Structural Change and Economic Dynamics 57 (2021) 28–44

Contents lists available at ScienceDirect

Structural Change and Economic Dynamics

journal homepage: www.elsevier.com/locate/strueco

Relative Productivity, Country Size and Export Diversification

Andrzej Cieślik[a], Aleksandra Parteka[b,#,*]

[a] University of Warsaw, Faculty of Economic Sciences, Department of Macroeconomics and International Trade Theory, Dluga 44/50, 00241 Warszawa, Poland
[b] Gdansk University of Technology, Faculty of Management and Economics, Narutowicza 11/12; 80-233 Gdańsk, Poland

A R T I C L E I N F O

Article history:
Received 26 February 2020
Revised 15 November 2020
Accepted 4 January 2021
Available online 22 January 2021

JEL:
F11
F14
F43
O40
O11

Keywords:
economic development
export diversification
export variety
relative productivity
relative country size
Ricardian model

A B S T R A C T

This paper analyses the effects of productivity and country size on the extent of trade structure diversification. Using a testable version of the Ricardian model, we show that relative export variety is an outcome of two forces: a relative productivity change (technological progress) and a relative country size change (labour force growth). The model predictions are validated empirically using product-level trade data for a sample of 132 countries (1988–2014), including 53 low-income countries. We find a robust positive relationship between export variety and the countries' relative productivity, as well as a negative relationship between export variety and the expansion of foreign economies (i.e., the growth of the RoW). The effect of technology differences on export variety is driving diversification especially at the beginning of the development process. The results are robust to changes in the measurement of export variety (also in terms of economic complexity), in the set of control variables, or in the estimation methods.

Introduction

Understanding of the factors related to the countries' ability to diversify their production and exports structures is vital for sustainable economic development and structural changes (Mania and Rieber, 2019). Many developing countries are still dependent on a narrow range of primary products (Newfamer et al., 2009). On average, low-income countries have 50% less diversified exports than high-income ones.[1] Given the possibility of diversification-led growth (Herzer and Nowak-Lehnmann, 2006; Hesse, 2008; Gozgor and Can, 2016; Kaitila, 2019), the positive relationship between the extensive margin and the total trade growth[2] (Kehoe and Ruhl, 2013), the complexity of export diversification strategies dependent on the already established production capabilities (Hausmann et al., 2019; Hidalgo et al., 2007; Tacchella et al., 2013) and the distance to comparative advantage at different income levels (Lectard and Rougier, 2018), as well as the role that a high export variety plays in reducing risk and volatility related to the presence in international markets (Haddad et al., 2013; Balavac and Pugh, 2016), it is crucial to understand what drives the differences in export variety.[3]

The seminal paper by Imbs and Wacziarg (2003, hereafter referred to as 'IW') began a strand of literature on 'stages of de-

* Corresponding author.
E-mail addresses: cieslik@wne.uw.edu.pl (A. Cieślik), aparteka@zie.pg.gda.pl (A. Parteka).

This research has been conducted within a Fulbright Senior Award project - Aleksandra Parteka acknowledges the support of the Polish-U.S. Fulbright Commission and the hospitality of University of California, Berkeley. All remaining errors are the authors'.

[1] Based on the Theil index of export concentration computed with HS 6-digit export data (source: UN Comtrade), 2014 (see Section 4.1 for details).

[2] Kehoe and Ruhl (2013) found that during the 1995–2005 period the extensive margin accounted for 9.9 per cent of the trade growth for the NAFTA country pairs and 26.0 per cent of trade growth between the United States and Chile, China, and Korea.

[3] The degree of export diversification (export variety) is a counterpart of the degree of export base specialisation. Throughout this paper we use the expressions 'export variety' and 'export diversification' interchangeably. Note the difference between the determinants of the *degree* of export diversification (measured by the number of products exported) and the determinants of export specialisation (meaning the *characteristics* of products that countries trade intensively). See Shikher (2013) for the explanations of specialisation and comparative advantage patterns across countries.

https://doi.org/10.1016/j.strueco.2021.01.002

velopment'. IW described an empirical observation of a U-shaped path of diversification of economic structures accompanying the growth path, followed by re-specialisation at higher levels of income. Since then, given the greater detail of trade statistics, the focus has shifted towards the analysis of the variety observed in trade patterns. Numerous empirical papers (including Basile et al., 2018; Cadot et al., 2011; De Benedictis et al., 2009; Klinger & Lederman, 2006; Parteka, 2010; Parteka & Tamberi, 2013a,b; and Mau, 2016) study the evolution of the variety (diversity) occurring in export structures as countries grow. These papers, with respect to the original IW (2003) contribution, expanded the set of the analysed countries and either added more explanatory variables or modified the methodology of measuring diversification and estimating diversification curves.

However, there is still an important research gap that this paper is meant to bridge: the lack of connection between the theory and empirics in most export diversification-economic development studies. This means that the theoretical explanation of an empirically revealed export diversification phenomenon along the development path is still missing. In contrast to the (few) existing theoretical explanations of export variety patterns, we deviate from the Heckscher–Ohlin (HO) framework (adopted in Regolo, 2013 and Cadot et al., 2011) in favour of the approach rooted in the recently-revived (mainly due to Eaton and Kortum, 2002; hereafter referred to as 'EK') Ricardian theory of international trade. Thus, this paper studies the evolution of export variety (export diversification) along the path of economic growth in a multi-good Ricardian country-level setting.

We set a theoretical background for our empirical analysis using a testable version of a Ricardian model, in which the countries' relative export variety, assessed with respect to the rest of the world (RoW), is a function of relative technology differences across countries (relative productivity), as well as a relative country size. Using this framework, we put forth three main propositions. Firstly, we demonstrate that increased relative productivity of a country with respect to the RoW results in increased export diversification. Secondly, we show that the increased relative size of the RoW with respect to the country size decreases the country's export diversification. The combined effect of these two forces depends on their relative strength.

To test the predictions of our model empirically, we use a panel of 132 countries, for which we compute relative export variety measures based on product level (HS 6-digit) trade data throughout the 1988–2014 period. Our sample includes 53 low-income economies. We find strong empirical support for the theoretical predictions. In particular, we fnd that – ceteris paribus – a one per cent rise in relative productivity (with respect to the RoW) can be associated with a 0.5 per cent rise in the number of active export lines.[4] In light of the estimated coefficients, the influence of productivity on export variety is stronger than the effect exhibited by cross-country differences in relative country size. At the same time, we find that the importance of technology differences is non-linear and depends on the development stage: it drives diversification at the beginning of the development process, whereas, at higher levels of growth, the expansion of the relative country size plays a major role in promoting export variety growth.

To test the sensitivity of the results against the underlying theoretical framework, we also introduce factor endowment differences between countries (physical capital, human capital, arable land, and petroleum abundance), in addition to productivity differentials. As expected, they do influence export variety, although productivity still acts as the main driver of the diversification process.[5] Our results are robust, with the variety of diversification indices, the economic complexity measures, additional measures of relative country size and estimation methods all having been accounted for. Moreover, while we focus on a setting where each country is assessed with respect to the RoW, we also test the bilateral specification of the panel regressions between country pairs. The results still hold.

The paper is structured as follows: Section 2 includes a literature review and describes our approach with respect to the existing research. Section 3 presents the analytical framework used to derive our empirical specification. The results are presented in Section 4. Section 5 contains a summary and conclusions. The appendix provides a set of complementary materials.

Literature review

Empirical studies on the diversity of economic structures of countries and its relationship with the stages of economic development were initiated by the seminal work of Imbs and Wacziarg (2003), who were the first to estimate the 'general' diversification curve using data on multiple countries. They used sectoral employment and value-added statistics concerning 99 economies at various stages of development to compute the countries' concentration indicators, subsequently combined with their GDP per capita for the 1969-1997 period. As a result of a non-parametric estimation, they obtained a U-shaped diversification curve which, according to their interpretation, reflects two successive stages: at an early stage of development, there is an increase in the diversity of economic structures as GDP per capita increases, while at a later stage, they argue for a reconcentration.[6]

The U-shaped diversification curve reported in IW (2003) is only an empirical observation without any attempts to explain it on the grounds of formal economic theory and it has been questioned in later export diversification studies (e.g. De Benedictis et al., 2009; Parteka and Tamberi, 2013a, see Mau, 2016 for a review). While the low level of economic diversification – typical for underdeveloped economies – can be easily explained (i.e. as associated with the economy's immaturity, the inability to use the production factors efficiently, or the excessive dependence on natural resources), the reconcentration (re-specialisation) tendency is less clear. The first wave of Imbs and Wacziarg's (2003) follow-up papers was essentially a repetition of their methods on other datasets. These studies, with a few exceptions, confirmed the existence of a diversification curve (for example, Klinger and Lederman, 2006 and Koren and Tenreyro, 2007 also achieved a U-shaped pattern based on production, employment and export data). In contrast, more recent empirical studies belonging to the second wave of related research, proposed significant methodological changes: the analyses were conducted using semi-parametric methods and aimed at examining the robustness of the results against changes in diversification measurement methods.[7] For ex-

[4] See Table 1.

[5] This is in line with Shikher (2013).

[6] Additionally, there are many papers that analyse the link between export diversification and economic growth (among others: Agosin, 2007; Herzer and Nowak-Lehnmann, 2006; Hesse, 2008; Gozgor and Can, 2016; Kaitila, 2019), often in terms of a causal relationship between the two. Recently, Mania and Rieber (2019) have studied the relationship between export diversification and *sustainable* growth, focusing on the structural impact of diversification on income elasticities of exports and imports.

[7] The has also been a debate in the literature on the appropriate setting (absolute vs. relative) to the study of export diversification. The early empirical literature is dominated by the absolute approach, where the degree of concentration of a given economy in measured in isolation from global trends and changes in the global basket of manufactured and exported goods. Cadot et al. (2011), Dennis and Shephard (2011), Klinger and Lederman (2011) and Agosin et al. (2012) all represent this trend in the literature. The second wave studies have adopted the relative approach, which makes it possible to refer the process of diversification of a given

ample, (De Benedictis, 2008, 2009) and Parteka (2010) showed that the shape of the estimated curves is rather consistent with the ongoing diversification in the economic development process. Parteka and Tamberi (2013a) showed that the re-specialisation part of the curve is driven by specific, highly-developed countries, that are usually small and/or rich in natural resources.

Important arguments contradicting the U-shaped patterns have also been presented in the economic complexity (EC) literature (including Cristelli et al., 2015; Hausmann et al., 2007, 2019; Hidalgo et al., 2007, 2018; Tacchella et al., 2012, 2013), built upon the Product Space (PS) framework (i.e., the network of relatedness between products: Hausmann and Klinger, 2007; Hidalgo et al., 2007). The EC approach proves empirically that diversification remains important at high levels of income as well. The mechanism is linked to the path-dependence of the export basket evolution – current capabilities (depending on technologies, factors endowments, institutions, etc.) determine the diversification of export structures towards new products because they originate from a re-combination of the current set of production capabilities.[8] The most prosperous economies have the most complex structures (as such, their production capability endowment is high), which enable them to export all products and at various levels of complexity. In a study covering 91 economies, Minondo (2011) builds an index of countries' diversification possibilities based on the commodities in which these countries have a comparative advantage and the proximity of those commodities to other products. He then shows that an indicator which reflects the degree of centrality in the product space is a strong predictor of a country's diversification level.

