in: R. Newfamer, W. Shaw, P. Walkenhorst (eds.), *Breaking into new markets. Emerging lessons for export diversification*, The World Bank, Washington.

de Benedictis L., Gallegati M., Tamberi M. (2009), Overall trade specialization and economic development: countries diversify, *Review of World Economics*, 145(1), 37–55.

de Benedictis L., Tajoli L. (2007), Economic integration and similarity in trade structures, *Empirica*, 34(2), 117–137.

Brenton P., Newfarmer R. (2007), *Watching more than the Discovery channel: export cycles and diversification in development*, Policy Research Working Paper Series, 4302, The World Bank, Washington.

Brenton P., Newfarmer R. (2009), Watching more than the Discovery channel to diversify exports, in: R. Newfamer, W. Shaw, P. Walkenhorst (eds.). *Breaking into new markets. Emerging lessons for export diversification*, The World Bank, Washington.

Brodzicki T. (2009), *Product variety and the export pattern of Poland (1999–2006)*, Analizy i Opracowania KEIE UG, 3/2009(002), http://gnu.univ.gda.pl/~keie/aio16.pdf.

Bruzda J. (2011), Business cycle synchronization according to wavelets – the case of Poland and the euro zone member countries, *Bank i Kredyt*, 42(3), 5–33.

Cadot O., Carrère C., Kukenova M., Strauss-Kahn V. (2010), *OECD imports: diversification and quality search*, Policy Research Working Paper, 5258, The World Bank, Washington.

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

Cadot O., Carrère C., Strauss-Kahn V. (2012), Trade diversification, income, and growth: what do we know?, *Journal of Economic Surveys* (forthcoming), DOI: 10.1111/j.1467-6419.2011.00719.x.

Cleveland W. (1979), Robust locally weighted regression and smoothing scatterplots, *Journal of American Statistic Association*, 74(368), 829–836.

Di Giovanni J., Levchenko A. (2010), *The risk content of exports: a portfolio view of international trade*, NBER Working Paper, 16005, National Bureau of Economic Research, Cambridge.

Evenett S.J., Venables A. (2002), *Export growth by developing economies: market entry and bilateral trade*, http://www.alexandria.unisg.ch/Publikationen/22177.

Fertó I., Soós K.A. (2008), Trade specialization in the European Union and in postcommunist European countries, *Eastern European Economics*, 46(3), 5–28.

Fidrmuc J., Korhonen I. (2006), Meta-analysis of the business cycle correlation between the euro area and the CEECs, *Journal of Comparative Economics*, 34(3), 518–537.

Funke M., Ruheweld R. (2005), Export variety and economic growth in East European transition economies, *Economics of Transition*, 13(1), 25–50.

Ghosh A.R., Ostry J.D. (1994), *Export instability and the external balance in developing countries*, IMF Staff Papers, 41(2), Palgrave Macmillan, Washington.

Grossman G., Helpman E. (1991), *Innovation and growth in the global economy*, MIT Press, Cambridge.

Helpman E., Krugman P. (1985), *Trade policy and market structure*, MIT Press, Cambridge.

Herzer D., Danzinger F.N.L. (2006), What does export diversification do for growth? An econometric analysis, *Applied Economics*, 38(15), 1825–1838.

Hesse H. (2009), Economic diversification and economic growth, in: R. Newfamer, W. Shaw, P. Walkenhorst (eds.), *Breaking into new markets. Emerging lessons for export diversification*, The World Bank, Washington.

Hummels D., Klenow P. (2005), The variety and quality of a nation's exports, *American Economic Review,* 94(1), 704–723.

Imbs J., Wacziarg R. (2003), Stages of diversification, *American Economic Review*, 93(1), 63–86.

Jaimovich E. (2012), Import diversification along the growth path, *Economics Letters*, 117(1), 306–310.

Klinger B., Lederman D. (2006), *Diversification, innovation, and imitation inside the global technology frontier*, World Bank Policy Research Working Paper, 3872, The World Bank, Washington.

Konopczak K. (2009), Analiza zbieżności cyklu koniunkturalnego gospodarki polskiej ze strefą euro na tle krajów Europy Środkowo-Wschodniej oraz państw członkowskich strefy euro, in: *Raport na temat pełnego uczestnictwa Rzeczypospolitej Polskiej w trzecim etapie Unii Gospodarczej i Walutowej. Projekty badawcze*, Narodowy Bank Polski, Warszawa.

Konopczak K., Marczewski K. (2011), Why so different from other CEECs – Poland's cyclical divergence from the euro area during the recent financial crisis, *Bank i Kredyt*, 41(2), 7–30.

Koźluk T. (2005), CEEC accession countries and the EMU. An assessment of relative suitability and readiness for euro-area membership, *Journal of Economic Integration*, 51, 439–474.

Krugman P.R. (1979a), Increasing returns, monopolistic competition, and international trade, *Journal of International Economics*, 9(4), 469–479.

Krugman P.R. (1979b), A model of innovation, technology transfer, and the world distribution of income, *Journal of Political Economy*, 87(2), 253–266.

Krugman P.R. (1980), Scale economies, product differentiation, and the pattern of trade, *American Economic Review*, 70(5), 950–959.

Krugman P.R. (1981), Intraindustry specialization and the gains from trade, *Journal of Political Economy*, 89(5), 959–973.

Melitz M.J. (2003), The impact of trade on intra-industry reallocations and aggregate industry productivity, *Econometrica*, 71, 1695–1725.

