LLEE Working Paper Series

EXPORT DIVERSIFICATION AND ECONOMIC DEVELOPMENT: A DYNAMIC SPATIAL DATA ANALYSIS

Roberto Basile   *Aleksandra Parteka*   *Rosanna Pittiglio*

Export diversification and economic development: a dynamic spatial data analysis

Roberto Basile,[*] Aleksandra Parteka[**] and Rosanna Pittiglio[***]

Abstract

This paper contributes to the empirical literature on the relationship between 'export variety' (export diversification) and economic development by relaxing the assumption of cross-country independence and allowing for spatial diffusion of shocks in observed and unobserved factors. Export variety is measured for a balanced panel of 114 countries (1992-2012) using very detailed information on their exports (HS 6-digit product level). The estimation results of a dynamic spatial panel data model confirm the relevance of spatial network effects: indirect effects (spatial spillovers) strongly reinforce direct effects, while spatial proximity to large countries accelerates the diversification process. These results are robust to the choice of the weights matrix (an inverse-distance matrix, an exponential distance matrix and a matrix based on bilateral trade flows are used).

Keywords: export diversification, economic development, panel spatial data models
JEL codes: F14, F43, C31, O11

[*] LUISS Lab, Rome and Second University of Naples, Department of Economics (Capua - Caserta, Italy); roberto.basile@unina2.it
[**]Gdansk University of Technology, Faculty of Management and Economics (Gdansk, Poland); aparteka@zie.pg.gda.pl (corresponding author).
[***]Second University of Naples, Department of Economics (Capua - Caserta, Italy); rosanna.pittiglio@unina2.it

1 Introduction

Diversification paths during the process of economic development is a topic that has attracted the attention of many economists (Imbs & Wacziarg 2003; Koren & Tenreyro 2007; Klinger & Lederman 2006; Cadot et al. 2011; Cadot et al. 2013; Minondo 2011; Parteka & Tamberi 2013a; Parteka & Tamberi 2013b; Mau 2016). Diversifying exports is one of the main strategies that a country may follow to reduce uncertainty (di Giovanni & Levchenko 2011; Koren & Tenreyro 2007; Koren & Tenreyro 2013). This ability is especially crucial in the case of developing countries, which are typically characterised by low diversification of their economic structure (Amurgo-Pacheco & Pierola 2008; Carrere & Strauss-Kahn 2014). From a theoretical point of view, increasing the variety of goods produced is expected to exert a positive impact on productivity and economic growth (as shown for instance in models of 'expanding product variety' (Barro & Sala-i-Martin 2004: 285-315; Grossman & Helpman 1991:43-83; Grossman & Helpman 1991). Consequently, it is not surprising that the topic of evolving diversification along the path of growth has been widely explored.

Discussion so far has mainly regarded the relationship between GDP *per capita* and levels of diversification of economic activity. Some authors (Imbs & Wacziarg 2003; Koren & Tenreyro 2007; Klinger & Lederman 2006; Cadot et al. 2011) argued that in the first stage, at low levels of income, growth goes in line with an increase in the level of diversification; however, once countries reach a certain level of income, further growth is accompanied by re-concentration.[1] (De Benedictis et al. 2008; De Benedictis et al. 2009; Parteka 2010; Parteka & Tamberi 2013a; Parteka & Tamberi 2013b) show scepticism about the robustness of these patterns, correcting conventional, absolute measures of product diversity and find a nonlinear but monotonically decreasing trend reflecting progressive relative de-specialization along the path of economic growth. More recently, (Mau 2016) has stressed that the above-cited non-monotonic hump-shaped pattern is mainly due to an omitted log-transformation of the income variable, as well as sample selection bias and lack of control variables[2].

By focusing on measurement issues (absolute vs. relative measures of export diversification)[3], the functional form of the model (linear vs. quadratic) and other model specification issues (log-transformation, dynamic specification and so on), the empirical literature has totally neglected another important source of bias, namely the existence of cross-country (or spatial) dependence in the data-generating process. Indeed, all the aforementioned studies analyse the relationship between trade diversification and economic development under the (implicit) assumption of spatial independence.[4] In other words, they do not consider any kind of spatial contagion among countries in the specialization process. This is quite surprising, given the strong links between countries involved in the

---

[1] Theoretically, such a situation can take place when countries are 'travelling across multiple cones of diversification' (Deardorff 2000; Schott 2003; Cadot et al. 2011) . Countries initially diversify at the extensive margin, but when a high level of development is reached it is more profitable to abandon the production of labour-intensive goods and, thus, re-specialize.

[2] (Mau 2016) also provides a relevant theoretical illustration of the link between diversification and economic growth based on (Eaton & Kortum 2002) Ricardian framework: richer countries are likely to export more goods because their superior production techniques endow them with an absolute advantage in global markets, and there is no re-specialization.

[3] Absolute measures (such as Herfindahl index or the Theil index), based on indices of concentration or inequality, were used by (Imbs & Wacziarg 2003; Koren & Tenreyro 2007; Cadot et al. 2011; Agosin et al. 2012; Klinger & Lederman 2006). Relative measures were employed by (De Benedictis et al. 2008; De Benedictis et al. 2009; Parteka & Tamberi 2013a; Parteka & Tamberi 2013b, Mau 2016).

[4] Some authors only take into account the role played by distance between trade partners, e.g. (Agosin et al. 2012) consider the GDP-weighted average distance of each country from its trading partners. (Dennis & Shepherd 2011) take into account the distance between the exporting country and Germany. These measures proxy for transportation costs.

global trade network (de Benedictis & Tajoli 2011; Chaney 2014) and the network structure of economic output (Hausmann & Hidalgo 2011).

Several terms are used in the literature to describe the phenomenon of the interaction between agents (e.g. countries) being shaped by geography: spatial diffusion, spatial contagion, spatial spillover effects and external effects. Leaving aside other disciplines (such as sociology or urban studies), the main areas of application of these concepts in economics include: economic geography and agglomeration economics (Krugman 1991; Fujita et al. 1999; Fujita & Thisse 2002; Duranton & Puga 2004; Glaeser 2008), the spatial diffusion of knowledge, technology and innovation (Comin et al. 2012; Ertur & Koch 2007; Ertur & Koch 2011) and mechanisms of contagion in financial markets (Allen & Gale 2000).