Another feature of the early literature on the relationship between diversification and economic development is its empirical focus and strong detachment from the economic theory. Some elements of theoretical considerations explaining why a reconcentration of exports is possible are present in the studies of, among others, Cadot et al. (2011) and Mau (2016), who refer to the theory of relative factor abundance. They try to interpret the process of diversification referring to the neoclassical Heckscher-Ohlin (HO) model using the extended version of the Lerner diagram for goods differing in the intensity of production factor use. In Cadot et al. (2011), the existence of multiple cones of diversification means that cross-country differences in factor endowments determine specialisation patterns: rich countries produce goods that are different from those produced by poor countries. The process of economic development (capital accumulation) can be interpreted as a 'travel' across diversification cones, so Cadot et al. (2011) explained the inverted U-shaped relationship between diversification and development as a slow adjustment in the two phases of this 'travel'. Firstly, as countries accumulate capital, they shift from one cone to another, but old lines remain active (which means diversification); then the old lines slowly die, which reduces diversification. However, as recently noted by Lectard and Rougier (2018), this issue may be more complex than the one illustrated by the diversification cone theory. Developing countries face the dilemma of choosing the right diversification strategy: promoting productive transformation by defying or following their comparative advantage. In the light of empirical results obtained by Lectard and Rougier (2018), the former choice (defying comparative advantage) seems to enhance export diversification of middle-income and resource-rich countries but tends to concentrate export baskets in the case of lower-income economies.[9]

However, the theoretical explanations of diversification patterns based purely on the HO setting put aside the potential impact of other factors that may shape diversification as well, such as relative productivity. The important reference point, crucial in our paper, is thus the Ricardian framework which has been often used in the literature to investigate the factors that shape the nature of international specialisation in production and foreign trade.[10] In particular, the evolution of production and export diversification can be well explained by referring to the neoclassical Ricardian framework with the continuum of goods developed by Dornbusch, Fisher and Samuelson (1977). In that model, changes in export variety may be caused by changes in both the relative productivity of a given country, resulting from the differences in the rates of technological progress in the home country and abroad, and changes in the relative size of the country, related to differences in the rates of population growth at home and abroad.

In the classic Ricardian framework, export diversification is driven mainly by technology differences between countries. Mau (2016) argues that the HO setting cannot explain export diversification at the extensive margin but holds only for the intensive margin; i.e., when the range and type of goods are unaffected, and relative output and factor allocations vary. He also uses the EK framework to yield a gravity equation for export diversification at the extensive margin. The model predicts that after controlling for factor costs and geography, a more technologically advanced country exports a wider variety of goods. A dynamic version of the model is presented in Naito (2017), who extends the two-country setting presented in Naito (2012) into a multi-country framework and combines the EK Ricardian model with a continuum of goods with a multi-country AK model by Acemoglu and Ventura (2002). Naito (2017) focuses on the effects of falling trade costs on the countries' growth and the extensive margin of trade over time. He showed that trade liberalisation raises the balanced growth rate and provides a theoretical explanation for why fast-growing countries experience a rise in export variety (which is in line with the findings of Hummels and Klenow, 2005 and Kehoe and Ruhl, 2013). Levchenko and Zhang (2016) extend the EK approach to a multi-sector framework and study the role played by *relative sectoral* productivity differences between countries, finding strong evidence of relative productivity convergence. Recent developments in the Ricardian theory depart from the assumption of uncorrelated productivity parameters across industries towards a setting where 'comparative advantage is correlated across technologically related industries' (Hausmann et al., 2019).

The next section focuses on modelling relative country-level differences in size, productivity and diversification. Hence, our research complements the Ricardian-based comparative advantage perspective where relative technology differentials are studied in the presence of multiple industries (Levchenko and Zhang, 2016). Moreover, the theoretical approach we use should also be considered complementary to the Ricardian-based diversification models presented in Mau (2016) and Naito (2017), and referring to the EK model. We aim to provide a close link between the theoretical model and its empirical validation. The basic assumptions and

country to the trends occurring in the entire sample (De Benedictis et al., 2008, 2009, Basile et al., 2018; Parteka 2013; Parteka and Tamberi, 2013a, 2013b). Our paper is closer to the latter approach.

[8] On the contrary, radical changes (path-defying diversification) take place if countries diversify towards areas of the PS that are unrelated to the initial production basket (Coniglio et al., 2018).

[9] Regolo (2013) also uses the HO setting, but she focuses on the determinants of bilateral differences in export concentration. Specifically, she uses a North–South setting, similar to Romalis (2004), where exports between similarly endowed countries (either South–South or North–North) become more diversified than exports between countries with different factor endowments (South–North).

[10] Its most important extensions include the works of Krugman (1987), who introduced the learning curve, as well as Eaton and Kortum (2002), Alvarez and Lucas (2007), Mau (2016) and Naito (2017), who included firm heterogeneity in their models.

research hypotheses derived from the classic Ricardian framework, used in the empirical part of the paper, are presented below.

Analytical framework

To study the determinants of export diversification, we use the model proposed by Dornbusch, Fisher, and Samuelson (1977, DFS hereafter) as our simple analytical framework. We demonstrate that the extent of export diversification can be related to both the relative level of technology and the relative country size. Our approach differs from the alternative ones (especially from the pure H-O framework) by emphasising the role played by relative productivity, which may be an important explanation of relative export variety patterns and the degree of economic complexity, at least partly reflected in trade diversification.

The DFS model is a traditional neoclassical model based on several simplifying assumptions. It assumes two countries, Home and Foreign, each using only one factor of production – labour – to produce an arbitrary number of goods.[11] In our setting, the latter 'country' represents the Rest of the World (RoW). Each country exhibits constant returns to scale technology, but they differ in the relative amounts of labour required to produce different goods. Relative unit labour requirements for specific goods can be ranked according to the diminishing comparative advantage of the Home country. This generates an incentive for each country to specialise in the production of only a certain set of goods, which, in turn, generates the gains from trade.

Rather than working with a finite number of goods, the model assumes a continuum of goods represented by the unit interval [0,1]. This assumption makes it possible to study the changes in export diversification directly. In this case, the *z*-th good from this interval simply reflects the share of Home country in the total number of tradable goods that are produced in the world economy. This share can be identified as one of the simplest measures of export diversification used in the empirical literature. Therefore, it is also used in our paper as the measure of relative export diversification.[12]

Production of the *z*-th good is feasible in the Home country if its price, equal to the unit cost, is either equal to or lower than the price in the RoW; i.e., $p(z) = a(z)w \leq a^*(z)w^* = p(z)^*$, where $p(z)$ denotes the price of *z*-th good at Home, w denotes the wage level at Home, $p(z)^*$ denotes the price of *z*-th good in the RoW, and w^* the wage level in the RoW.[13] This is the efficient specialisation condition: $\omega \leq A(z)$, where $\omega = w/w^*$, denotes the relative wage and $A(z) = a^*(z)/a(z)$ is the relative unit labour requirement, and $A'(z) < 0$. Given the wage level, it is possible to find the threshold commodity $\tilde{z}$ that will be produced in both countries. However, to find the relative wage we need to specify the demand side of the model.

[11] Extensions of the analysis by DFS (1977) with respect to the demand structure and the number of countries are developed in Wilson (1980). He demonstrates that several sharp comparative statics results are still possible.

[12] The assumptions of the model are quite restrictive, but our goal is to set up a simple base for the subsequent empirical analysis. Several extensions of the DFS framework include elements of heterogeneity in the model (e.g. like in multi-country and multi-product Ricardian models by, among others, Eaton and Kortum, 2002 and Costinot et al., 2012). The assumption of uncorrelated productivity parameters across industries is relaxed in economic complexity frameworks (among others: Hausmann et al., 2019), which put emphasis on technological relatedness and path-dependence of export variety patterns. We take this approach into account in Section 4, using economic complexity measures in the spirit of Hausmann and Klinger (2007), Hidalgo et al. (2007), Hausmann et al. (2013) and Tacchella et al. (2013) as an alternative to a pure export variety index.

[13] Given the basic assumptions of the model, i.e., perfect competition and one factor of production (labour) only, it can be noted that the wage level is equal to the GDP per capita. Hence, the GDP equals the aggregate wage bill, i.e. the product of the wage level and the total number of workers.

The baseline model assumes homogenous and identical preferences in both countries. Specifically, each tradable good receives a fixed fraction of expenditure $b(z)$, which is the same in both countries $b(z) = b(z^*)$. Hence, the fraction of income spent in both countries on Home-produced goods produced can be defined as $\nu(\tilde{z}) = \int_0^{\tilde{z}} b(z)dz > 0$, where: $\nu'(\tilde{z}) = b(\tilde{z}) > 0$, and $0 \leq \nu(\tilde{z}) \leq 1$. Similarly, the fraction of income spent on goods produced in the RoW can be expressed as $1 - \nu(\tilde{z}) = \int_{\tilde{z}}^1 b(z)dz$. In equilibrium, the RoW expenditure on Home-produced goods equals the Home expenditure on goods produced in the RoW: $[1 - \nu(\tilde{z})]wL = \nu(\tilde{z})w^*L^*$.

Therefore, the equilibrium in the baseline DFS model can be described by a set of two conditions: the efficient specialisation condition (1) and the trade balance condition[14] (2), respectively:

$$A(z) = \omega \tag{1}$$

$$B(z,\ L^*/L) = \frac{\nu(z)}{1-\nu(z)}\frac{L^*}{L} \tag{2}$$

where ω denotes the relative wage of the Home country with respect to the RoW.

Firstly, we study the effects of technological progress at Home with respect to the RoW on the extent of export diversification and relative wages, assuming that the relative country size remains constant (i.e., at a given allocation of labour across countries, $d(L^*/L) = 0$). By totally differentiating Eqs. (1) and (2), we obtain, respectively:

$$dA(z) = d\omega - A'(z)dz \tag{1a}$$

$$0 = d\omega - \left[\frac{\nu'(z)}{(1-\nu(z))^2}\frac{L^*}{L}\right]dz \tag{2a}$$

Using the matrix notation, we can write Eqs. (1a) and (2a) as:

$$\begin{bmatrix} 1 & -A(z) \\ 1 & -\dfrac{\nu'(z)}{(1-\nu(z))^2} \end{bmatrix} \begin{bmatrix} d\omega \\ dz \end{bmatrix} = \begin{bmatrix} dA(z) \\ 0 \end{bmatrix} \tag{3}$$

Employing Cramer's rule, we calculate the changes in the extent of export diversification dz and relative wages $d\omega$ resulting from the positive change in the relative productivity $dA(z)$, respectively, as:

$$dz = \left(\frac{\nu'(z)}{(1-\nu(z))^2}\frac{L^*}{L} - A'(z)\right)^{-1} dA(z) > 0 \tag{4}$$

$$d\omega = \frac{\frac{\nu'(z)}{(1-\nu(z))^2}\ \frac{L^*}{L}}{\frac{\nu'(z)}{(1-\nu(z))^2}\ \frac{L^*}{L} - A'(z)} dA(z) > 0 \tag{5}$$

Ceteris paribus, the change in the extent of export diversification dz resulting from increased relative productivity at Home, $dA(z) > 0$, is positive. Also, the change in the relative wage resulting from increased relative productivity at Home $dA(z)$ is positive. Proposition 1 summarises this finding:

Proposition 1. *An increase in the relative productivity in the Home country – with an unchanged relative country size – results in increased Home country export diversification compared to the RoW. This increase is accompanied by an increase in the relative wage (GDP per capita).*[15]

[14] Of course, this is a simplification of the empirical reality that is required to set up the modelling framework.

[15] Note that the relative wage increase is less than proportional to an increase in the relative productivity level.