Misztal P. (2011), Export diversification and economic growth in European Union member states, *Oeconomia*, 10(2), 55–64.

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

Parteka A. (2013), Trade diversity and stages of development – evidence on EU countries, *Ekonomia*, 30, 23–44.

Parteka A., Tamberi M. (2013), What determines export diversification in the development process? Empirical assessment, *The World Economy*, DOI: 10.1111/twec.12064

Reis J.G., Farole T. (2012), *Trade competitiveness diagnostic toolkit,* The International Bank for Reconstruction and Development / The World Bank, Washington, https://openknowledge.worldbank.org/bitstream/handle/10986/2248/673620PUB0EPI007869B009780821389379.pdf?sequence=1.

Royston P., Ambler G. (2012), *GAM: Stata module for general additive models*, Statistical Software Components, S428701, Boston College Department of Economics, Boston.

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

Shaw W., Newfarmer R., Walkenhorst P., eds. (2009), *Breaking into new markets: emerging lessons for export diversification*, The World Bank, Washington.

Shepherd B. (2010), Geographical diversification of developing country exports, *World Development*, 38(9), 1217–1228.

Shepotylo O. (2012), Export diversification across countries and products: Do Eastern European (EE) and Commonwealth of Independent States (CIS) countries diversify enough?, *The Journal of International Trade & Economic Development: An International and Comparative Review*, DOI: 10.1080/09638199.2011.577797.

Skrzypczyński P. (2009), Wahania aktywności gospodarczej w Polsce i strefie euro, in: *Raport na temat pełnego uczestnictwa Rzeczypospolitej Polskiej w trzecim etapie Unii Gospodarczej i Walutowej. Projekty badawcze*, Warszawa, Narodowy Bank Polski.

Wood S.N. (2006), *Generalized additive models: an introduction with R,* Chapman and Hall, Washington.

Zaghini A. (2005), Evolution of trade patterns in the new EU member states, *Economics of Transition*, 13(4), 629–658.

Acknowledgements

The paper presents the results of a research project realized while the author was visiting Universitat Pompeu Fabra (UPF, Barcelona, Spain). Financial support from the Polish Ministry of Science and Higher Education (Mobilnosc Plus 652/MOB/2011/0) and the hospitality of UPF Department of Economics are gratefully acknowledged. I would like to thank two anonymous referees for detailed and constructive comments on an earlier draft of the paper. All the remaining errors are mine.

Appendix

Table 1
Summary statistics of international panel with country-year observations matching commodity diversification indices with income *per capita* data

		All countries[a]	EU27 countries[a]	Poland
Total number of country-year observations		1,905	486	17
Time span		1988–2010	1988–2010	1994–2010
Number of countries		163	27	1
Number of year observations per country	Mean	11	18	17
	Min	1	12	17
	Max	23	23	17
Number of country observations per year	Mean	82	21	1
	Min	11	4	1
	Max	130	27	1

[a] Unbalanced panels.

Table 2
List of European countries (EU27) present in the dataset summarized in Table 1

Country name	Country code	Years covered	
		first	last
Austria	AUT	1994	2010
Belgium	BEL	1999	2010
Bulgaria	BGR	1996	2010
Cyprus	CYP	1989	2010
Czech Republic	CZE	1993	2010
Germany	GER	1988	2010
Denmark	DNK	1989	2010
Spain	ESP	1989	2010
Estonia	EST	1995	2010
Finland	FIN	1988	2010
France	FRA	1994	2010
United Kingdom	GBR	1993	2010
Greece	GRC	1988	2010
Hungary	HUN	1992	2007
Ireland	IRL	1992	2010
Italy	ITA	1994	2010
Lithuania	LTU	1994	2010
Luxembourg	LUX	1999	2010
Latvia	LVA	1994	2010
Malta	MLT	1994	2010
Netherlands	NLD	1992	2006
Poland	POL	1994	2010
Portugal	PRT	1988	2010
Romania	ROM	1989	2010
Slovak Republic	SVK	1994	2010
Slovenia	SVN	1994	2010
Sweden	SWE	1992	2010

Table 3

Summary statistics of commodity diversification measures

	Commodity diversification measures						
	index		**obs.**	**mean**	**sd**	**min**	**max**
Number of products	*N_prod*	Poland	17	4,176	147.45	3,863	4,426
Normalised Herfindahl	*HH_prod*	Poland	17	0.0062	0.00084	0.0049	0.0077
Theil index	*Theil_prod*	Poland	17	2.29	0.070	2.17	2.41
Number of products	*N_prod*	EU27	472	3,784	872.69	829	4,858
Normalised Herfindahl	*HH_prod*	EU27	467	0.021	0.030	0.0033	0.2009
Theil index	*Theil_prod*	EU27	467	2.79	0.791	1.626	5.962
Number of products	*N_prod*	All countries[a]	1,867	2,689	1,461.45	171	4,858
Normalised Herfindahl	*HH_prod*	All countries[a]	1,867	0.1059	0.154	0.0033	0.798
Theil index	*Theil_prod*	All countries[a]	1,867	4.193	1.53	1.62	7.838

Note: sample of 163 countries.
[a] observations below 1st or above 99th percentile excluded.