What kind of channels can lead to similar patterns of export structure (in particular, the level of export diversification) among countries close to each other in geographical and/or economic terms[5]? The first obvious channel is trade itself.[6] Whatever its driving force (differences in endowments in the Heckscher-Ohlin framework, differences in productivity in the Ricardian framework, or others), international trade inevitably leads to the creation of ties among countries and to cross-country interdependence. In particular, an important reference point for the study of diversification dynamics is still the New Trade Theory (Krugman 1995; Neary 2009), which explains why similar countries trade intensively, exploiting economies of scale and drawing utility gains from access to a wider variety of goods ('love of variety'). Useful insights into possible transmission channels are also provided by endogenous growth models with international R&D spillovers, imitation of innovation and technology diffusion (Aghion & Howitt 1997; Howitt 2000; Grossman & Helpman 1991; Coe et al. 1997) especially in a Schumpeterian multi-country setting (Ertur & Koch 2011).

In this paper, we address this issue and contribute to the export diversification literature by removing the assumption of spatial independence. Specifically, we study how export variety evolves as a function of economic development (GDP *per capita*) in the presence of spatial contagion effects. In this way, we control for the fact that a shock in the level of development of a country may affect not only the degree of specialization of this country (direct effect), but also that of its neighbours (indirect effect). To our knowledge, this is the first paper studying the relationship between trade diversification and economic development from a spatial econometrics perspective. We employ a dynamic spatial panel data model and consider three alternative weight matrices (an inverse-distance matrix, an exponential distance matrix and one based on bilateral trade flows), which allow us to show that a trade network is a more important driver of spillovers than simple geographical distance.

The remainder of the paper is structured as follows. Section 2 discusses the properties of the data and presents some descriptive evidence. Section 3, being the core of our paper, outlines the empirical model and the methodology and presents the results. Finally, Section 4 concludes.

---

[5] Two countries can be located far away in geographical terms but characterized by similar socio-economic, cultural or institutional settings (for instance, due to a common colonial past), which raises their indirectly measured proximity. This will be taken into account by considering alternative interaction matrices.

[6] Some diversification studies only address this issue indirectly by including in the set of additional explanatory variables participation in common regional trade agreements (Parteka & Tamberi 2013b).

2 Data and descriptive evidence

In our empirical analysis, we consider a balanced panel dataset of 114 (both developed and developing) reporter countries[7] and 2,394 observations over the years 1992-2012, covering the overwhelming proportion of world trade.[8] Export data (in line with Cadot et al. 2011; Parteka & Tamberi 2013a) we use mirrored exports) are drawn from the United Nations Commodity Trade Statistics Database (UN Comtrade database, retrieved through the WITS database) at the highest level of disaggregation available for international comparisons (5,016 product lines) in the Harmonized System of goods classification.[9]

Trade data are used to compute measures of export diversification for each country and time period. Given that we are interested in links between countries and dependence between their trade structures, we choose to assess each country's export composition with respect to the overall trend (referring each country's degree of export diversity to other countries). Hence, we employ relative measures of diversification in the spirit of (De Benedictis et al. 2008; De Benedictis et al. 2009; Parteka & Tamberi 2013a; Mau 2016). The Relative Theil entropy index (*RelTheil*) is our benchmark measure and is computed for each time period as:

$$RelTheil_i = \sum_{j=1}^{m} \left(s_{ij} \cdot \ln \frac{s_{ij}}{w_j} \right), \quad RelTheil_i \in \langle 0, \ln(m) \rangle, \quad\quad [1]$$

where $s_{ij} = \dfrac{x_{ij}}{\sum_j x_{ij}}$ are the shares of the exports (x) of product j (j=1,2,….m) in the total

exports of country i (i=1,2,….,n) and $w_j = \dfrac{\sum_i x_{ij}}{\sum_i \sum_j x_{ij}}$ is the average share of product j in

total world exports. We have also considered Relative Gini index (*RelGini*) and the Dissimilarity Index (*DI*) as additional measures (see (Parteka 2010) for exact formulas). The higher the values of the indices, the less diversified (the more specialised) is the export structure of the country under investigation.

Additional variables used in the analysis include GDP *per capita* (*GDPpc*) at *PPP* constant 2005 international $, and population (*POP*), both obtained from the World Bank's World Development Indicators Database. *GDPpc* is used as a measure of economic development level, whereas *POP* is a proxy for country size (Mau 2016).[10] Additionally, in

---

[7] The countries are: Albania; Algeria; Angola; Armenia; Australia; Austria; Azerbaijan; Bangladesh; Belarus; Benin; Bolivia; Brazil; Bulgaria; Burkina Faso; Burundi; Cameroon; Canada; Central African Republic; Chad; Chile; China; Colombia; Congo, Rep.; Costa Rica; Cote d'Ivoire; Denmark; Dominican Republic; Ecuador; Egypt, Arab Rep.; El Salvador; Finland; France; Gabon; Gambia, The; Georgia; Germany; Ghana; Greece; Guatemala; Guinea; Guinea-Bissau; Honduras; Hong Kong SAR, China; Hungary; India; Indonesia; Israel; Italy; Japan; Jordan; Kazakhstan; Kenya; Korea, Rep.; Kyrgyz Republic; Lao PDR; Latvia; Lebanon; Lithuania; Madagascar; Malawi; Malaysia; Mali; Mauritania; Mauritius; Mexico; Moldova; Mongolia; Morocco; Nepal; Netherlands; New Zealand; Nicaragua; Niger; Nigeria; Norway; Pakistan; Panama; Papua New Guinea; Paraguay; Peru; Philippines; Poland; Portugal; Romania; Russian Federation; Rwanda; Saudi Arabia; Senegal; Sierra Leone; Singapore; Slovenia; South Africa; Spain; Sri Lanka; Sweden; Switzerland; Tajikistan; Tanzania; Thailand; Togo; Trinidad and Tobago; Tunisia; Turkey; Turkmenistan; Uganda; Ukraine; United Kingdom; United States; Uruguay; Uzbekistan; Venezuela, RB; Vietnam; Yemen, Rep.; Zambia.

[8] The estimation of dynamic spatial panel data models requires balanced data. The countries considered correspond to 90.7% of world trade (own calculations based on export data, 2012, from UN Comtrade). Microstates (defined as countries with a population below 1m) are excluded from the analysis.