Secondly, we study the effects of the increased size of the RoW, compared to Home, on the extent of export diversification and relative wages, assuming that the relative level of productivity remains constant (i.e. at a given unit labour requirement level across countries, $dA(z) = 0$). Using the same procedure as above, we obtain:

$$\begin{bmatrix} 1 & -A(z) \\ \frac{1-\nu(z)}{\nu(z)} & -\frac{\nu'(z)}{\nu(z)(1-\nu(z))} \end{bmatrix} \begin{bmatrix} d\omega \\ dz \end{bmatrix} = \begin{bmatrix} 0 \\ d\frac{L^*}{L} \end{bmatrix} \tag{6}$$

and the resulting changes in the extent of export diversification dz and relative wages $d\omega$ are as follows:

$$dz = \left(A'(z)\frac{1-\nu(z)}{\nu(z)} - \frac{\nu'(z)}{\nu(z)(1-\nu(z))}\frac{L^*}{L} \right)^{-1} d\left(\frac{L^*}{L}\right) < 0 \tag{7}$$

$$d\omega = \left(\frac{1-\nu(z)}{\nu(z)} - \frac{\nu'(z)}{\nu(z)(1-\nu(z))A'(z)}\frac{L^*}{L} \right)^{-1} d\left(\frac{L^*}{L}\right) > 0 \tag{8}$$

The change in the extent of export diversification – dz – resulting from the increased relative size of the RoW, $d(L^*/L) > 0$, is negative, *ceteris paribus*. However, the change in the relative wage resulting from the increased relative size of the RoW $d(L^*/L)$ is positive.

Proposition 2. *An increase in the relative size of the RoW (a decrease in the relative size of the Home country) results in decreased Home country export diversification – with the relative technology level remaining unchanged. This increase is accompanied by an increase in the relative wage (GDP per capita).*

Finally, we examine the combined effect of the changes in relative productivity and relative country size on the extent of export diversification. Through a total differentiation of conditions (1) and (2), assuming that $dA(z) > 0$ and $d(L^*/L) > 0$, we obtain:

$$\begin{bmatrix} 1 & -A(z) \\ \frac{1-\nu(z)}{\nu(z)} & -\frac{\nu'(z)}{\nu(z)(1-\nu(z))} \end{bmatrix} \begin{bmatrix} d\omega \\ dz \end{bmatrix} = \begin{bmatrix} dA(z) \\ d\frac{L^*}{L} \end{bmatrix} \tag{9}$$

The resulting changes in the extent of export diversification dz and relative wages $d\omega$ are provided as:

$$dz = \frac{dA(z) - \frac{\nu(z)}{1-\nu}d\ \frac{L^*}{L}}{\frac{\nu'(z)}{(1-\nu(z))^2}\ \frac{L^*}{L} -\ A'(z)} \tag{10}$$

$$d\omega = \frac{\frac{\nu'(z)}{(1-\nu(z))^2}\ \frac{L^*}{L}dA(z) - \frac{\nu(z)}{1-\nu}A'(z)d\ \frac{L^*}{L}}{\frac{\nu'(z)}{(1-\nu(z))^2}\ \frac{L^*}{L} -\ A'(z)} > 0 \tag{11}$$

Proposition 3. *The combined effect of an increase in the relative productivity at Home and an increase in the relative size of the RoW on the extent of export diversification at Home is not clear – technological progress at Home increases it while the increased relative size of the RoW decreases it. The extent of export diversification at Home increases only if the effect of increased relative productivity dominates over the effect of the increased relative size of the RoW, i.e., $dA(z) > \frac{\nu(z)}{1-\nu(z)}d\frac{L^*}{L}$. At the same time, the combined effect of an increase in relative productivity and an increase in the relative size of the RoW on relative wages (GDP per capita) is positive.*

Therefore, the above analysis shows that the relationship between the level of development, measured by GDP per capita, and the extent of export diversification is not as evident as it may seem at first glance. For example, in a situation where the population growth rate in the Home country is slowing down or declines compared to the RoW, the diversification of the country's exports may begin to decline despite a further increase in per capita GDP related to technological progress. Similarly, an increase in productivity in the RoW, which may exceed the increase in productivity in the Home country, may result in a decrease in diversification depending on the changes in the relative worker population in both countries. Table A1 in Appendix A contains a summary of the possible outcomes. The next section is focused on the empirical validation of the model predictions using an international panel dataset.

Empirical verification of the model predictions

The data

To test the model predictions, we use a balanced panel dataset (Table A2 in the Appendix lists the 132 countries included in the sample, which are subsequently classified into high, medium-high, medium-low, and low income groups based on their development level) covering the period between 1988 and 2014. To construct our dependent variable,[16] we use the highest degree of detail available for international comparisons in a panel-data setting, i.e. the HS 6-digit (product-level) mirrored exports statistics from UN Comtrade. With these statistics, we construct country- and time-specific synthetic indices of relative export variety (*REV)*, which, in line with the theoretical model, measure the degree of export diversification compared to the rest of the world – RoW. To address the presence of the lower bound (export variety cannot be negative) in a relative dimension setting, we follow Regolo's (2013) method and use logarithms of a ratio of export variety of a 'home' country i and the RoW:[17]

$$lnREV_{it}^{k} = ln\left(k_{it}/k_t^{RoW}\right) \tag{12}$$

where i denotes the 'home' country, t – time and the *RoW* value is computed as the average for all the countries in the sample minus country i (the same method is applied with respect to the explanatory variables).[18] The set of indices *k={N, Theil, Gini, HH}* refers to alternative measures of export variety/export diversification (their exact definitions and formulas can be found in Table A11 while their comparison to relative diversification measures is provided in Parteka, 2010). Thus, *REV* is firstly calculated based on the number of products exported (REV^N) – which is the closest to the theoretical model – or the inverse (for the ease of interpretation) of export concentration indices: the Theil index (REV^{Theil}), the Gini index (REV^{Gini}), and the Herfindahl–Hirschman Index (REV^{HH}). Additionally, we consider indices that capture the relative aspect of export diversification due to their construction (namely, the Relative Theil index, $REV^{RelTheil}$, and the Relative Gini index, $REV^{RelGini}$) employed by De Benedictis et al. (2009) and Parteka and Tamberi (2013a). Finally, we use the Economic Complexity Index (ECI)[19]

[16] Empirically, export variety and the degree of export diversification have been measured using trade data at various levels of detail: from sector level (Parteka 2010, Parteka and Tamberi, 2013b; De Benedictis et al., 2009) to product level (Dennis and Shepherd, 2011; Hummels and Klenow, 2005; Cadot et al., 2011; Mau, 2016; Parteka and Tamberi 2013a; Basile et al., 2018) to the firm level and the shipment level (Martincus and Carballo, 2008; Hillberry and Hummels, 2008).

[17] Regolo (2013) considers the determinants of bilateral diversification (between country pairs) and focuses on bilateral differences in relative endowments. She is interested in measuring the degree of similarity between trading partners and, thus, adopts the *absolute* differences' measurement. However, we are also interested in the sign of the differences (for instance, if the productivity of a given country is either lower or higher than that of the RoW), so our *REV* variables are not expressed in absolute terms. See also Parteka (2020) for a bilateral panel analysis of export variety differentials.

[18] Consequently, we measure the degree of the country i's export variety compared to the average export variety of all the remaining countries.

[19] To assure the maximum level of concordance with our benchmark indices, we rely on the Economic Complexity Index, ECI, computed using the HS 6-digit product

to confront pure export variety with broader measures of countries' economic composition based on the product space concept (Hausmann and Klinger, 2007; Hidalgo et al., 2007; Hausmann et al., 2013; Tacchella et al., 2013).

In the case of the main explanatory variables – ones directly linked to the theoretical model – we use the data from PWT 9.0 (Feenstra et al., 2015) to measure productivity in terms of output per worker (Y/L) and the number of workers – persons engaged (L) – to proxy the country size.[20] The set of additional explanatory variables (which control for the importance of factor endowments other than labour, bringing our model closer to the HO-based view of international trade) includes physical capital stock (K – from PWT 9.0), territory (T – the land area from WDI) and fuel exports (F – expressed as a percentage of merchandise exports, also from WDI). To control for the quality of labour, we use the *hc* index from PWT 9.0 (based on the Barro and Lee dataset) to compute a human-capital-adjusted measure of employment ($L_{HC} = hc^{*}L$). We also combine the education enrolment ratios from the original Barro and Lee dataset (2016 update), Barro and Lee, 2013[21] with annual data on the number of persons engaged (in mln, from PWT 9.0) to obtain the number of workers with low (L_{low}), medium (L_{med}), and high education level (L_{high}). Alternatively (as a robustness check),[22] we use the data on employment by education from ILO to split the labour force variable (L) into three components: advanced (L_{adv}) intermediate (L_{int}), and basic (L_{bas}). [23] However, such data are obtainable only for the period since 1990 and for a limited subsample of countries.

In line with the theoretical model, productivity is expressed directly concerning the RoW, ln $(\frac{Y}{L}/\frac{Y^{RoW}}{L^{RoW}})$, while all the other explanatory variables are inverted and expressed as RoW concerning the home country, namely $ln(X_t^{RoW}/X_{it})$ *for* $X \in \{L; K; L_{HC}, L_{high}, L_{med}, L_{low}, L_{adv}, L_{int}, L_{bas}, T, F\}$. Table A3 includes the summary statistics while Table A4 and Table A5 [24] in the Appendix show the pairwise correlation coefficients concerning all variables (in logs). Table A12 reports the variables' variation over time.

The empirical model and benchmark estimation results

To test the crucial Proposition 3 of the theoretical model, we regress the measure of the relative export variety of country i in year t on two main factors – productivity and size:

$$lnREV_{it}^{k} = \alpha + \beta_1 ln\left(\frac{Y_{it}}{L_{it}}/\frac{Y^{RoW}}{L^{RoW}}\right) + \beta_2 ln\left(L^{RoW}/L_{it}\right) + D_t + \varepsilon_{it} \quad (13)$$

where REV^k, as in Eq. (12) with k = *{N, Theil, Gini, HH, RelTheil, RelGini}*, alternated with *ECI*, is a function of relative productivity $(\frac{Y_{it}}{L_{it}}/\frac{Y^{RoW}}{L^{RoW}})$ and relative country size (L^{RoW}/L_{it}), with i referring to the countries and t to the time. Table 1 shows the estimations of the simplest models, related to Proposition 1 and 2, in which *REV* is only a function of one of these two factors.

The basic results obtained through OLS[25] (Table 2) with time-fixed effects confirm the key model predictions: export variety is positively related to the countries' relative productivity and negatively related to the expansion of foreign economies (the increase in the size of the RoW). Importantly, this result also holds when considering the diversity, ubiquity, and complexity of products in the countries' exports (reflected in the *ECI* measure) instead of a pure export variety (column 7).[26] As various measures are highly correlated (see Fig.1 showing a strong positive relationship between the *ECI* and *REV*, as well as Table A4 in the Appendix) and provide comparable results (in terms of the sign), we only used REV^N and REV^{Theil} for descriptions and regressions in the main text.

To check the robustness of this result, we include the results obtained with alternative estimators that take into account the specificity of the dependent variable in Table A6 in the Appendix. The upper bound of the measures of export variety, i.e. the number of exported products in the simplest case, is zero while their upper bound is the number of classes (e.g., product categories) registered in trade classification schemes[27]; similarly, export concentration measures are bounded. The calculation of export variety with respect to the RoW (see Eq. (12)) at least partly eliminates the problem of the limited upper bound (theoretically, there is no limit on having the export structure n times more diversified than is the RoW). From below the *REV* approaches zero[28], so we express the dependent variable in logs (Hillberry and Hummels, 2008). Nevertheless, the results obtained through other estimation solutions addressing the problem of the bounded dependent variable, namely Poisson and the *flex* estimates [flexible pseudo maximum likelihood estimation, Silva et al. (2014)], give similar results.

We also check if the results are not driven by specific countries, such as very big economies (China, India), very small states, or petroleum-abundant countries (typically having highly-concentrated exports). Table 3 shows the results of Eq. (13) estimates performed on limited country sets. The results once again remain stable. The benchmark results shown in Table 2 are also robust once we consider income per capita and TFP (from PWT) instead of productivity (Y/L) – see Table A13.

classification (source: The Observatory of Economic Complexity, OEC, Simoes and Hidalgo, 2011; accessed on 11 August, 2020). It is possible to match ECI and our dataset in the case of 106 countries throughout the 1985-2014 period (2120 observations, unbalanced panel due to ECI availability restrictions). The EC measure from 'The Atlas of Economic Complexity' (Center for International Development at Harvard University) could be used as an alternative but its matching with our database results in a lower country coverage than the one obtained using the OEC data.