Table 4

Summary statistics of geographical diversification measures

	Geographical diversification measures						
	index		**obs.**	**mean**	**sd**	**min**	**max**
Number of partners	*N_geo* (*A*)	Poland	84,371	8.02	11.44	0	109
Normalised Herfindahl	*HH_geo* (*A*)	Poland	58,321	0.344	0.068	0.200	0.491
Number of partners	*N_geo* (*B*)	Poland	84,371	5.6	5.40	0	32
Normalised Herfindahl	*HH_geo* (*B*)	Poland	60,806	0.351	0.069	0.231	0.496
Index of export market penetration	*IEMP*	Poland[a]	7	18.52	1.95	16.47	21.99
Index of export market penetration	*IEMP*	EU27[b]	462	16.59	13.27	1.93	53.60
Index of export market penetration	*IEMP*	All countries[b]	1,821	10.09	11.24	1.06	53.60

Note: alternative thresholds used to calculate *N_ geo* and *HH_ geo* indices.
[a] 2004–2010 (see note 13).
[b] Observations below 1st or above 99th percentile excluded.

Table 5

Measures of exported products diversification, Poland versus global and European trends (average values 1994−2010)

	Number of products *N_prod*	Normalised Herfindahl *HH_prod*	Theil index *Theil_prod*
Poland	4,176	0.006	2.29
EU27	3,864	0.211	2.75
All countries (163)	2,701	0.112	4.22
Developing countries	2,051	0.146	4.82
Developed countries	3,691	0.061	3.30

Note: theoretical max = 4,963; based on World Bank's classification; numbers given for the EU27 and all countries are averages across individual countries and not region-wide aggregates.

Source: HS0 6 digit trade data from UNComtrade.

Table 6

Correlation coefficients between measures of exported products diversity and country size (GDP, population)

		Number of products *N_prod*	Normalised Herfindahl *HH_prod*	Theil index *Theil_prod*
All countries (163), n = 1,902	*pop*	0.26***	-0.10***	-0.21***
All countries (163), n = 1,896	*GDP*	0.32***	-0.15***	-0.29***
EU27, n = 486	*pop*	0.51***	-0.30***	-0.56***
EU27, n = 484	*GDP*	0.51***	-0.24***	-0.50***

Notes: n – number of observations; all available observations from unbalanced panel 1988−2010 used.
pop – population (in thousands), *GDP* – real income (in constant prices in USD taken from 2000).
*** denote significance at 1% level.

Source: trade data from UNComtrade, GDP and population data from World Development Indicators (2011).

Table 7

Regression results: measures of geographical diversification of export lines versus time trend, Poland (1994–2010)

	Dependent variable: measure of geographical diversity of export lines					
	basic		**robustness**			
	N_geo **(1) FE**	***HH_geo*** **(2) FE**	***N_geo*** **(3) FE**	***HH_geo*** **(4) FE**	***N_geo*** **(5)NB**	***HH_geo*** **(6) LT**
Time	0.7496*** [0.0135]	-0.0041*** [0.0001]	0.7419*** [0.0135]	-0.0041*** [0.0001]	0.0945*** [0.001]	-0.0183*** [0.0003]
Number of observations	84,371	58,321	84,371	58,321	84,371	58,321
Number of products	4,963	4,649	4,963	4,649	4,963	4,649
R2/pseudoR2	0.3	0.14	0.3	0.15	0.02	0.15
Product fixed effects	yes	yes	yes	yes	yes	yes
Dummy 2004			yes	yes		
Time period	1994–2010	1994–2010	1994–2010	1994–2010	1994–2010	1994–2010

Notes:
Indices obtained with cut-off point (A) *exp_ikj* > 10,000 USD.
*** denote significance at 1% level.
SE (robust in case of FE) in parenthesis under coefficients, FE – fixed effects estimation, NB – negative binomial, LT – logistic transformation. Constant included – not reported.

Source: trade data from UNComtrade.

Table 8

Index of export market penetration (in %). Poland versus other EU27 countries (2010)

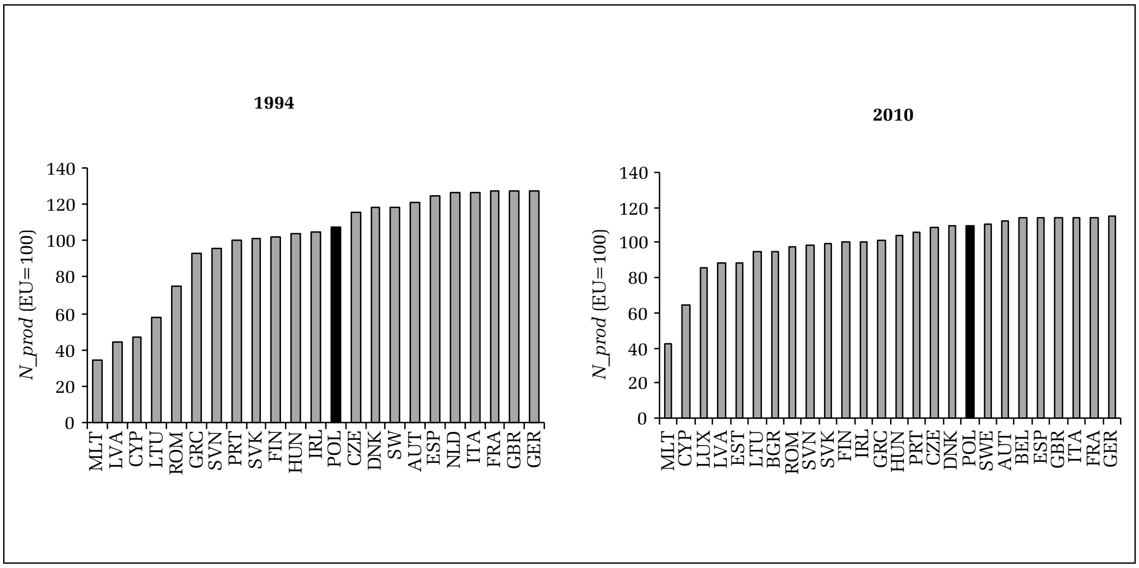
Country	***IEMP***	**Country**	***IEMP***	**Country**	***IEMP***	**Country**	***IEMP***
MLT	2.8	LTU	8.1	FIN	14.6	ESP	34.9
CYP	3.1	GRC	8.8	CZE	19.5	BEL	36.2
EST	5.8	SVK	9.2	DNK	20.3	GBR	41.2
LVA	6.3	ROM	9.6	SWE	21.2	FRA	42.1
LUX	6.4	SVN	10.4	POL	22.0	ITA	44.8
IRL	7.8	HUN	11.3	AUT	27.7	GER	52.5
BGR	8.1	PRT	13.5	NLD	32.4		