[9] A similar level of detail is adopted by: (Klinger & Lederman 2006; Cadot et al. 2011; Parteka & Tamberi 2013a; Mau 2016).

[10] Given the high correlation (0.69) between the logs of our crucial dependent variable (*per capita* income) and GDP, in order to avoid multicollinearity issues we decide to use data on population instead (the correlation coefficient between the logs of *POP* and *GDPpc* is equal to only 0.05). In our sample, the correlation between

order to account for the degree of dependence on petrol, using UN Comtrade data[11] we compute the share of petrol-related products in overall country exports (*Oil*). Finally, to construct the weight matrices, we use great circle distances between the centroids of countries and bilateral trade data in 1992 (described in (Feenstra et al. 2005) and available at: http://cid.econ.ucdavis.edu/).

Table 1 shows summary statistics for each of the aforementioned variables. We observe a high variability (in terms of min/max differences) in the values of the three export diversification indices, indicating that countries with a very highly specialised export structure coexist in our panel with ones with a very diversified structure. Similarly, the *per capita* income of the countries ranges from only $420 to $53,578, with a mean of $10,125, suggesting great heterogeneity in the level of development among countries. In our sample there are also countries with a considerable share of petrol in their total exports (the maximum of the *Oil* variable equals 0.98); they are likely to have a different export structure to all the other countries so this variable is usually taken into account in panel data studies on export diversification as an additional covariate (Cadot et al. 2011; Parteka & Tamberi 2013a; Mau 2016).

Table 1. Description of variables and summary statistics

Variable	Description	Obs.	Mean	Std Dev	Min	Max
RelTheil	Relative Theil index	2,394	2.8409	1.4986	0.3739	8.0463
RelGini	Relative Gini index	2,394	0.8872	0.1268	0.4540	0.9999
DI	Dissimilarity index	2,394	1.5606	0.3202	0.6425	1.9790
GDPpc	GDP *per capita*, PPP const. 2005 US$	2,394	10125	11330	420	53878
POP	Population (millions)	2,394	49.6	158	0.980	1350
Oil	Share of petrol in total exports	2,394	0.13	0.24	0	0.98

Source: Authors' calculations.

To assess whether the data exhibit any specification problem, serial autocorrelation and cross-sectional dependence are tested. Breusch-Godfrey/Wooldridge and Pesaran tests reveal the presence of autocorrelation and cross-sectional dependence in the data (Table 2). For each export diversification index considered, we reject both the null hypothesis of no first-order autocorrelation (Wooldridge test) and the null hypothesis of cross-sectional independence (Pesaran 2004).

Table 2. Serial autocorrelation and spatial autocorrelation tests

|	*RelTheil*	*RelGini*	*DI*
Woolridge test (Chi sq)	1663.712*	2099.532*	2048.986*
Pesaran CD test for cross-sectional independence in panels	10.918*	3.7003*	-1.9262**

Note: *, ** and *** denote significance at the 1, 5 and 10 per cent levels respectively.
Source: authors' calculations.

---

the logs of *GDP* and *POP* is equal to 0.75, while that between the logs of land area and *POP* is equal 0.62, so population can be considered a good proxy of country size.

[11] Specifically, this variable is obtained with the use of product-level export statistics (HS 6-digit level) as a share of product lines 270900, 271000, 271011, 271119, 271129, 271210, 271311, 271312, 271320 and 271390 in overall country exports.

Figure 1. Spatial distribution of export diversification
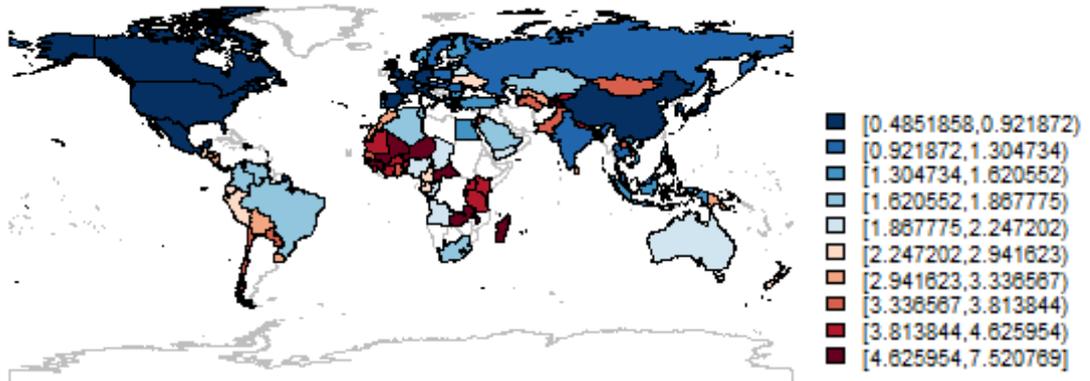
Note: map obtained using the *RelTheil* measure of export diversification, 2012
Source: authors' calculations

Finally, a choropleth map (Figure 1) provides evidence of the non-random spatial distribution of the degree of export diversification. Most of the countries lie close to other economies with a similar level of diversification, indicating strong spatial dependence[12].

3 Econometric analysis

3.1 A dynamic spatial model specification

In most developed and developing countries, the level of specialization is highly persistent (Section 2), and this feature has to be taken into account in the econometric model to avoid a misspecification bias. Hence, a dynamic approach is more appropriate for investigating the relationship between GDP *per capita* and trade specialization (Mau 2016). Additionally, in the presence of cross-country interdependence, standard panel estimators are likely to be biased and inconsistent (Elhorst 2014). In order to simultaneously take into account time persistence and spatial interdependence along with spatial and temporal heterogeneity, a dynamic spatial panel model with fixed spatial and time effects is needed.