[20] We use the *rgdpo* and *emp* series.

[21] We use 'Education Attainment for Population Aged 25 and Over' (1950–2010, 5-year averages) from http://www.barrolee.com. We combine their *lu* and *lp* categories into one (low), so our three categories correspond to workers with primary (or lower), secondary, and tertiary schooling.

[22] The results are shown in Table A8 in the Appendix.

[23] Specifically, we use information on 'Labour force by sex, age and education (thousands)'. The categories are as follows: basic – primary and lower secondary education (ISCED-97 1 and 2), intermediate – upper-secondary and post-secondary non-tertiary education (ISCED-97 3 and 4) and advanced – first and second stage of tertiary education (ISCED-97 3 and 4).

[24] In particular, in line with economic complexity literature, countries that are more developed and technologically advanced are characterised by higher export complexity: the *ECI* is positively correlated with productivity (coefficient of correlation equals 0.71 – see Table A5). In the final year of our analysis, the top 5 countries in terms of *ECI* are Japan, Switzerland, Germany, Korea and Sweden while the bottom 5 (Chad, Equatorial Guinea, Gabon, Burma, Guinea) belong to the developing group.

[25] The inclusion of country-fixed effects removes all differences in productivity (Y/L) between countries – see Table A6 in the Appendix – so we decided not to include them into the model and focus on measurable characteristics of countries instead.

[26] Additionally, Table A15 presents the estimates of the diversification model in a dynamic set-up where we also analyse the past technological capabilities. The results show that export diversification is a persistent phenomenon and confirm the role played by relative size and productivity. We also test the potential endogeneity of REV – relative productivity relationship, rejecting it on the basis of the p-value of C statistic.

[27] Silva et al. (2014) propose a 'flexible pseudo maximum likelihood estimation of models for doubly-bounded data method', which may be implemented in Stata using the *flex* command. Their application (a bilateral model for the number of sectors exporting from a given country to a destination, ranging between 0 and 5132 categories), shows that the choice of the estimator may lead to significant differences in assessing the role played by different determinants of the extensive margin of trade.

[28] For example, the log of *REV_N* in our sample ranges from -4.7 to 1.63. For more details, see the summary statistics presented in Table A3 in the Appendix.

Table 1
Relative export variety and productivity (Proposition 1) vs. relative export variety and country size (Proposition 2)
Panel regression (whole sample: 132 countries, 1988-2014)[a], alternative *REV* measures, log-log.

	Dep.var.: Expected sign	(1) $\ln(REV^N)$	(2) $\ln(REV^{Theil})$	(3) $\ln(REV^{Gini})$	(4) $\ln(REV^{HH})$	(5) $\ln(REV^{RelTheil})$	(6) $\ln(REV^{RelGini})$	(7) *ECI*
$\ln(\frac{Y}{L}/\frac{Y^{RoW}}{L^{RoW}})$	(+)	0.504***	0.155***	0.019***	0.487***	0.172***	0.079***	0.661***
		[0.049]	[0.026]	[0.003]	[0.094]	[0.027]	[0.015]	[0.060]
R^2		0.358	0.214	0.234	0.165	0.29	0.224	0.513
Observations		3564	3564	3564	3564	3564	3564	1809
	Dep.var.: Expected sign	(1) $\ln(REV^N)$	(2) $\ln(REV^{Theil})$	(3) $\ln(REV^{Gini})$	(4) $\ln(REV^{HH})$	(5) $\ln(REV^{RelTheil})$	(6) $\ln(REV^{RelGini})$	(7) *ECI*
$\ln(L^{RoW}/L)$	(-)	-0.327***	-0.101***	-0.012***	-0.317***	-0.100***	-0.051***	-0.117*
		[0.030]	[0.013]	[0.002]	[0.046]	[0.019]	[0.009]	[0.064]
R^2		0.364	0.222	0.234	0.169	0.243	0.225	0.029
Observations		3564	3564	3564	3564	3564	3564	1809

Source: authors' calculations
Note. *, ** and *** denote significance at 10%, 5% and 1% level, respectively
[a] – restricted sample (1995-2014, 106 countries due to the *ECI* availability and the matching with our dataset). OLS estimates and robust clustered standard errors are provided in parentheses. Time-fixed effects are included in all models. Dependent variable based on the number of active export lines, *N* (column 1), the inverse of export concentration measures, *Theil, Gini* and *HH* (columns 2-4), the inverse of relative export diversification measures, *RelTheil* and *RelGini* (columns 5-6), Economic complexity index (*ECI* – column 7). RoW denotes the Rest of the World.

Table 2
Relative export variety, productivity and country size (Proposition 3)
Panel regression (whole sample: 132 countries, 1988-2014)[a], alternative *REV* measures, log-log.

	Dep.var.: Expected sign	(1) $\ln(REV^N)$	(2) $\ln(REV^{Theil})$	(3) $\ln(REV^{Gini})$	(4) $\ln(REV^{HH})$	(5) $\ln(REV^{RelTheil})$	(6) $\ln(REV^{RelGini})$	(7) *ECI*
$\ln(\frac{Y}{L}/\frac{Y^{RoW}}{L^{RoW}})$	(+)	0.552***	0.170***	0.021***	0.534***	0.187***	0.087***	0.704***
		[0.031]	[0.021]	[0.003]	[0.080]	[0.022]	[0.013]	[0.056]
$\ln(L^{RoW}/L)$	(-)	-0.358***	-0.111***	-0.013***	-0.347***	-0.111***	-0.056***	-0.201***
		[0.021]	[0.011]	[0.002]	[0.040]	[0.015]	[0.008]	[0.039]
R^2		0.71	0.48	0.51	0.36	0.568	0.477	0.595
Observations		3564	3564	3564	3564	3564	3564	1809

Source: authors' calculations
Note. *, ** and *** denote significance at 10%, 5% and 1% level, respectively
[a] – restricted sample (1995-2014, 106 countries due to the *ECI* availability and the matching with our dataset). OLS estimates and robust clustered standard errors are provided in parentheses. Time-fixed effects are included in all models. Dependent variable based on the number of active export lines, *N* (column 1), the inverse of export concentration measures, *Theil, Gini* and *HH* (columns 2-4), the inverse of relative export diversification measures, *RelTheil* and *RelGini* (columns 5-6), Economic complexity index (*ECI* – column 7). RoW denotes the Rest of the World.

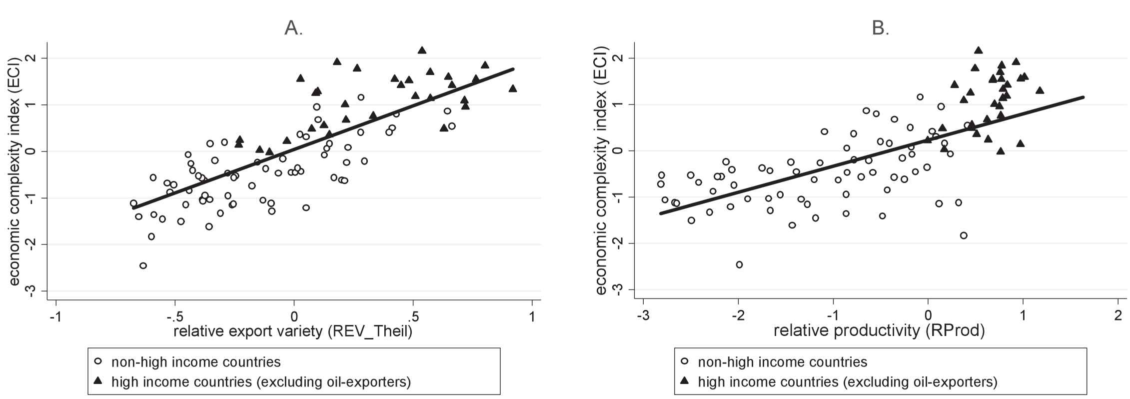

Figure 1. Relative export variety vs. economic complexity
Note. *REV* based on the inverse of export concentration measures, *Theil; ECI* from The Observatory of Economic Complexity; *restricted sample due to the limited availability of the ECI and imperfect matching with our dataset (106 countries, data for 2014).
Source: own elaboration.

4.3. Estimation results – extensions

In the next step, following the 'stages of diversification' literature (Imbs and Wacziarg, 2003; Cadot et. al., 2011), we explore nonlinearity in the export diversification process. The results reported in Table 4 refer to the Eq. (13), estimated in the subgroups of countries based on the World Bank income class that they belonged to in a particular year (we used country-specific classifications to account for the possibility that countries move between income categories over time). The coefficient associated with ln (L^{RoW}/L) is always negative and statistically significant; *ceteris paribus*, at all levels of economic development, the bigger the RoW size, the lower the relative export variety of the Home country. This result holds for both types of *REV* measures (based on *N*

Table 3
Relative export variety, productivity and country size (Proposition 3)
Panel regression (subsamples of countries, 1988-2014), alternative *REV* measures, log-log.

Panel A. Dep.var.: $\ln(REV^N)$	Expected sign	(1) Excluding China and India	(2) Excluding small states (<=300'000)	(3) Excluding small states (<=500'000)	(4) Excluding Fuel>90%	(5) Excluding Fuel>80%	(6) Excluding Fuel>60%
$\ln(\frac{Y}{L}/\frac{Y^{RoW}}{L^{RoW}})$	(+)	0.553***	0.555***	0.555***	0.493***	0.506***	0.516***
		[0.032]	[0.032]	[0.032]	[0.028]	[0.028]	[0.028]
$\ln(L^{RoW}/L)$	(-)	-0.364***	-0.360***	-0.347***	-0.307***	-0.300***	-0.297***
		[0.023]	[0.026]	[0.028]	[0.020]	[0.020]	[0.020]
R^2		0.71	0.698	0.686	0.705	0.718	0.726
Observations		3510	3380	3221	2652	2566	2468
No of countries		130	128	122	125	121	116
Panel B. Dep.var.: $\ln(REV^{Theil})$	**Expected sign**	**(1) Excluding China and India**	**(2) Excluding small states (<=300'000)**	**(3) Excluding small states (<=500'000)**	**(4) Excluding Fuel>90%**	**(5) Excluding Fuel>80%**	**(6) Excluding Fuel>60%**
$\ln(\frac{Y}{L}/\frac{Y^{RoW}}{L^{RoW}})$	(+)	0.170***	0.172***	0.179***	0.199***	0.215***	0.228***
		[0.021]	[0.021]	[0.022]	[0.019]	[0.017]	[0.016]
$\ln(L^{RoW}/L)$	(-)	-0.107***	-0.121***	-0.122***	-0.119***	-0.117***	-0.114***
		[0.012]	[0.012]	[0.013]	[0.009]	[0.008]	[0.008]
R^2		0.461	0.484	0.487	0.574	0.648	0.702
Observations		3510	3380	3221	2652	2566	2468
No of countries		130	128	122	125	121	116

Source: authors' calculations
Note. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. Robust clustered standard errors are provided in parentheses. Time-fixed effects are included in all models. Countries excluded: (1) China and India, (2) Bahamas, Barbados, Belize, Brunei Darussalam, Iceland, Maldives, Sao Tome and Principe, St. Lucia, St. Vincent and the Grenadines; (3) as in (2) plus Bahrain, Cabo Verde, Comoros, Equatorial Guinea, Malta, Qatar, Suriname; (4) Algeria, Angola, Brunei Darussalam, Congo Rep., Iraq, Kuwait, Nigeria, Oman, Qatar, Saudi Arabia, United Arab Emirates, Venezuela; (5) as in (4) plus Bahrain, Gabon, Iran; (6) as in (5) plus Cameroon, Colombia, Ecuador, Norway, Paraguay, Syrian AR, Trinidad and Tobago.