Note: *IEMP* based on HS0 6 digit mirrored exports.

Source: trade indicators data from UNComtrade/WITS.

Figure 1

Number of exported products Poland versus other EU countries (1994 and 2010, EU27 = 100)

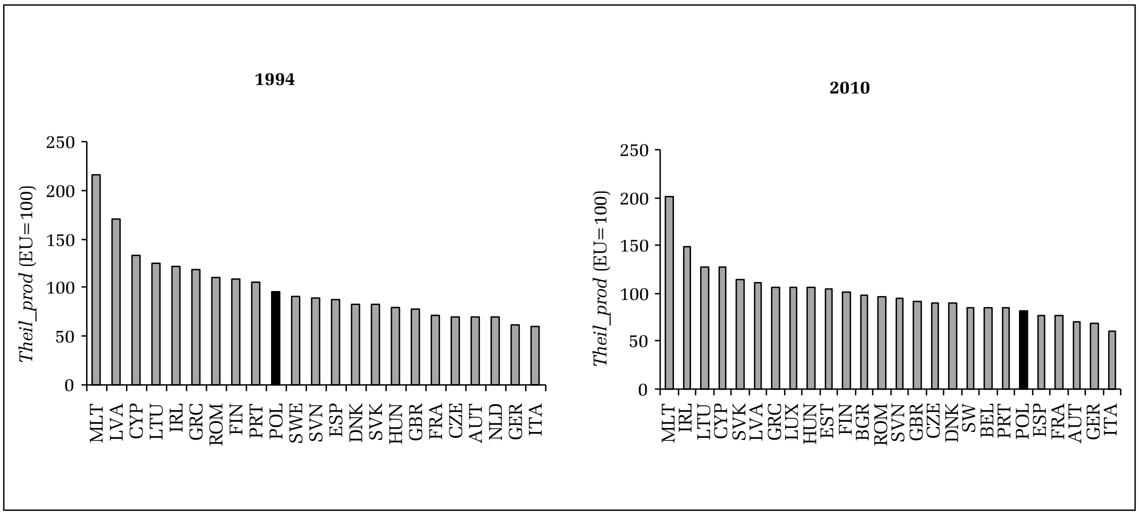


Notes: *N_prod* for each country expressed with respect to average value for all other EU countries in the group in the particular year. Countries not present in 1994 sample: Belgium and Luxembourg (separate data available since 1999), Bulgaria (data since 1996), Estonia (data since 1995).

Source: HS0 6 digit trade data from UNComtrade.

Figure 2

Measure of exported products' concentration (Theil_prod index) – Poland versus other EU countries (1994 and 2010, EU27 = 100)

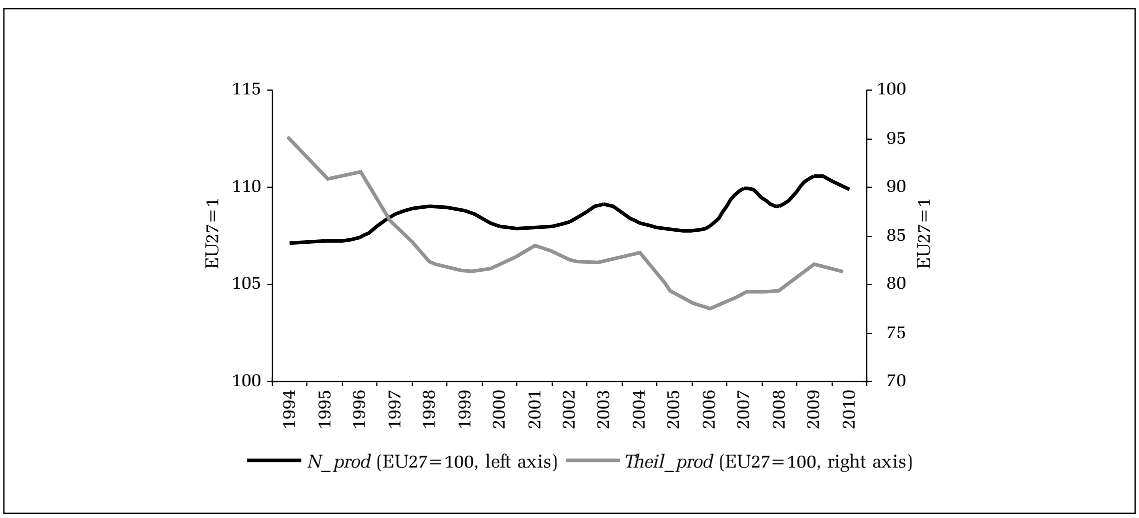


Notes: *Theil_prod* for each country expressed with respect to average value for all other EU countries in the group in the particular year. Countries not present in 1994 sample: Belgium and Luxembourg (separate data available since 1999), Bulgaria (data since 1996), Estonia (data since 1995).