The spatial econometric literature provides several alternative specifications of spatial dynamic models. A very general one includes time lags of both the dependent and independent variables, contemporaneous spatial lags of both, and lagged spatial lags of both. However, as (Elhorst 2014) points out, this generalized model suffers from identification problems and is not useful for empirical research. A more parsimonious model (written in vector form for a cross-section of observations at time t) can be expressed as

$$\mathbf{Y}_t = \tau \mathbf{Y}_{t-1} + \delta \mathbf{W} \mathbf{Y}_t + \eta \mathbf{W} \mathbf{Y}_{t-1} + \mathbf{X}_t \boldsymbol{\beta} + \boldsymbol{\mu} + \xi_t \boldsymbol{\iota}_N + \boldsymbol{\varepsilon}_t, \qquad [2]$$

where $\mathbf{Y}_t$ denotes a $N \times 1$ vector consisting of one observation of the dependent variable for every spatial unit ($i = 1,...,N$) in the sample (here, 114 countries) at time t ($t = 1,...,T$) and $\mathbf{X}_t$ is an $N \times K$ matrix of the explanatory variables. A vector or a matrix with subscript $t-1$ denotes its serially lagged value, while a vector or a matrix pre-multiplied by the spatial weights matrix $\mathbf{W}$ denotes its spatially lagged value.[13]

---

[12] Similar patterns can be observed with regard to the *RelGini* and *DI* indices (the maps are included in the full working paper version) → please see the additional material for the referee

[13] We also estimated a spatial Durbin version of the model. Nevertheless, the spatial lags of the explanatory variables did not enter the model significantly. Thus, we preferred to use the spatial lag representation given in equation [2].

The $K \times 1$ vector $\boldsymbol{\beta}$ includes the parameters of the explanatory variables, while τ, δ and η are the response parameters of the dependent variable lagged in time, $\mathbf{Y}_{t-1}$, the dependent variable lagged in space, $\mathbf{WY}_t$, and the dependent variable lagged in both space and time, $\mathbf{WY}_{t-1}$. The $N \times 1$ vector $\boldsymbol{\mu}$ contains spatial specific effects, μ_i, meant to control for all spatial-specific, time-invariant variables, the omission of which could bias the estimates in a typical cross-sectional study. Similarly, ξ_t denotes time-period specific effects, where ι_N is an $N \times 1$ vector of ones, controlling for all time-specific unit-invariant variables, the omission of which could also bias the estimates. These spatial and time-period specific effects are treated as fixed effects in our analysis, as it is very likely that unobserved effects are correlated with the regressor already included in the model. Finally, the elements of the disturbance term $\boldsymbol{\varepsilon}_t$ are assumed to be *i.i.d.* across i and t, i.e. $\boldsymbol{\varepsilon}_t \sim \left(\mathbf{0}, \sigma^2 \right)$.

The stationarity conditions on the spatial and temporal parameters in a dynamic spatial panel data model like [2] go beyond the standard condition $|\tau| < 1$ in serial models, and the standard condition $1/\omega_{\min} < \delta < 1/\omega_{\max}$ in spatial models (with $\omega_{\min}$ and $\omega_{\max}$ indicating the minimum and maximum eigenvalues of the $\mathbf{W}$ matrix). Indeed, to achieve stationarity in the dynamic spatial panel data model [2], the characteristic roots of the matrix $\left(\mathbf{I}_N - \delta \mathbf{W} \right)^{-1} \left(\tau \mathbf{I}_N + \eta \mathbf{W} \right)$ should lie within the unit circle (Elhorst 2001; Debarsy et al. 2012) which is the case when

$$\begin{aligned}
\tau + \left(\delta + \eta \right) \varpi_{\max} &< 1 \quad if \quad \delta + \eta \geq 0 \\
\tau + \left(\delta + \eta \right) \varpi_{\min} &< 1 \quad if \quad \delta + \eta < 0 \\
\tau - \left(\delta - \eta \right) \varpi_{\max} &> -1 \quad if \quad \delta - \eta \geq 0 \\
\tau - \left(\delta - \eta \right) \varpi_{\min} &> -1 \quad if \quad \delta - \eta < 0
\end{aligned} \qquad [3]$$

The parameters of model [2] are estimated using bias-corrected quasi-maximum likelihood (QML) estimators (Lee & Yu 2010).

3.2 Direct and indirect effects

Assuming that the matrix $\left(I_N - \delta \mathbf{W} \right)^{-1}$ is invertible, we can re-write model [2] as

$$\mathbf{Y}_t = \left(I_N - \delta \mathbf{W} \right)^{-1} \left(\tau I_N + \eta \mathbf{W} \right) \mathbf{Y}_{t-1} + \left(I_N - \delta \mathbf{W} \right)^{-1} \mathbf{X}_t \boldsymbol{\beta} + \left(I_N - \delta \mathbf{W} \right)^{-1} \boldsymbol{\varepsilon}_t. \qquad [4]$$

This equation allows us to compute the partial derivatives of the expected value of $\mathbf{Y}$ with respect to each k-th variable in $\mathbf{X}$ in each unit i at each time t in the short run:

$$\left[\frac{\partial E(\mathbf{Y})}{\partial x_{1k}} \ldots \frac{\partial E(\mathbf{Y})}{\partial x_{Nk}} \right]_t = \left(I_N - \hat{\delta} \mathbf{W} \right)^{-1} \hat{\beta}_k I_N \qquad [5]$$

and in the long run:

$$\left[\frac{\partial E(\mathbf{Y})}{\partial x_{1k}} \ldots \frac{\partial E(\mathbf{Y})}{\partial x_{Nk}} \right]_t = \left[\left(1 - \hat{\tau} \right) I_N - \left(\hat{\delta} + \hat{\eta} \right) \mathbf{W} \right]^{-1} \hat{\beta}_k I_N. \qquad [6]$$

The diagonal elements of both matrices [5] and [6] give a measure of the so-called 'direct effect', i.e. how much a change in the explanatory variable k for country i would affect the dependent variable for the same country i. This effect is different from the estimated parameter $\hat{\beta}_k$, since it includes the feedback effects that arise as a result of impacts passing through neighbouring countries and back to the countries themselves. The off-diagonal elements of the matrices [5] and [6] give a measure of the so-called 'indirect or spillover

effect', i.e. how much a change in the explanatory variable for country j might affect the dependent variable for any other country i.