Table 4
Relative export variety, productivity and country size (Proposition 3) – results by income group
Panel regression (132 countries, 1988-2014), log-log.

Panel A. Dep.var.: $\ln(REV^N)$	Expected sign	(1) Low income	(2) Low-middle income	(3) Upper-middle income	(4) High income	(5) High income – excluding petroleum-abundant countries
$\ln(\frac{Y}{L}/\frac{Y^{RoW}}{L^{RoW}})$	?	0.534***	0.383**	0.069	-0.115	0.225
		[0.092]	[0.150]	[0.116]	[0.180]	[0.148]
$\ln(L^{RoW}/L)$	?	-0.464***	-0.361***	-0.312***	-0.288***	-0.193***
		[0.029]	[0.031]	[0.031]	[0.044]	[0.033]
R^2		0.762	0.631	0.635	0.582	0.6
Observations		1110	940	599	915	743
No of countries		53	68	49	44	37
Panel B. Dep.var.: $\ln(REV^{Theil})$	**Expected sign**	**(1) Low income**	**(2) Low-middle income**	**(3) Upper-middle income**	**(4) High income**	**(5) High income – excluding petroleum-abundant countries**
$\ln(\frac{Y}{L}/\frac{Y^{RoW}}{L^{RoW}})$	?	0.140***	0.085	-0.077	-0.408***	-0.13
		[0.029]	[0.072]	[0.114]	[0.070]	[0.140]
$\ln(L^{RoW}/L)$	?	-0.086***	-0.074***	-0.098***	-0.192***	-0.155***
		[0.016]	[0.016]	[0.017]	[0.020]	[0.021]
R^2		0.448	0.248	0.296	0.64	0.563
Observations		1110	940	599	915	743
No of countries		53	68	49	44	37

Source: authors' calculations
Note. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. Robust clustered standard errors are provided in parentheses. Time-fixed effects are included in all models. Division of countries into income groups according to historical (year specific) World Bank's classifications. Column (5) – excluding observations for countries classified as high income but having more than 60% of merchandise exports in fuel products in a given year (Bahrain, Brunei Darussalam, Kuwait, Norway, Oman, Qatar, Saudi Arabia, Trinidad and Tobago, United Arab Emirates).

– panel A and *Theil* – panel B). At the same time, the relationship between *REV* and relative productivity, $\ln(\frac{Y}{L}/\frac{Y^{RoW}}{L^{RoW}})$, is not stable along the growth path. It is positive at low levels of income (column 1), but it then vanishes as countries move to higher levels of development.[29] This suggests that increases in relative labour

[29] For robustness, we match our dataset with HDI from United Nations Development Programme webpage (http://hdr.undp.org/en/data, accessed on 17 August, 2020). This ispossible in the case of 130 countries, in a period from 1990 onwards. As expected, the degree of correlation between income per capita (underlying the classification adopted in Table 4 to group countries by income) and HDI is high and equals 0.7 in our sample. HDI correlates even stronger with productivity (0.82). Table A10 shows the additional set of results obtained with the HDI measure, with countries being classified into ones with low, medium and high level of human development (using the thresholds suggested by the UN). The coefficient associated with the relative size is always negative and statistically significant while the positive relationship between *REV* and relative productivity is confirmed only for countries with low human development level.

Table 5
Relative export variety, productivity and country size (Proposition 3)
Panel regression (whole sample: 132 countries, 1988-2014), with additional control variables, log-log.

Panel A. Dep.var.: $\ln(REV^N)$	Expected sign	(1) basic	(2) controlled for capital	(3) controlled for capital and territory (land)	(4) controlled for resources (fuel)	(5) controlled for capital, land and resources (fuel)	(6) controlled for participation in trade agreements
$\ln(\frac{Y}{L}/\frac{Y^{RoW}}{L^{RoW}})$	(+)	0.552***	0.393***	0.358***	0.474***	0.331***	0.534***
		[0.031]	[0.081]	[0.076]	[0.030]	[0.071]	[0.033]
$\ln(L^{RoW}/L)$	(-)	-0.358***	-0.216***	-0.280***	-0.312***	-0.257***	-0.341***
		[0.021]	[0.071]	[0.067]	[0.023]	[0.063]	[0.020]
$\ln(K^{RoW}/K)$			-0.146**	-0.164**		-0.117*	
			[0.072]	[0.067]		[0.060]	
$\ln(T^{RoW}/T)$				0.091***		0.071***	
				[0.023]		[0.022]	
$\ln(F^{RoW}/F)$					0.025*	0.016	
					[0.013]	[0.013]	
GATT							0.399***
							[0.134]
R^2		0.71	0.715	0.714	0.642	0.66	0.727
Observations		3564	3564	3564	2695	2695	3564
No of countries		132	132	132	127	127	132
Panel B. Dep.var.: $\ln(REV^{Theil})$	**Expected sign**	**(1) basic**	**(2) controlled for capital**	**(3) controlled for capital and territory (land)**	**(4) controlled for resources (fuel)**	**(5) controlled for capital, land and resources (fuel)**	**(6) controlled for participation in trade agreements**
$\ln(\frac{Y}{L}/\frac{Y^{RoW}}{L^{RoW}})$	(+)	0.170***	0.134***	0.110**	0.213***	0.107**	0.164***
		[0.021]	[0.044]	[0.042]	[0.020]	[0.046]	[0.022]
$\ln(L^{RoW}/L)$	(-)	-0.111***	-0.079**	-0.124***	-0.144***	-0.095**	-0.105***
		[0.011]	[0.033]	[0.033]	[0.011]	[0.041]	[0.012]
$\ln(K^{RoW}/K)$			-0.032	-0.045		-0.088**	
			[0.035]	[0.033]		[0.042]	
$\ln(T^{RoW}/T)$				0.064***		0.045***	
				[0.013]		[0.013]	
$\ln(F^{RoW}/F)$					0.046***	0.041***	
					[0.008]	[0.008]	
GATT							0.129*
							[0.070]
R^2		0.478	0.479	0.537	0.559	0.595	0.491
Observations		3564	3564	3564	2695	2695	3564
No of countries		132	132	132	127	127	132

Source: authors' calculations
Note. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. Robust clustered standard errors are provided in parentheses. Time-fixed effects are included in all models. K – capital, T – territory (arable land), F – Fuel (as% of merchandise exports), GATT – participation in the WTO/GATT system.

productivity are associated with increased diversification of export activity only at the beginning of the growth process. Note that the result of a negative relationship between productivity and export variety at high levels of income (column 4, panel B), which might be interpreted as a sign of re-specialisation, is not robust. It does not hold when export variety is measured based on the number of products (column 4, panel A). Additionally, even if it is significant in the case of Theil-based *REV* measure (column 4, panel B), it is evident that this result is driven by rich, petroleum-abundant countries (with a very concentrated export basket). Once we exclude from the rich countries group the economies with export structures dependent on fuel (column 5, panel A and B), $\ln(\frac{Y}{L}/\frac{Y^{RoW}}{L^{RoW}})$ is not among statistically significant determinants of *REV* in high-income countries. This finding is in line with the strand of literature that has questioned the U-shaped export diversification pattern and re-specialisation at higher stages of growth in favour of progressing diversification along the development path (De Benedictis et al., 2009; Parteka and Tamberi, 2013a).[30] The results are also in line with the economic complexity literature (Cristelli et al., 2015; Hausmann et al., 2007, 2019). The general relationship between productivity and the *ECI* is positive (Fig. 1B), which supports the view that the most complex and prosperous economies can export all products, ranging from the most to the least complex ones (which is reflected by the high *ECI* scores) – in Fig. 1B, the group of high-income countries (excluding oil-exporters) is situated above the rest of the sample.

To control for the importance of other endowments, Table 5 shows the results of the estimation referring to the extended model:

$$lnREV_{it}^{k} = \alpha + \beta_1 ln\left(\frac{Y_{it}}{L_{it}}/\frac{Y^{RoW}}{L^{RoW}}\right) + \beta_2 ln\left(L^{RoW}/L_{it}\right) + \beta_3 ln\left(K^{RoW}/K_{it}\right) + \beta_4 ln\left(T^{RoW}/T_{it}\right) + \beta_5 ln\left(F^{RoW}/F_{it}\right) + D_t + \varepsilon_{it} \quad (14)$$

[30] We also consider a quadratic polynomial specification (Table A14): there is no statistically significant U-shaped pattern of export diversification, and as such, re-specialisation at higher levels of development (higher productivity) is not observed in our sample.

Table 6
Relative export variety, productivity and country size (Proposition 3) – labour force adjusted for human capital
Panel regression (whole sample, 1988-2014), log-log.

Panel A. Dep.var.: $\ln(REV^N)$	Expected sign	(1) controlled for human capital-adjusted labour	(2) controlled for a low-skilled labour force	(3) controlled for a medium-skilled labour force	(4) controlled for a highly-skilled labour force
$\ln(\frac{Y}{L}/\frac{Y^{RoW}}{L^{RoW}})$	(+)	0.462*** [0.029]	0.611*** [0.034]	0.393*** [0.032]	0.320*** [0.033]
$\ln(L_{HC}^{RoW}/L_{HC})$		-0.314*** [0.022]			
$\ln(L_{low}^{RoW}/L_{low})$			-0.297*** [0.022]		
$\ln(L_{med}^{RoW}/L_{med})$				-0.301*** [0.021]	
$\ln(L_{high}^{RoW}/L_{high})$					-0.262*** [0.020]
R^2		0.693	0.65	0.695	0.679
Observations		3159	3051	3051	3051
No of countries		117	113	113	113
Panel B. Dep.var.: $\ln(REV^{Theil})$	**Expected sign**	**(1) controlled for human capital-adjusted labour**	**(2) controlled for a low-skilled labour force**	**(3) controlled for a medium-skilled labour force**	**(4) controlled for a highly-skilled labour force**
$\ln(\frac{Y}{L}/\frac{Y^{RoW}}{L^{RoW}})$	(+)	0.153*** [0.021]	0.204*** [0.025]	0.121*** [0.021]	0.095*** [0.023]
$\ln(L_{HC}^{RoW}/L_{HC})$		-0.125*** [0.012]			
$\ln(L_{low}^{RoW}/L_{low})$			-0.106*** [0.014]		
$\ln(L_{med}^{RoW}/L_{med})$				-0.124*** [0.011]	
$\ln(L_{high}^{RoW}/L_{high})$					-0.102*** [0.011]
R^2		0.502	0.412	0.518	0.473
Observations		3159	3051	3051	3051
No of countries		117	113	113	113

Source: authors' calculations
Note. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. Robust clustered standard errors are provided in parentheses. Time-fixed effects are included in all models.

where the following additional explanatory variables are added: relative abundance in the capital (K), land/territory (T), and fuel (F) – all expressed as RoW with respect to the home country. The number of observations available for such extended estimations is lower due to the limited availability of some of the regressors. To address the role played by trade integration and the participation in trade agreements, we also include the information on GATT/WTO membership. The two basic explanatory variables are very robust even if the importance of endowments of the RoW in relation to the given country is controlled for. Relative export variety is positively related to the Home country's relative productivity and negatively related to the size expansion of the RoW.

Moreover, as shown in Table 6, a positive relationship between REV and relative productivity, as well as the negative relationship between REV and the country size hold once the country size measure is adjusted based on the quality of human capital (Jetter and Ramírez Hassan, (2015). To do so, we first substitute L from the model (13) with $\ln(L_{HC})$; the results are shown in column 1 of Table 6. Then (columns 2–4), instead of an aggregate number of workers, we considered the sizes of the labour force elements with three levels of education – low, medium and high – separately. It turns out that regardless of how the relative country size is measured (i.e., either as a total labour force or only a labour force of a certain type), the main model predictions still hold.[31]

[31] The robustness check obtained with alternative indicators of human capital is included in Table A8 in the Appendix.