Source: HS0 6 digit trade data from UNComtrade.

Figure 3

Evolution of exported products' diversity in time – Poland versus EU27 average (1994–2010)

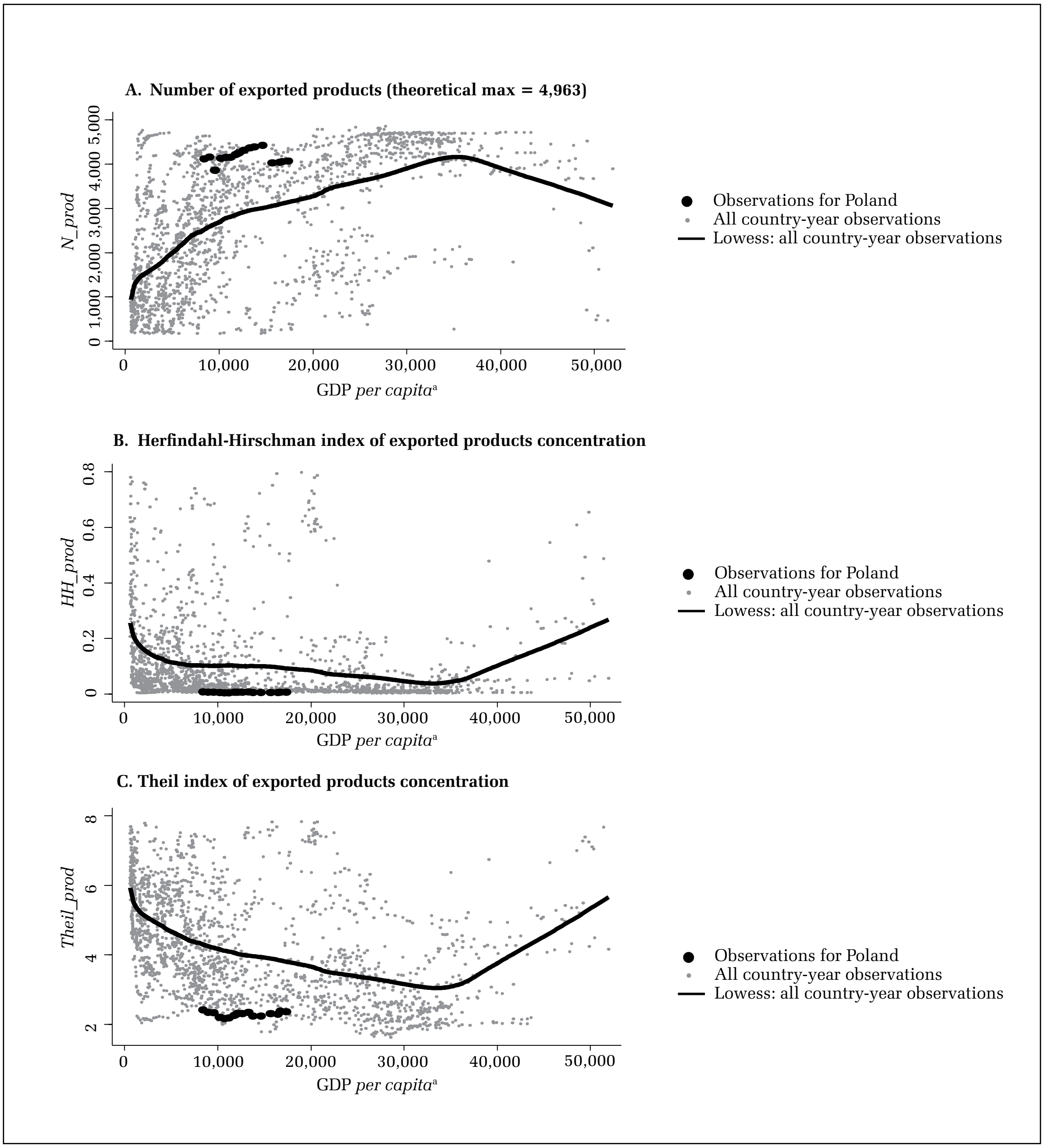


Notes: Polish value with respect to the average for all other EU27 countries in the group in the particular year. Belgium and Luxembourg in the sample since 1999, Bulgaria – since 1996, Estonia – since 1995.

Source: HS0 6 digit trade data from UNComtrade.