Using [5] and [6], we can compute short-term and long-term average direct (ADE) and indirect (AIE) marginal effects:

$$\text{Short-term ADE for variable } k = \left[\left(I_N - \hat{\delta}\mathbf{W}\right)^{-1}\hat{\beta}_k I_N\right]^{\overline{d}} \qquad [7]$$

$$\text{Short-term AIE for variable } k = \left[\left(I_N - \hat{\delta}\mathbf{W}\right)^{-1}\hat{\beta}_k I_N\right]^{\overline{rsum}} \qquad [8]$$

$$\text{Long-term ADE for variable } k = \left[\left[\left(1-\hat{\tau}\right)I_N - \left(\hat{\delta}+\hat{\eta}\right)\mathbf{W}\right]^{-1}\hat{\beta}_k I_N\right]^{\overline{d}} \qquad [9]$$

$$\text{Long-term AIE for variable } k = \left[\left[\left(1-\hat{\tau}\right)I_N - \left(\hat{\delta}+\hat{\eta}\right)\mathbf{W}\right]^{-1}\hat{\beta}_k I_N\right]^{\overline{rsum}}, \qquad [10]$$

where the superscript $\overline{d}$ denotes the operator that calculates the mean diagonal element of the matrix and the superscript $\overline{rsum}$ denotes the operator that calculates the mean row sum of the non-diagonal elements.

3.3   Estimation results

Table 4 reports the quasi-maximum likelihood (QML) estimation results of the spatial dynamics equation [2] obtained using the balanced panel dataset described in Section 2 and three alternative weighting matrices W (described below). The benchmark dependent variable ($\mathbf{Y}_t$) is *RelTheil*, which captures the role of the extensive margin of diversification and reveals the best distributional properties for disaggregated export data while its qualitative interpretation is equal to the other measures (Mau 2016)[14].

The two main explanatory variables are included in the $\mathbf{X}_t$ matrix: (i) the log of GDP per capita $\left(\mathbf{lnGDPpc}_t\right)$, approximating the level of technological development of the country; and (ii) the log of the population size $\left(\mathbf{lnPop}_t\right)$, to capture the effect of exporter sizes, and thus to proxy for factor costs (assumed to be lower in large countries due to internal factor competition)[15]. As mentioned in the introduction, some authors found a hump shape relation between export diversification and growth, suggesting to include polynomial expansion terms of per capita income in the model. However, some preliminary checks support the argument of (Mau 2016) that the non-monotonic pattern is mainly due to an omitted log-transformation of the income variable. Thus, in our model specification we do not include a quadratic term for $\mathbf{lnGDPpc}_t$. We additionally control for the strong heterogeneity in the patterns of trade diversification between oil exporters

---

[14] We have checked the correlation between the *RelTheil* index and other measures of diversification used in the literature when applied to our data: the Gini index and the conventional Theil entropy measure (both in absolute terms, as in (Cadot et al. 2011). The correlation between them and our index – *RelTheil* – is very high (0.71 for the Gini index and 0.74 for the absolute Theil index).

[15] Additionally, we have considered a measure of 'remoteness' (distance of country i to export market j) among the explanatory variables (Baldwin & Harrigan 2011; Mau 2016): $remot_i = \left[\sum_j GDP_j \big/ dist_{ij}\right]^{-1}$ but this variable does not turn out to be statistically significant, so we decided to exclude it.

and other countries[16] and the model also includes country $(\mathbf{\mu})$ as well as time-specific effects (ξ_t).[17]

As *RelTheil* decreases with diversification and increases with specialization, we expect a negative impact of $\mathbf{lnGDPpc}_t$ and $\mathbf{lnPop}_t$. While the explanatory variables are in log values, the dependent variable is not logged because its distribution is not as extreme as GDP *per capita*. Thus, the marginal effects must be interpreted as semi-elasticities. The r.h.s. of the model also includes time and spatial lags of the dependent variable, that is, $\mathbf{Y}_{t-1}$, $\mathbf{WY}_t$ and $\mathbf{WY}_{t-1}$.

Separate columns of Table 2 refer to the results obtained with different weights matrices $(\mathbf{W})$ used for the estimation. The first one $(\mathbf{W}1)$ is an inverse-distance matrix, whose general term is defined as:

$$w1_{ij} = \begin{cases} 0 & \text{if } i = j \text{ and if } d_{ij} > \overline{d} \\ d_{ij}^{-1} / \sum_{j \neq i} d_{ij}^{-1} & \text{otherwise} \end{cases}, \qquad [11]$$

where d_{ij} is the great circle distance between the centroids of the countries and $\overline{d}$ is a cut-off value equal to 3,843 km, which corresponds to the minimum distance which allows all countries to have at least one neighbour. The second matrix $(\mathbf{W}2)$ is an exponential distance matrix, whose general term is defined as:

$$w2_{ij} = \begin{cases} 0 & \text{if } i = j \text{ and if } d_{ij} > \overline{d} \\ e^{-d_{ij}} / \sum_{j \neq i} e^{-d_{ij}} & \text{otherwise} \end{cases}. \qquad [12]$$

The third matrix, $\mathbf{W}3$, is based on bilateral flows. (Grossman & Helpman 1991; Coe et al. 1997), among others, suggest that international trade may be considered as a major diffusion vector of technological progress so that, in our framework, trade flows may proxy for multi-country technological interactions. The general term of $\mathbf{W}3$ is defined as:

$$w3_{ij} = \begin{cases} 0 & \text{if } i = j \\ m_{ij} / \sum_{j \neq i} m_{ij} & \text{otherwise} \end{cases}, \qquad [13]$$

where m_{ij} is the quantity of imports to country i coming from country j registered in 1992, that is at the beginning of the sample period to prevent endogeneity problems that might arise. All $\mathbf{W}$ matrices are row-standardized.

The stability conditions are always satisfied. Moreover, to assess whether the extension of the non-dynamic model to the dynamic spatial panel data model increases the explanatory power of the model, we estimate a static spatial autoregressive model (SAR) and test whether the coefficients of the variables $\mathbf{Y}_{t-1}$ and $\mathbf{WY}_{t-1}$ are jointly significant using an LR test. Using the matrix $\mathbf{W}1$, the outcome of this test (2x(-700.68-117.16)=1635.68 with 2 *df*) justifies the extension to the model with dynamic effects. Similar results are obtained using the alternative $\mathbf{W}$ matrices.

Table 4 shows that the two explanatory variables, $\mathbf{lnGDPpc}_t$ and $\mathbf{lnPop}_t$ are always statistically significant and the coefficients associated with these variables have the expected negative sign. Moreover, the results obtained with the alternative $\mathbf{W}$ matrices are

---

[16] In every specification we include the log of oil export share. Computing $\mathbf{ln(Oil}_t)$ we lose 151 observations for which $\mathbf{Oil}_t = 0$. Thus, we decide to take the log of $\mathbf{ln(Oil}_t + 0.01)$.