Finally, even though this paper mainly focuses on a setting where each country is assessed with respect to the RoW, we also tested the bilateral specification of the panel regressions between country pairs (i.e., instead of RoW in model 13, we have an export variety of each country i assessed with respect to country j: $lnREV_{it}^k = ln(k_{it}/k_{jt}) = ln(k_{it}) - \ln(k_{jt})$ for each index of export variety k). The empirical results, reported in Table A8 in the Appendix, are very similar to the estimates of baseline specification (13) reported in Table 2, which confirms their robustness.

Conclusions

Several research contributions have addressed the issue of changes in export variety (export diversification) along the path of economic development. However, they mainly focus on the empirical side of the analysed phenomenon and rarely refer to the theoretical foundations. This makes it difficult to interpret their results. Consequently, we aimed to address the lack of connection between the theory and empirics in most export diversification-economic development studies. We showed that export diversification can be modelled using a testable multi-good Ricardian export diversification model, in which a country's relative export variety is a function of the level of technology and country size – all assessed with respect to the rest of the world (RoW). Relative export diversification can be then viewed as an outcome of two forces: a relative change in productivity due to technological progress, and a relative country size change due to labour force growth.

We developed three crucial propositions that show the relationships between export diversification, relative productivity and relative country size. Firstly, we demonstrated that increased relative productivity of the country, compared to the RoW, results in increased export diversification. Secondly, we showed that the increased relative size of the RoW with respect to the size of the country decreases export diversification. The combined effect of these two forces depends on their relative strength.

We tested the model predictions on the sample of 132 countries (1988–2014), for which we measured the degree of export variety using product-level (HS 6-digit) trade data. We computed several measures of the degree of diversification, based both on conventional indices of export concentration and relative indices of trade diversification. Additionally, we have taken into account the developments in the economic complexity literature, considering the metric of country structure compositions which reflect the diversity, ubiquity, and complexity of the products in the countries' exports. Estimation results confirmed the key model predictions. In particular, we found a robust, positive relationship between export variety and the countries' relative productivity and a negative relationship between export variety and the expansion of foreign economies (i.e., the increase in the size of the RoW). The empirical results were robust against changes in the way of measuring export variety (including the economic complexity index), as well as the inclusion of other control variables (referring to the countries' endowments, human capital), and estimation methods (accounting for the doubly-bounded nature of the dependent variable).

Additionally, we found that the influence of technology differences on export variety was: (i) stronger than the effect of cross-country differences in relative size; (ii) non-linear and drove diversification at the beginning of the development process but then disappeared. This suggests that increases in relative labour productivity are associated with increased diversification of export activity at the beginning of the growth process. This finding has very important economic policy implications: to fight excessive export concentration, developing countries should focus on technological improvements. Technology transfer, which drives productivity change, eventually dies out and can no longer sustain the growth of the export variety; however, it remains a crucial source of diversification in early development. Nonetheless, it must be emphasized that we considered a classic framework in our study. We included some elements of the economic complexity literature into the analysis but possible expansions upon our work could consider technological relatedness (inspired by the multi-country and multi-product Ricardian frameworks: Eaton and Kortum, 2002; Costinot et al., 2012; Hausmann et al., 2019) and further focus on export variety in the economic complexity context (Tacchella et al., 2012, 2013; Hidalgo et al, 2018).

Appendix

Tables A1–A15.

Table A1
Theoretical background – summary of the possible results.

	The early stage of diversification			The late stage of diversification		
	case (1)	case (2)	case (3)	case (4)	case (5)	case (6)
relative productivity	Increase+	increase+	increase+	decrease-	decrease-	decrease-
relative country size	no change 0	increase+	decrease-	no change0	increase+	decrease-
GDP per capita (Absolute wage)	Increase+	unknown+,0,-	increase+	increase+	decrease-	unknown+,0,-
relative wage	Increase+	unknown+,0,-	increase+	decrease-	decrease-	unknown+,0,-
diversification	Increase+	increase+	unknown+,0,-	decrease-	unknown+,0,-	decrease-

Source: own elaboration.

Table A2
List of countries.

Albania	Cyprus	Jordan	Poland
Algeria	Denmark	Kenya	Portugal
Angola	Djibouti	Korea, Rep.	Qatar
Argentina	Dominican Republic	Kuwait	Romania
Australia	Ecuador	Lao PDR	Rwanda
Austria	Egypt, Arab Rep.	Lebanon	Sao Tome and Principe
Bahamas, The	El Salvador	Liberia	Saudi Arabia
Bahrain	Equatorial Guinea	Madagascar	Senegal
Bangladesh	Fiji	Malawi	Sierra Leone
Barbados	Finland	Malaysia	Singapore
Belize	France	Maldives	Spain
Benin	Gabon	Mali	Sri Lanka
Bhutan	Gambia, The	Malta	St. Lucia
Bolivia	Germany	Mauritania	St. Vincent and the Grenadines
Brazil	Ghana	Mauritius	Suriname
Brunei Darussalam	Greece	Mexico	Sweden
Bulgaria	Guatemala	Mongolia	Switzerland
Burkina Faso	Guinea	Morocco	Syrian Arab Republic
Burundi	Guinea-Bissau	Mozambique	Tanzania
Cabo Verde	Haiti	Myanmar	Thailand
Cambodia	Honduras	Nepal	Togo
Cameroon	Hong Kong SAR, China	Netherlands	Trinidad and Tobago
Canada	Hungary	New Zealand	Tunisia
Central African Republic	Iceland	Nicaragua	Turkey
Chad	India	Niger	Uganda
Chile	Indonesia	Nigeria	United Arab Emirates
China	Iran, Islamic Rep.	Norway	United Kingdom
Colombia	Iraq	Oman	United States
Comoros	Ireland	Pakistan	Uruguay
Congo, Dem. Rep.	Israel	Panama	Venezuela, RB
Congo, Rep.	Italy	Paraguay	Vietnam
Costa Rica	Jamaica	Peru	Zambia
Cote d'Ivoire	Japan	Philippines	Zimbabwe

Table A3
Summary statistics.

		n	mean	st. dev.	min	max
Dependent variable - *REV*	$\ln(REV^{N})$	3,564	-0.469	1.130867	-4.70486	1.632755
	$\ln(REV^{Theil})$	3,564	-0.0866	0.404169	-0.733	1.125711
	$\ln(REV^{Gini})$	3,564	-0.00115	0.047598	-0.03822	0.228994
	$\ln(REV^{HH})$	3,564	-1.0107	1.473566	-3.66074	2.697653
	$\ln(REV^{RelTheil})$	3,564	0.467759	0.394503	0.106602	2.716888
	$\ln(REV^{RelGini})$	3,564	1.126896	0.207982	1.000014	2.202718
Relative productivity	$ln(\frac{Y}{L}/\frac{Y^{RoW}}{L^{RoW}})$	3,564	0.597327	1.209355	-2.3087	4.188967
Relative size (labour)	$ln(L^{RoW}/L)$	3,564	1.759584	1.88312	-4.10437	6.283935
Relative capital	$ln(K^{RoW}/K)$	3,564	2.105803	2.229059	-3.63592	7.408419
Relative territory (land)	$ln(T^{RoW}/T)$	3,564	1.732436	2.297397	-2.64491	7.809105
Relative fuel exports	$ln(F^{RoW}/F)$	2,695	1.863281	2.805532	-2.12679	16.87773
Relative human-capital-adjusted labour	(lnL_{HC}^{RoW}/L_{HC})	3,159	1.589927	1.707027	-3.88994	6.008731
Relative human capital (1)	$ln(L_{low}^{RoW}/L_{low})$	3,051	1.667498	1.779511	-4.02353	6.661209
	$ln(L_{med}^{RoW}/L_{med})$	3,051	1.910237	1.824288	-4.33484	6.387411
	$ln(L_{high}^{RoW}/L_{high})$	3,051	2.032483	2.126308	-4.21934	8.31246
Relative human capital (2)	$ln(L_{bas}^{RoW}/L_{bas})$	1,029	1.208749	1.672869	-3.4315	9.004905
	$ln(L_{int}^{RoW}/L_{int})$	1,025	1.153711	1.7088	-3.08108	7.628739
	$ln(L_{adv}^{RoW}/L_{adv})$	1,029	1.207467	1.721873	-3.22222	7.097333

Source: authors' calculations
Note. *n* varies due to the limited availability of some of the variables.

Table A4
Pairwise correlations between alternative measures of relative export variety (*REV*) and economic complexity index (*ECI*)
(whole sample: 132 countries, 1988-2014, n=3564).

	ECI*	ln(REV^{N})	ln(REV^{Theil})	ln(REV^{Gini})	ln(REV^{HH})	ln(REV^{RelTheil})	ln(REV^{RelGini})
ECI*	1						
ln(REV^{N})	0.75	1					
ln(REV^{Theil})	0.73	0.7368	1				
ln(REV^{Gini})	0.74	0.6287	0.9071	1			
ln(REV^{HH})	0.64	0.6951	0.9689	0.8013	1		
ln(REV^{RelTheil})	0.71	0.8218	0.6906	0.6899	0.6113	1	
ln(REV^{RelGini})	0.71	0.6593	0.888	0.9632	0.7881	0.7409	1

Source: authors' calculations
Note: REV based on the number of active export lines, *N*, the inverse of export concentration measures, *Theil, Gini* and *HH*, the inverse of relative export diversification measures, *RelTheil* and *RelGini. ECI* from The Observatory of Economic Complexity; *restricted sample due to the limited availability of the ECI and imperfect matching with our dataset (1995-2014, 106 countries, n=2120).

Table A5
Pairwise correlations between explanatory variables (whole sample: 132 countries, 1988-2014, N=3564*).

	ECI**	$(\frac{Y}{L}/\frac{Y^{RoW}}{L^{RoW}})$	(L^{RoW}/L)	(K^{RoW}/K)	(T^{RoW}/T)	(F^{RoW}/F)	(L^{RoW}_{HC}/L_{HC})	(L^{RoW}_{low}/L_{low})	(L^{RoW}_{med}/L_{med})	$(L^{RoW}_{high}/L_{high})$	(L^{RoW}_{bas}/L_{bas})	(L^{RoW}_{int}/L_{int})	(L^{RoW}_{adv}/L_{adv})
ECI													
$(\frac{Y}{L}/\frac{Y^{RoW}}{L^{RoW}})$	0.71	1.00											
(L^{RoW}/L)	-0.16	0.09	1.00										
(K^{RoW}/K)	-0.64	-0.52	0.77	1.00									
(T^{RoW}/T)	0.08	0.15	0.74	0.53	1.00								
(F^{RoW}/F)	0.03	-0.24	0.23	0.37	0.32	1.00							
(L^{RoW}_{HC}/L_{HC})	-0.29	-0.02	0.98	0.81	0.67	0.21	1.00						
(L^{RoW}_{low}/L_{low})	0.10	0.33	0.94	0.55	0.67	0.14	0.88	1.00					
(L^{RoW}_{med}/L_{med})	-0.38	-0.17	0.91	0.86	0.58	0.22	0.96	0.76	1.00				
$(L^{RoW}_{high}/L_{high})$	-0.51	-0.33	0.83	0.90	0.58	0.25	0.90	0.66	0.92	1.00			
(L^{RoW}_{bas}/L_{bas})	0.01	0.18	0.90	0.71	0.58	0.25	0.86	0.90	0.83	0.69	1.00		
(L^{RoW}_{int}/L_{int})	-0.40	-0.17	0.91	0.87	0.59	0.27	0.94	0.71	0.95	0.89	0.77	1.00	
(L^{RoW}_{adv}/L_{adv})	-0.44	-0.24	0.89	0.90	0.61	0.29	0.93	0.68	0.92	0.94	0.76	0.93	1.00

Source: authors' calculations
Note. * *n* varies due to limited availability of some of the variables. ** This column refers to the restricted sample (1995-2014, 106 countries, n=2120). All variables (except *ECI*) are provided in logs.