Figure 4

Relationship between measures of exported products diversity and income *per capita* levels – nonparametric plots

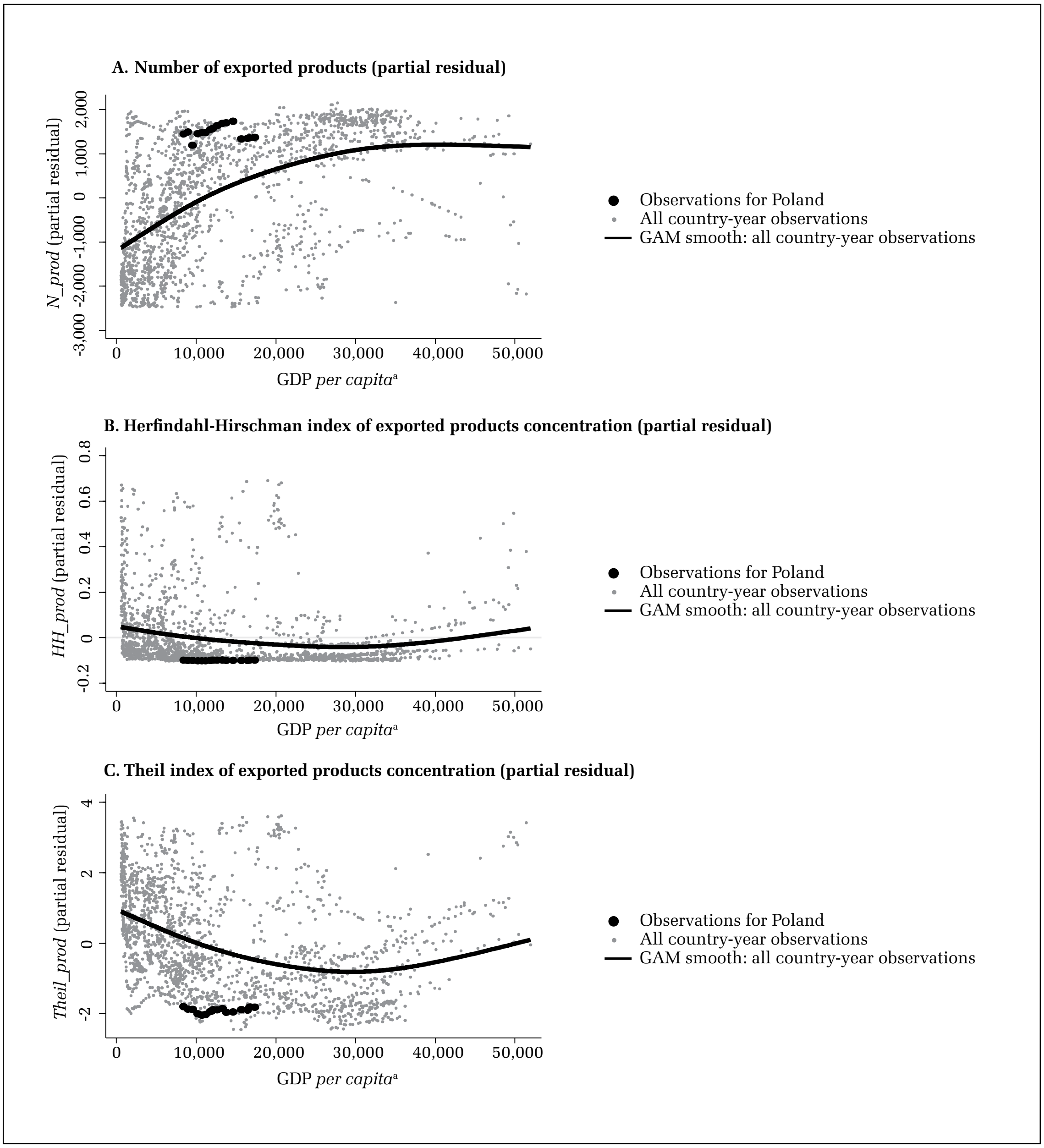


Notes: solid lines corresponds to lowess prediction (span = 0.5), estimated equation: $Y\_prod = \alpha + s(GDPpc) + \varepsilon$. Outliers (defined as observations below 1st or above 99th percentile) excluded; number of observations =1,828.
[a] PPP, constant prices in USD taken from 2005.

Source: UNComtrade (HS0 6 digit, 4,963 product lines) and GDP *per capita* from WB WDI.