[17] Model [2] can also be extended to include the spatial lags of the $\mathbf{X}_t$ variable. We have estimated this Spatial Durbin specification, but the results systematically reveal that the $\mathbf{WX}_t$ terms do not enter significantly. Hence, we report the results of the restricted model [2].

very similar, confirming the robustness of the choice of spatial weights. However, the estimated coefficients cannot be interpreted as marginal effects, as they do not take into account spillover and feedback effects.

<div align="center">

Table 4. Dynamic spatial model – QML estimates.
Sample: all countries (114). Period: 1992-2012

</div>

	Dep. Variable: *RelTheil*		
	W1	**W2**	**W3**
$\mathbf{Y}_{t-1}$	0.577***	0.579***	0.579***
	(35.313)	(35.452)	(35.354)
$\mathbf{WY}_t$	0.206***	0.240***	0.525***
	(4.659)	(6.067)	(9.820)
$\mathbf{WY}_{t-1}$	-0.061	-0.117**	-0.280***
	(-1.012)	(-2.125)	(-2.010)
$\mathbf{lnGDPpc}_t$	-0.217***	-0.220***	-0.212***
	(-5.376)	(-5.459)	(-5.173)
$\mathbf{lnPop}_t$	-0.297***	-0.289***	-0.330***
	(-3.826)	(-3.739)	(-4.155)
$corr^2$	0.571	0.571	0.571
Log-lik.	-117.16	-113.62	-123.83

Notes: **W1** is an inverse distance matrix; **W2** is an exponential distance matrix; **W3** is a bilateral trade matrix; *t* statistics in parenthesis; *, ** and *** denote significance at the 1, 5 and 10 per cent levels respectively.
Source: authors' calculations.

The marginal effects estimates of the two main explanatory variables ($\mathbf{lnGDPpc}_t$ and $\mathbf{lnPop}_t$) are reported in Table 5. The short- and long-run direct effects are all significantly different from zero and have the expected negative sign. In line with the predictions of endogenous growth models (Grossman & Helpman 1991; Grossman & Helpman 1991; Acemoglu & Zilibotti 1997), the Ricardian-based Eaton and Kortum model (Mau 2016) and the related empirical literature on the determinants of export diversification (Parteka & Tamberi 2013b), a higher level of development and a bigger country size stimulate diversification.

How to interpret it? Making investments in new fields of activities is associated with uncertainty about future outcomes, and potentially also with sunk costs that cannot be recovered in the case of failure. Capital indivisibilities require a minimum stock of capital in order to make such investments possible. Consequently, richer (in terms of *per capita* income) and larger countries have more possibilities of starting risky projects. In other words, richer countries export more goods because their superior production technology endows them with an absolute advantage in global markets. Moreover, large countries can compensate for lower fundamental productivity with lower factor costs.

The semi-elasticity of the short-run direct effects is slightly higher than the estimated parameter, due to the feedback effects that arise as a result of the impacts passing through neighbouring countries and back to the countries themselves. Consistently with our expectations, the long-run direct effects are much stronger than the short-run direct effects. This is because it takes time before trade diversification levels change.

The indirect effects have the same sign as the direct effect. Spatial spillover effects are indeed negative and significant, both in the short and in the long run. Increases in GDP *per capita* and the population of a country have a positive impact not only on its own trade diversification, but also on the diversification of other countries, with a distance decay effect. Thus, knowledge spillovers reinforce the absolute technological advantage of countries and allow them to export more goods. Moreover, spatial proximity to large countries accelerates the diversification process, since the lower factor costs of the neighbours (which compensate for lower fundamental productivity) can easily be imported. Curiously, the spillover effect is even higher than the direct effect. This means that changes in the level of development of neighbouring countries are more important than changes in the characteristics of the country itself. In line with our expectations again, the long-run

indirect effects are much stronger than the short-run indirect effects. Finally, spatial spillover effects are much larger when using the bilateral trade matrix, **W3**, suggesting that the trade network is a more important driver of technological spillovers than simple geographical distance.

Table 5. Marginal effects – QML estimates.
Sample: all countries (114). Period: 1992-2012.

	Dep. Variable: *RelTheil*		
	W1	**W2**	**W3**
	Short-run effects		
ADE - **lnGDPpc**$_t$	-0.222***	-0.226***	-0.218***
	(-4.623)	(-4.661)	(-4.606)
AIE - **lnGDPpc**$_t$	-0.325***	-0.352***	-0.617***
	(-4.327)	(-4.348)	(-4.189)
ATE- **lnGDPpc**$_t$	-0.547***	-0.578***	-0.889***
	(-4.509)	(-4.529)	(-4.356)
ADE - **lnPop**$_t$	-0.306***	-0.297***	-0.343***
	(-3.546)	(-3.310)	(-3.937)
AIE - **lnPop**$_t$	-0.449***	-0.462***	-1.054***
	(-3.393)	(-3.196)	(-3.702)
ATE - **lnPop**$_t$	-0.756***	-0.759***	-1.397***
	(-3.485)	(-3.265)	(-3.807)
	Long-run effects		
ADE - **lnGDPpc**$_t$	-0.538***	-0.544***	-0.554***
	(-4.710)	(-4.766)	(-4.684)
AIE - **lnGDPpc**$_t$	-1.066***	-0.953***	-1.488***
	(-3.018)	(-3.208)	(-3.060)
ATE- **lnGDPpc**$_t$	-1.604***	-1.496***	-2.042***
	(-3.652)	(-3.900)	(-3.066)
ADE - **lnPop**$_t$	-0.743***	-0.715***	-0.882***
	(-3.570)	(-3.328)	(-3.525)
AIE - **lnPop**$_t$	-1.479***	-1.255***	-2.845***
	(-2.587)	(-2.553)	(-3.052)
ATE - **lnPop**$_t$	-2.222***	-1.970***	-2.727***
	(-2.977)	(-2.918)	(-3.056)

Notes: **W1** is an inverse distance matrix; **W2** is an exponential distance matrix; **W3** is a bilateral trade matrix; *t* statistics in parenthesis; *, ** and *** denote significance at the 1, 5 and 10 per cent levels respectively. ADE – direct marginal effect, AIE – indirect marginal effect, ATE – average total effect (ADE+AIE).
Source: authors' calculations

Finally, we take into account a considerable level of country heterogeneity in our sample in a more direct way. We estimate the spatial dynamic model excluding petrol-rich countries (those with an average oil export share greater than or equal to 50 percent). The **X**$_t$ matrix of explanatory variables now includes only **lnGDPpc**$_t$ and **lnPop**$_t$. The estimated marginal effect in this restricted model (Table 6) confirms the negative relationship between **lnGDPpc**$_t$ and *RelTheil*, while country size (**lnPop**$_t$) is no longer significant. This suggests that exclusion of petrol-rich countries also removes a substantial amount of variability in country size from the data.