Table A6
Relative export variety, productivity and country size (Proposition 3) – robustness check (FE estimates)
Panel regression (whole sample: 132 countries, 1988-2014), alternative *REV* measures, log-log.

	Dep.var.: Expected sign	(1) ln(REV^{N})	(2) ln(REV^{Theil})	(3) ln(REV^{Gini})	(4) ln(REV^{HH})	(5) ln($REV^{RelTheil}$)	(6) ln($REV^{RelGini}$)
$ln(\frac{Y}{L}/\frac{Y^{RoW}}{L^{RoW}})$	(+)	0.053	0.004	0.003**	-0.023	0.058***	0.014*
		[0.085]	[0.020]	[0.002]	[0.095]	[0.011]	[0.008]
$ln(L^{RoW}/L)$	(-)	-1.175***	-0.124***	-0.008	-0.572***	0.03	0.061***
		[0.251]	[0.041]	[0.006]	[0.169]	[0.048]	[0.021]
R^2		0.46	0.037	0.022	0.076	0.287	0.275
Observations		3564	3564	3564	3564	3564	3564

Source: authors' calculations
Note. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. Robust clustered standard errors are provided in parentheses. Time- and country-fixed effects are included in all models. Dependent variables based on the number of active export lines (column 1), the inverse of export concentration measures (columns 2-4), the inverse of relative export diversification (5)-(6)

Table A7
Relative export variety, productivity and country size (Proposition 3)- robustness check (Flex and Poisson estimates)
Panel regression (whole sample: 132 countries, 1988-2014).

	Dep. var.: Expected sign	(1) flex (REV^N)	(2) flex (REV^N)	(3) flex (REV^N)	(4) Poisson (REV^N)	(5) Poisson (REV^N)	(6) Poisson (REV^N)
$ln(\frac{Y}{L}/\frac{Y^{RoW}}{L^{RoW}})$	(+)	0.148***		0.152***	0.393***		0.419***
		[0.015]		[0.007]	[0.042]		[0.028]
$ln(L^{RoW}/L)$	(-)		-0.101***	-0.100***		-0.226***	-0.244***
			[0.008]	[0.004]		[0.022]	[0.019]
R^2							
Observations		3564	3564	3564	3564	3564	3564

Source: authors' calculations
Note. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. Robust clustered standard errors are provided in parentheses. Time-fixed effects are included in all models. Dependent variable: relative number of active export lines (theoretical max=5016, rescaled 0-1). flex- Flexible pseudo maximum likelihood estimation. Poisson –Poisson regression.

Table A8
Relative export variety, productivity and country size (Proposition 3) – robustness check – labour force adjusted for human capital
Panel regression (whole sample, 1988-2014), log-log.

Panel A. Dep. var.: (REV^N)	Expected sign	(1) controlled for basic labour	(2) controlled for intermediate labour	(3) controlled for advanced labour
$ln(\frac{Y}{L}/\frac{Y^{RoW}}{L^{RoW}})$	(+)	0.389***	0.268***	0.246***
		[0.031]	[0.031]	[0.031]
$ln(L^{RoW}_{bas}/L_{bas})$		-0.155***		
		[0.021]		
$ln(L^{RoW}_{int}/L_{int})$			-0.156***	
			[0.020]	
$ln(L^{RoW}_{adv}/L_{adv})$				-0.156***
				[0.021]
R^2		0.667	0.706	0.688
Observations		1029	1025	1029
No of countries		110	109	110
Panel B. Dep. var.: (REV^{Theil})	**Expected sign**	**(1) controlled for low-skilled labour**	**(2) controlled for medium-skilled labour**	**(3) controlled for high-skilled labour**
$ln(\frac{Y}{L}/\frac{Y^{RoW}}{L^{RoW}})$	(+)	0.246***	0.161***	0.151***
		[0.030]	[0.031]	[0.034]
$ln(L^{RoW}_{bas}/L_{bas})$		-0.106***		
		[0.016]		
$ln(L^{RoW}_{int}/L_{int})$			-0.116***	
			[0.012]	
$ln(L^{RoW}_{adv}/L_{adv})$				-0.102***
				[0.013]
R^2		0.495	0.564	0.493
Observations		1029	1025	1029
No of countries		110	109	110

Source: authors' calculations
Note. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. Robust clustered standard errors are provided in parentheses. Time-fixed effects are included in all models.

Table A9
Relative export variety, productivity and country size (Proposition 3) – bilateral specification
Bilateral panel regression (whole sample: 132 country pairs, 1988-2014), alternative *REV* measures, log-log.

	Dep. var.: Expected sign	(1) $\ln(REV^N)$	(2) $\ln(REV^{Theil})$	(3) $\ln(REV^{Gini})$	(4) $\ln(REV^{HH})$
$\ln(\frac{Y^i}{L^i}/\frac{Y^j}{L^j})$	(+)	0.551***	0.167***	0.021***	0.527***
		[0.003]	[0.002]	[0.000]	[0.007]
$\ln(L^j/L^i)$	(-)	-0.359***	-0.110***	-0.013***	-0.344***
		[0.002]	[0.001]	[0.000]	[0.004]
R^2		0.69	0.474	0.509	0.353
Observations		452790	452790	452790	452790

Source: authors' calculations
Note. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. OLS estimates and robust clustered standard errors are provided in parentheses. Time-fixed effects are included in all models. Dependent variable based on the number of active export lines, *N* (column 1), the inverse of export concentration measures, *Theil, Gini* and *HH* (columns 2-4).

Table A10
Relative export variety, productivity and country size (Proposition 3) – results by the human development group
Panel regression (132 countries, 1988-2014), log-log.

Dep. var.: ln(REV^{Theil})	Expected sign	(1) Low human development	(2) Middle human development	(3) High human development
$\ln(\frac{Y}{L}/\frac{Y^{RoW}}{L^{RoW}})$	?	0.150***	-0.036	0.081
		[0.037]	[0.063]	[0.070]
$\ln(L^{RoW}/L)$	?	-0.085***	-0.091***	-0.145***
		[0.018]	[0.015]	[0.019]
R^2		0.426	0.379	0.394
Observations		997	680	1376
No of countries		56	57	74

Source: authors' calculations
Note. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. Robust clustered standard errors are provided in parentheses. Time-fixed effects are included in all models. The division of countries into human development groups according to UN thresholds based on Human Development Index (low: HDI< 0.55; medium: HDI>0.55 and HDI<0.7; high: HDI>0.7).

Table A11
Indices of specialisation (concentration) used to compute REV.

	Formula	where:
Number of active export lines	$N_{it} = \sum_{j=1}^{n} \mu_{ijt}$	$\mu_{ijt} = \{0, 1\}$, $\mu_{ijt} = 1$ if $x_{ijt} > 0$
Theil index	$Theil_{it} = \frac{1}{n}\sum_{j=1}^{n} \frac{x_{ijt}}{\overline{x_{it}}} ln(\frac{x_{ijt}}{\overline{x_{it}}})$	$\overline{x_{it}} = \frac{\sum_{j=1}^{n} x_{jt}}{n}$
Gini index	$Gini_{it} = 1 - (\sum_{j=1}^{n} q_{ijt} - q_{i(j-1)t})/n$	$q_{imt} = \sum_{j=1}^{m} \frac{x_{ijt}}{X_{it}}, \forall m = 1, \ldots, n$
Herfindahl–Hirschman index	$HH_{it} = \frac{\sum_{j=1}^{n} (\frac{x_{ijt}}{X_{it}})^2 - \frac{1}{n}}{1-\frac{1}{n}}$	$X_{it} = \sum_{j=1}^{n} x_{jt}$

Note: *i* – country, *j* - product, t – time, x- export value, *X* – total country exports, *q*- cumulative export share

Table A12
Explanatory variables – time trend estimates ($y_t = \alpha + \beta t + \varepsilon$).

y:	Productivity (Y/L)	Size (L)	Capital (K)	territory (arable land, T)	Fuel (F)	Minerals (M)	Human capital (HC)
T	778.573***	0.303*	61922.214***	-25.394	0.239***	0.059***	0.020***
	[69.328]	[0.158]	[10475.149]	[3649.013]	[0.068]	[0.008]	[0.002]
R^2	0.034	0.001	0.01	0	0.004	0.017	0.05
N	3564	3564	3564	3564	2782	3501	3159

Note: constant included, not reported. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. Standard errors are provided in parentheses.

Table A13
Proposition 3 – robustness check: income per capita and TFP instead of productivity (Y/L)
Panel regression (whole sample: 132 countries, 1988-2014), alternative *REV* measures, log-log.

A.	Dep. var.: Expected sign	(1) ln(REV^{N})	(2) ln(REV^{Theil})	(3) ln(REV^{Gini})	(4) ln(REV^{HH})	(5) ln($REV^{RelTheil}$)	(6) ln($REV^{RelGini}$)
$ln(\frac{GDPpc}{GDPpc^{RoW}})$	(+)	0.527***	0.166***	0.020***	0.525***	0.180***	0.084***
		[0.030]	[0.020]	[0.003]	[0.078]	[0.021]	[0.012]
$ln(L^{RoW}/L)$	(-)	-0.341***	-0.105***	-0.013***	-0.331***	-0.105***	-0.053***
		[0.020]	[0.011]	[0.002]	[0.040]	[0.015]	[0.008]
R^2		0.719	0.498	0.531	0.377	0.582	0.492
Observations		3564	3564	3564	3564	3564	3564
B.	Dep. var.: Expected sign	(1) ln(REV^{N})	(2) ln(REV^{Theil})	(3) ln(REV^{Gini})	(4) ln(REV^{HH})	(5) ln($REV^{RelTheil}$)	(6) ln($REV^{RelGini}$)
$ln(\frac{TFP}{TFP^{RoW}})$	(+)	0.930***	0.302***	0.042***	0.884***	0.403***	0.187***
		[0.126]	[0.063]	[0.008]	[0.245]	[0.066]	[0.037]
$ln(L^{RoW}/L)$	(-)	-0.295***	-0.125***	-0.018***	-0.369***	-0.149***	-0.076***
		[0.029]	[0.017]	[0.002]	[0.064]	[0.025]	[0.012]
R^2		0.499	0.341	0.421	0.235	0.486	0.418
Observations		2457	2457	2457	2457	2457	2457

Source: authors' calculations
Note. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. Robust clustered standard errors are provided in parentheses. Time-fixed effects are included in all models. Dependent variables based on the number of active export lines (column 1), the inverse of export concentration measures (columns 2-4), the inverse of relative export diversification (5)-(6)

Table A14
Relative export variety, productivity and country size (Proposition 3) –quadratic polynomial specification
Bilateral panel regression (whole sample: 132 country pairs, 1988-2014), alternative *REV* measures.

	Dep.var.: Expected sign	(1) $\ln(REV^{N})$	(2) $\ln(REV^{Theil})$	(3) $\ln(REV^{Gini})$	(4) $\ln(REV^{HH})$
RelProd	(+)	0.524***	0.152***	0.026***	0.388**
		[0.052]	[0.042]	[0.005]	[0.150]
$RelProd^2$	?	-0.019	-0.012	0.003	-0.097
		[0.025]	[0.018]	[0.002]	[0.066]
$\ln(L^{Row}/L)$	(-)	-0.359***	-0.111***	-0.013***	-0.351***
		[0.020]	[0.011]	[0.002]	[0.039]
R^2		0.711	0.479	0.523	0.369
Observations		3564	3564	3564	3564

Source: authors' calculations
Note. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. OLS estimates and robust clustered standard errors are provided in parentheses. *RelProd*- relative productivity (with respect to RoW). Time-fixed effects are included in all models. Dependent variable based on the number of active export lines, *N* (column 1), the inverse of export concentration measures, *Theil, Gini* and *HH* (columns 2-4).