Figure 5

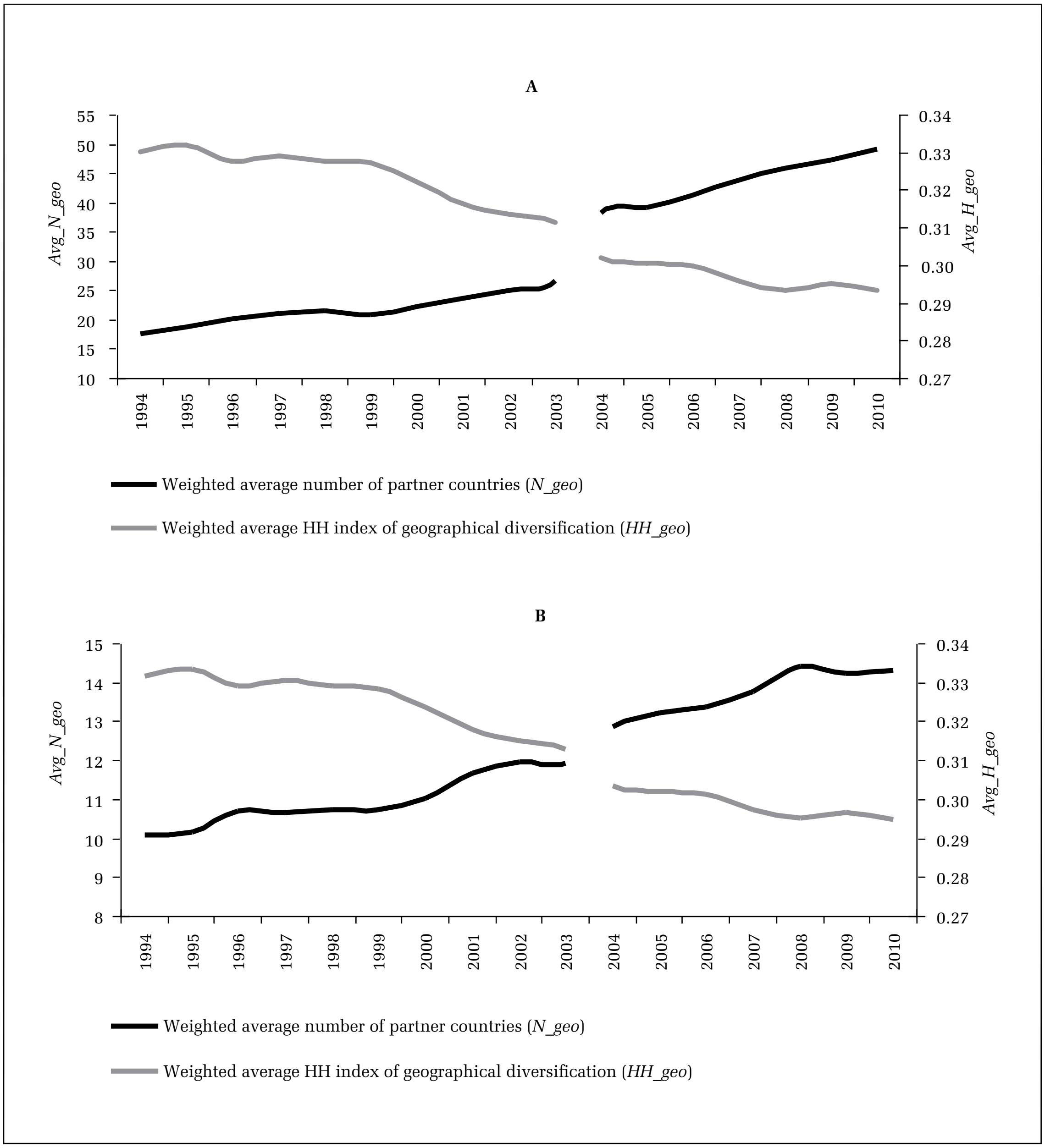
Relationship between measures of exported products diversity and income *per capita* levels – GAM plots with country size correction



Notes: solid lines corresponds to GAM smooth (df = 2), estimated equation: $Y\_prod = \alpha + s(GDPpc) + \beta GDP + \varepsilon$. Outliers (defined as observations below 1st or above 99th percentile) excluded; number of observations = 1,828.
[a] PPP, constant prices in USD taken from 2005.

Source: trade data from UNComtrade (HS0 6 digit, 4,963 product lines), GDP *per capita* and GDP from WB WDI.

Figure 6
Geographical diversification of Polish exports (1994–2010), alternative thresholds