Table 6. Marginal effects - QML estimates.
Sample: excluding 15 petrol-rich countries. Period: 1992-2012.

	Dep. Variable: *RelTheil*		
	W1	**W2**	**W3**
	Short-run effects		
ADE - **lnGDPpc**$_t$	-0.145***	-0.155***	-0.161***
	(-2.234)	(-2.635)	(-2.492)
AIE - **lnGDPpc**$_t$	-0.168***	-0.192***	-0.903**
	(-2.281)	(-2.560)	(-2.186)
ATE- **lnGDPpc**$_t$	-0.313***	-0.347***	-1.065***
	(-2.323)	(-2.609)	(-2.254)
ADE - **lnPop**$_t$	0.081	0.076	0.096
	(0.751)	(0.679)	(0.839)
AIE - **lnPop**$_t$	0.094	0.094	0.534
	(0.745)	(0.672)	(0.822)
ATE - **lnPop**$_t$	0.176	0.170	0.630
	(0.750)	(0.676)	(0.827)
	Long-run effects		
ADE - **lnGDPpc**$_t$	-0.426***	-0.452***	-0.441*
	(-2.347)	(-2.624)	(-1.788)
AIE - **lnGDPpc**$_t$	-0.614*	-0.581*	-0.106*
	(-1.736)	(-1.855)	(-1.755)
ATE- **lnGDPpc**$_t$	-1.040**	-1.034**	-0.547*
	(-2.065)	(-2.269)	(-1.728)
ADE - **lnPop**$_t$	0.238	0.221	0.270
	(0.749)	(0.678)	(0.785)
AIE - **lnPop**$_t$	0.322	0.271	0.613
	(0.667)	(0.611)	(0.053)
ATE - **lnPop**$_t$	0.561	0.493	0.901
	(0.713)	(0.651)	(0.074)

Notes: **W1** is an inverse distance matrix; **W2** is an exponential distance matrix; **W3** is a bilateral trade matrix; *t* statistics in parenthesis; *, ** and *** denote significance at the 1, 5 and 10 per cent levels respectively. ADE – direct marginal effect, AIE – indirect marginal effect, ATE – average total effect (ADE+AIE).
Source: authors' calculations

As a final remark, it is important to observe that QML estimators for dynamic spatial panel models (as used so far) are based on the assumption that there are only exogenous covariates except for the time and spatial lag terms. In order to check for the robustness of the results to endogeneity biases (due to for example simultaneity) [18], we have also used the System-GMM (Generalized Method of Moments) estimator (Blundell & Bond 1998) in place of QML to estimate model (2), as suggested by (Kukenova & Monteiro 2008). The results from two-step System-GMM robust estimations with (Windmeijer 2005)'s finite-sample correction[19], strongly confirms the main conclusions obtained using the QML estimator: a higher level of development and larger country size exert a positive effect on the export diversification of countries (bearing in mind that *RelTheil* is an inverse measure of export diversity), with indirect effects reinforcing the direct impact.

---

[18] Arguments that trade diversification generates economic growth are present in papers by (Al-Marhubi 2000; Feenstra & Kee 2008; Herzer & Nowak-Lehnmann D. 2006; Hesse 2008). Using System-GMM to estimate a non-spatial dynamic model and testing for reverse causality and potential feedback effects, (Mau 2016) shows that GPD *per capita* is weakly exogenous and that diversification also has an impact on GDP *per capita*.

[19] The results are presented in the full working paper version of the paper.) → please see the additional material for the referee

4 Conclusions

In this paper we have proposed an extension to the existing literature on the export diversification-development relationship. In particular, we have relaxed the implicit assumption of cross-country independence that has characterized all previous empirical works in this field. Our argument is that international trade in goods and cross-border mobility of factors of production make countries strongly interdependent. Consequently, a shock in the characteristics of one country (e.g. with respect to its income) is likely to have an impact not only on its own performance but also on the performance of all other countries, with a distance decay effect. Given the relationship between export diversity and GDP *per capita*, the transmission of shocks results in spatial dependence in terms of diversification too. We are not aware of any other trade diversification study which addresses this issue.

We have employed a spatial dynamics panel data specification, which has allowed us to capture short- and long-run, direct and indirect (spatial spillover) effects. The sample of countries analysed is very broad (114 economies at all stages of development, observed between 1992 and 2012, covering more than 90% of all trade exports).

Using QML and system GMM estimators, we have found that spatial network effects are indeed very important in determining the impact of GDP *per capita* and country size on the degree of export diversification. On the one hand, our results confirm the predictions of endogenous growth models: richer countries export more goods because their superior production technology endows them with an absolute advantage in global markets, while large countries exploit economies of scale and can compensate for lower fundamental productivity with lower factor costs. These are known as direct effects. On the other hand, our findings reveal that indirect effects strongly reinforce direct effects: spatial spillovers strengthen the absolute technological advantage of countries and allow them to export a greater variety of goods. Moreover, spatial proximity to large countries accelerates the diversification process, since lower factor costs of neighbours (which compensate for lower fundamental productivity) can easily be imported.

References

Acemoglu, D. & Zilibotti, F., 1997. Was Prometheus Unbound by Chance? Risk, Diversification, and Growth. *Journal of Political Economy*, 105(4), pp.709–751.

Aghion, P. & Howitt, P., 1997. Endogenous Growth Theory. *MIT Press Books*, 1, pp.1–12.

Agosin, M.R., Alvarez, R. & Bravo-Ortega, C., 2012. Determinants of export diversification around the world: 1962-2000. *The World Economy*, 35(3), pp.295–315.