Table A15
Relative export variety, productivity and country size – dynamic specifications
Panel regression (whole sample: 132 countries, 1988-2014), alternative *REV* measures, log-log.

	Dep.var.: Expected sign	(1) $\ln(REV^{N})$	(2) $\ln(REV^{Theil})$	(3) $\ln(REV^{Gini})$	(4) $\ln(REV^{HH})$	(5) $\ln(REV^{N})$	(6) $\ln(REV^{Theil})$	(7) $\ln(REV^{Gini})$	(8) $\ln(REV^{HH})$
$ln\ (REV)_{t-1}$	(+)	0.832***	0.798***	0.891***	0.764***				
		[0.014]	[0.019]	[0.014]	[0.023]				
$ln(\frac{Y}{L}/\frac{Y^{RoW}}{L^{RoW}})_{t-1}$	(+)					0.523***	0.168***	0.021***	0.528***
						[0.008]	[0.004]	[0.001]	[0.018]
$ln(L^{RoW}/L)_{t-1}$	(-)					-0.338***	-0.111***	-0.014***	-0.348***
						[0.005]	[0.002]	[0.000]	[0.009]
R^2		0.98	0.98	0.99	0.96	0.71	0.48	0.52	0.36
Observations		3300	3300	3300	3300	3300	3300	3300	3300
Under-identification		571.6939	441.6467	346.0912	405.3863	1594.101	1594.101	1594.101	1594.101
Under-identification (p-value)		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Weak identification		4799.072	2459.901	3402.019	1762.531	312307.9	312307.9	312307.9	312307.9

Source: authors' calculations
Note. *, ** and *** denote significance at 10%, 5% and 1% level, respectively. IV estimates and robust standard errors are provided in parentheses. Time-fixed effects are included in all models; country-fixed effects are included in models 1-4. Dependent variables based on the number of active export lines (column 1,5), the inverse of export concentration measures (columns 2-4 and 6-8). Under-identification refers to the Kleibergen-Paap rk LM test statistic, where a rejection of the null indicates that the instruments are not under-identified. Weak identification refers to the Kleibergen-Paap rk Wald F statistic test for the presence of weak instruments (with corresponding Stock-Yogo, 2005 critical values). The rule of thumb: the statistic should be >= 10 for weak identification not to be a problem (Staiger and Stock, 1997).

References

Acemoglu, D., Ventura, J., 2002. The world income distribution. The Quarterly Journal of Economics 117 (2), 659–694.
Agosin, M.R., 2007. Export Diversification and Growth in Emerging Economies. Working Paper No.233 2007. Department of Economics, University of Chile, Chile.
Agosin, M.R., Alvarez, R., Bravo-Ortega, C., 2012. Determinants of export diversification around the world: 1962–2000. The World Economy 35 (3), 295–315.
Alvarez, F., Lucas, R.E., 2007. General equilibrium analysis of Eaton-Kortum model of international trade. Journal of Monetary Economics 54 (6), 1726–1768.
Balavac, M., Pugh, G., 2016. The link between trade openness, export diversification, institutions and output volatility in transition countries. Economic Systems 40 (2), 273–287.
Barro, R.J., Lee, J.W., 2013. A new data set of educational attainment in the world, 1950–2010. Journal of Development Economics 104, 184–198.
Basile, R., Parteka, A., Pittiglio, R., 2018. Export diversification and economic development: a dynamic spatial data analysis. Review of International Economics 26 (3), 634–650.
Cadot, O., Carrère, C., Strauss-Kahn, V., 2011. Export diversification: What's behind the hump? Review of Economics and Statistics 93 (2), 590–605.
Coniglio, N.D., Vurchio, D., Cantore, N., Clara, M., 2018. On the evolution of comparative advantage: path-dependent versus path-defying changes. SERIES Working Papers N. 01/2018 Available at https://ssrn.com/abstract=3136471 .
Costinot, A., Donaldson, D., Komunjer, I., 2012. What goods do countries trade? A quantitative exploration of Ricardo's ideas. The Review of economic studies 79 (2), 581–608. http://dave-donaldson.com/wp-content/uploads/2015/12/CDK_Restud2012.pdf.
Cristelli, M., Tacchella, A., Pietronero, L., 2015. The heterogeneous dynamics of economic complexity. PloS one 10 (2), e0117174.
De Benedictis, L., Gallegati, M., Tamberi, M., 2008. Semiparametric analysis of the specialization–income relationship. Applied Economics Letters 15 (4), 301–306.
De Benedictis, L., Gallegati, M., Tamberi, M., 2009. Overall trade specialization and economic development: countries diversify. Review of World Economics 145 (1), 37–55.
Dennis, A., Shepherd, B., 2011. Trade facilitation and export diversification. The World Economy 34 (1), 101–122.
Dornbusch, R., Fischer, S., Samuelson, P.A., 1977. Comparative advantage, trade, and payments in a Ricardian model with a continuum of goods. The American Economic Review 67 (5), 823–839.
Eaton, J., Kortum, S., 2002. Technology, geography, and trade. Econometrica 70 (5), 1741–1779.
Feenstra, R.C., Inklaar, R., Timmer, M.P., 2015. The next generation of the Penn World Table. American Economic Review 105 (10), 3150–3182 available for download at www.ggdc.net/pwt.
Gozgor, G., Can, M., 2016. Effects of the product diversification of exports on income at different stages of economic development. Eurasian Business Review 6 (2), 215–235.
Haddad, M., Lim, J.J., Pancaro, C., Saborowski, C., 2013. Trade openness reduces growth volatility when countries are well diversified. Canadian Journal of Economics/Revue canadienne d'économique 46 (2), 765–790.
Hausmann, R., Hidalgo, C.A., Bustos, S., Coscia, M., Chung, S., Jimenez, J., Simoes, A., Yildirim, M., 2013. The Atlas of Economic Complexity. MIT Press, Cambridge MA pp. 978-26252542-8.
Hausmann, R., Klinger, B., 2007. *The Structure of the Product Space and the Evolution of Comparative Advantage* (No. 146). Center for International Development at Harvard University.

Hausmann, R., Hidalgo, C., Stock, D., & Yildirim, M. A. (2019). Implied comparative advantage. HKS Working Paper No. RWP14-003, Available at http://dx.doi.org/10.2139/ssrn.2410427.
Herzer, D., Nowak-Lehnmann, D, F., 2006. What does export diversification do for growth? An econometric analysis. Applied economics 38 (15), 1825–1838.
Hesse, H., 2008. Export diversification and economic growth. In: Newfamer, R., Shaw, W., Walkenhorst, P. (Eds.), Breaking into new markets: Emerging lessons for export diversification. The World Bank, Washington, DC, pp. 55–80.
Hidalgo, C.A., Klinger, B., Barabási, A.L., Hausmann, R., 2007. The product space conditions the development of nations. Science 317 (5837), 482–487.
... & Hidalgo, C.A., Balland, P.A., Boschma, R., Delgado, M., Feldman, M., Frenken, K., Neffke, F., 2018. The principle of relatedness. In: International conference on complex systems, Cham. Springer, pp. 451–457.
Hillberry, R., Hummels, D., 2008. Trade responses to geographic frictions: A decomposition using micro-data. European Economic Review 52 (3), 527–550.
Hummels, D., Klenow, P.J., 2005. The variety and quality of a nation's exports. American Economic Review 95 (3), 704–723.
Imbs, J., Wacziarg, R., 2003. Stages of diversification. American Economic Review 93 (1), 63–86.
Jetter, M., Ramírez Hassan, A., 2015. Want export diversification? Educate the kids first. Economic Inquiry 53 (4), 1765–1782.
Kaitila, V., 2019. Export product range and economic growth in the EU countries. The World Economy 42 (3), 711–737.
Kehoe, T.J., Ruhl, K.J., 2013. How important is the new goods margin in international trade? Journal of Political Economy 121 (2), 358–392.
Klinger, B., Lederman, D., 2006. Diversification, innovation, and imitation inside the global technological frontier (World Bank Policy Research Working). Washington, DC: World Bank Paper No. 3872.
Koren, M., Tenreyro, S., 2007. Volatility and development. Quarterly Journal of Economics 122 (1), 243–287.
Krugman, P., 1987. The narrow moving band, the Dutch disease, and the competitive consequences of Mrs. Thatcher: Notes on trade in the presence of dynamic scale economies. Journal of Development Economics 27 (1-2), 41–55.
Lectard, P., Rougier, E., 2018. Can developing countries gain from defying comparative advantage? Distance to comparative advantage, export diversification and sophistication, and the dynamics of specialization. World Development 102, 90–110.
Levchenko, A.A., Zhang, J., 2016. The evolution of comparative advantage: Measurement and welfare implications. Journal of Monetary Economics 78, 96–111.
Martincus, C.V., Carballo, J., 2008. Is export promotion effective in developing countries? Firm-level evidence on the intensive and the extensive margins of exports. Journal of International Economics 76 (1), 89–106.
Mania, E., Rieber, A., 2019. Product export diversification and sustainable economic growth in developing countries. Structural Change and Economic Dynamics 51, 138–151.
Mau, K., 2016. Export diversification and income differences reconsidered: The extensive product margin in theory and application. Review of World Economics 152 (2), 351–381.
Minondo, A., 2011. Does comparative advantage explain countries' diversification level? Review of World Economics 147 (3), 507–526.
Naito, T., 2012. A Ricardian model of trade and growth with endogenous trade status. Journal of International Economics 87 (1), 80–88.
Naito, T., 2017. An Eaton–Kortum model of trade and growth. Canadian Journal of Economics/Revue canadienne d'économique 50 (2), 456–480.
Newfamer, R., Shaw, W., Walkenhorst, P. (Eds.), 2009. Breaking into new markets: Emerging lessons for export diversification. The World Bank.
Parteka, A., 2010. Employment and export specialisation along the development path: some robust evidence. Review of World Economics 145 (4), 615–640.
Parteka, A., Tamberi, M., 2013a. Product diversification, relative specialisation and economic development: Import-export analysis. Journal of Macroeconomics 38, 121–135.
Parteka, A., Tamberi, M., 2013b. What determines export diversification in the development process? Empirical assessment. The World Economy 36 (6), 807–826.
Parteka, A., 2020. What drives cross-country differences in export variety? A bilateral panel approach. Economic Modelling 92, 48–56.
Regolo, J., 2013. Export diversification: How much does the choice of the trading partner matter? Journal of International Economics 91 (2), 329–342.
Romalis, J., 2004. Factor proportions and the structure of commodity trade. American Economic Review 94 (1), 67–97.
Shikher, S., 2013. Determinants of Trade and Specialization in the Organisation for Economic Co-Operation and Development Countries. Economic Inquiry 51 (1), 138–158.
Silva, J.S., Tenreyro, S., Wei, K., 2014. Estimating the extensive margin of trade. Journal of International Economics 93 (1), 67–75.
Tacchella, A., Cristelli, M., Caldarelli, G., Gabrielli, A., Pietronero, L., 2012. A new metrics for countries' fitness and products' complexity. Scientific reports 2, 723.
Tacchella, A., Cristelli, M., Caldarelli, G., Gabrielli, A., Pietronero, L., 2013. Economic complexity: conceptual grounding of new metrics for global competitiveness. Journal of Economic Dynamics and Control 37 (8), 1683–1691.
Wilson, C.A., 1980. On the general structure of Ricardian models with a continuum of goods: applications to growth, tariff theory, and technical change. Econometrica: Journal of the Econometric Society 1675–1702.