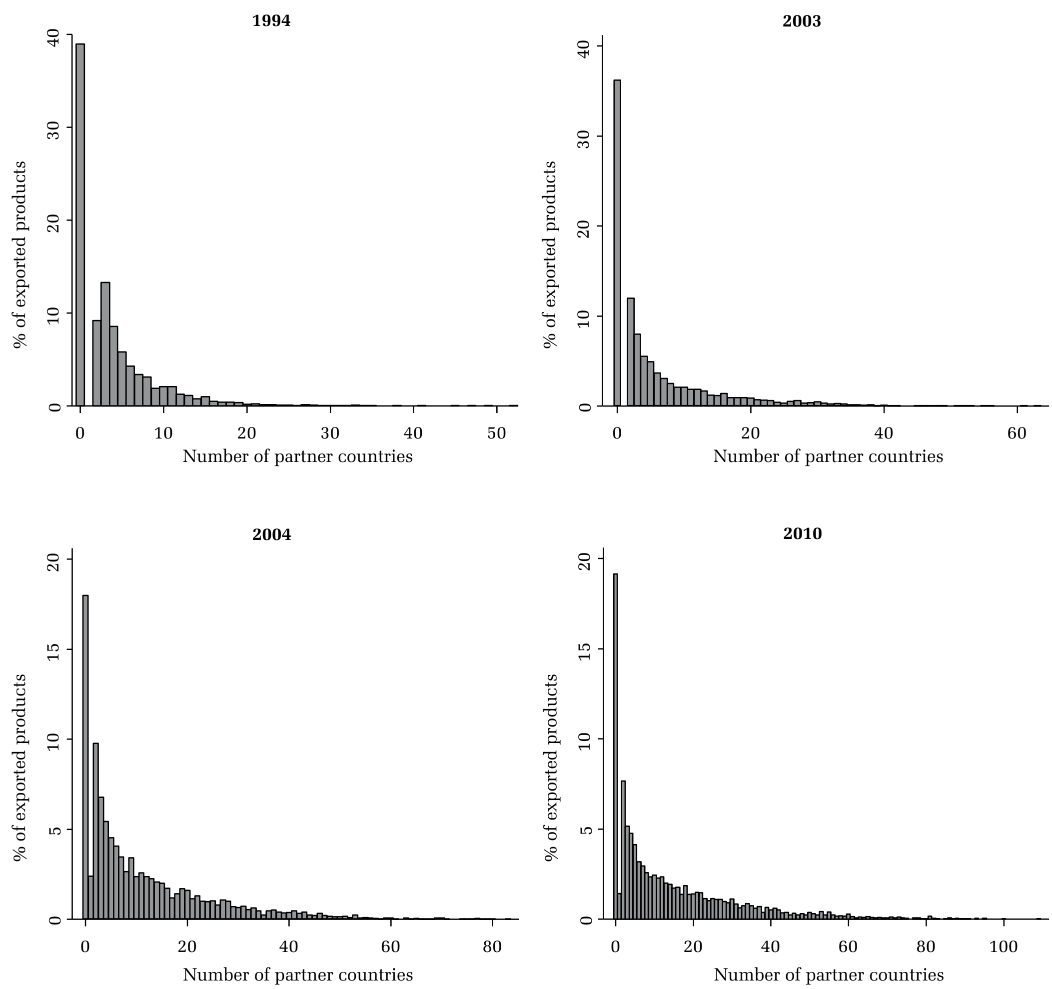


Notes: weighted averages across 4,963 product lines; weights – share of exports of each product line (sum across partners) in total Polish exports.
Break in series 2003–2004 (see remark on p. 442).

Source: trade data from UNComtrade (HS0 6 digit).

Figure 7

Geographical diversification of Polish exports (histograms, 1994–2003 and 2004–2010)

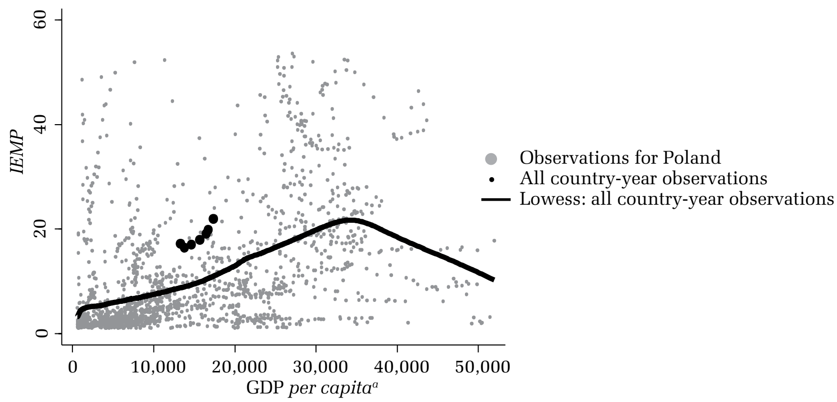


Notes: theoretical max 4,963 = 100%. Only exports above the cut-off point (A) *exp_ikj* > 10,000 USD taken into account. Each bar corresponds to one product line; sum of the bars' height =100; 2003–2004 – break in series (see remark on p. 422).

Source: trade data from UNComtrade (HS0 6 digit).

Figure 8

Relationship between geographical diversity of exports (*IEMP*) and income *per capita* levels – nonparametric plots

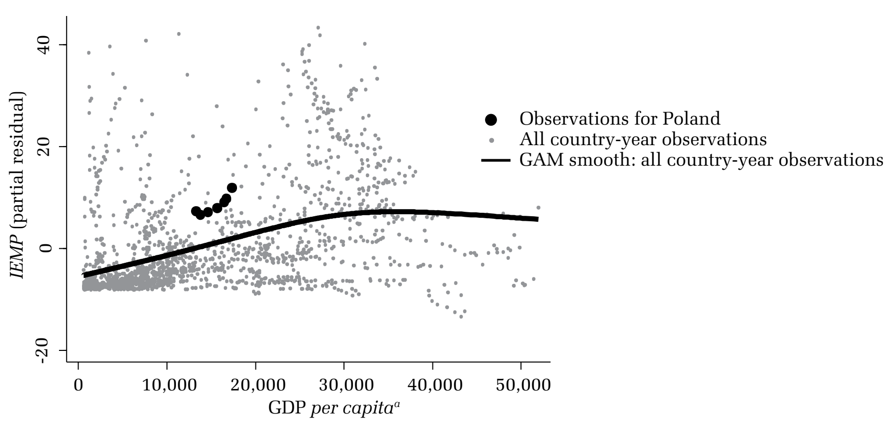


Notes: solid lines corresponds to lowess prediction (span = 0.5), estimated equation: $IEMP = \alpha + s(GDPpc) + \varepsilon$. Outliers (defined as observations below 1st or above 99th percentile) excluded, number of observations = 1,821.
[a] PPP, constant prices in USD taken from 2005.

Source: trade indicators data from UNComtrade/WITS (HS0 6 digit), GDP *per capita* from WB WDI.

Figure 9

Relationship between geographical diversity of exports (*IEMP*) and income *per capita* levels – GAM plots with country size correction



Notes: solid lines corresponds to GAM smooth (df = 2), estimated equation: $IEMP = \alpha + s(GDPpc) + \beta GDP + \varepsilon$. Outliers (defined as observations below 1st or above 99th percentile) excluded, number of observations =1,821.
[a] PPP, constant prices in USD taken from 2005.

Source: trade indicators data from UNComtrade/WITS (HS0 6 digit), GDP *per capita* and GDP from WB WDI.