Allen, F. & Gale, D., 2000. Financial Contagion. *Journal of Political Economy*, 108(1), pp.1–33.

Al-Marhubi, F., 2000. Export Diversification and Growth: an Empirical Investigation. *Applied Economics Letters*, 7(9), pp.559–562.

Amurgo-Pacheco, A. & Pierola, M.D., 2008. Patterns of exporting diversification in developing countries: Intensive and extensive margins. *World Bank Policy Research Working Paper*, No. 4473.

Baldwin, R. & Harrigan, J., 2011. Zeros, quality, and space: Trade theory and trade evidence. *American Economic Journal: Microeconomics*, 3(2), pp.60–88.

Barro, R.J. & Sala-i-Martin, X., 2004. *Economic growth* 2nd ed., The MIT Press.

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

De Benedictis, L., Gallegati, M. & Tamberi, M., 2008. Semiparametric analysis of the specialization-income relationship. *Applied Economics Letters*, 15, pp.301–306.

de Benedictis, L. & Tajoli, L., 2011. The World Trade Network. *World Economy*, 34(8), pp.1417–1454.

Blundell, R. & Bond, S., 1998. Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1), pp.115–143.

Cadot, O., Carrere, C. & Strauss-Kahn, V., 2013. Trade diversification, income, and growth: What do we know? *Journal of Economic Surveys*, 27(4), pp.790–812.

Cadot, O., Carrère, C. & Strauss-Kahn, V., 2011. Export Diversification: What's behind the Hump? *Review of Economics and Statistics*, 93(January), pp.590–605.

Carrere, C. & Strauss-Kahn, V., 2014. Developing Countries Exports Survival in the OECD: Does Experience Matter. *FERDI Working Paper*, P100.

Chaney, T., 2014. The network structure of international trade. *American Economic Review*, 104(11), pp.3600–3634.

Coe, D.T., Helpman, E. & Hoffmaister, A.W., 1997. North-South R&D Spillovers. *Economic Journal*, 107(440), pp.134–49.

Comin, D., Dmitriev, M. & Rossi-Hansberg, E., 2012. The Spatial Diffusion of Technology. *NBER Working Paper*.

Deardorff, A. V, 2000. Patterns of Trade and Growth across Cones. *De Economist*, 148(2).

Debarsy, N., Ertur, C. & LeSage, J.P., 2012. Interpreting dynamic space-time panel data models. *Statistical Methodology*, 9(1-2), pp.158–171.

Dennis, A. & Shepherd, B., 2011. Trade facilitation and export diversification. *World Economy*, 34(1), pp.101–122.

Duranton, G. & Puga, D., 2004. Micro-foundations of urban agglomeration economies. In J. V Henderson & J. F. Thisse, eds. *Handbook of Regional and Urban Economics*. Elsevier, pp. 2063–2117.

Eaton, J. & Kortum, S., 2002. Technology, Geography, and Trade. *Econometrica*, 70(5), pp.1741–1779.

Elhorst, J.P., 2001. Dynamic Models in Space and Time. *Geographical Analysis*, 33(2), pp.119–140.

Elhorst, J.P., 2014. Spatial panel data models. *Spatial Econometrics*, pp.37–93.

Ertur, C. & Koch, W., 2011. A contribution to the theory and empirics of Schumpeterian growth with worldwide interactions. *Journal of Economic Growth*, 16(3), pp.215–255.

Ertur, C. & Koch, W., 2007. Growth, technological interdependence and spatial externalities: Theory and evidence. *Journal of Applied Econometrics*, 22(6), pp.1033–1062.

Feenstra, R. & Kee, H.L., 2008. Export variety and country productivity: Estimating the monopolistic competition model with endogenous productivity. *Journal of International Economics*, 74(2), pp.500–518.

Feenstra, R.C. et al., 2005. World Trade Flows: 1962-2000. *National Bureau of Economic Research Working Paper Series*, No. 11040.

Fujita, M., Krugman, P. & Venables, A.J., 1999. *The Spatial Economy: Cities, Regions, and International Trade*, The MIT Press.

Fujita, M. & Thisse, J.F., 2002. *Economics of Agglomeration: Cities, Industrial Location, and Regional Growth*, Cambridge University Press.

di Giovanni, J. & Levchenko, A. a., 2011. *The Risk Content of Exports: A Portfolio View of International Trade*,

Glaeser, E.L., 2008. *Cities, Agglomeration, and Spatial Equilibrium*, Oxford University Press.

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

Grossman, G.M. & Helpman, E., 1991. Quality Theory Ladders of in the Growth. *The Review of Economic Studies*, 58(1), pp.43–61.

Hausmann, R. & Hidalgo, C.A., 2011. The network structure of economic output. *Journal of Economic Growth*, 16(4), pp.309–342.

Herzer, D. & Nowak-Lehnmann D., F., 2006. What does export diversification do for growth? An econometric analysis. *Applied Economics*, 38(15), pp.1825–1838.

Hesse, H., 2008. Export diversification and economic growth. In R. Newfamer, W. Shaw, & P. Walkenhorst, eds. *Breaking into New Markets Emergin Lessons for Export Diversification.* Washington, DC: The World Bank., pp. 55–80.

Howitt, P., 2000. Endogenous growth and cross-country income differences. *American Economic Review*, 90(4), pp.829–846.

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

Klinger, B. & Lederman, D., 2006. Diversification, innovation, and imitation inside the global technological frontier. *World Bank Policy Research Working Paper*, 3872.

Koren, M. & Tenreyro, S., 2013. Technological diversification. *American Economic Review*, 103(1), pp.378–414.

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

Krugman, P., 1995. Increasing returns, imperfect competition and the positive theory of international trade. In *Handbook of International Economics.* pp. 1243–1277.

Krugman, P.R., 1991. *Geography and trade*, The MIT Press.

Kukenova, M. & Monteiro, J.-A., 2008. Spatial dynamic panel model and system GMM: a Monte Carlo investigation. *October*, (13405), pp.1–23.

Lee, L. & Yu, J., 2010. A spatial dynamic panel data model with both time and individual fixed effects. *Econometric Theory*, 26(02), pp.564–597.

Mau, K., 2016. Export Diversification and Income Differences Reconsidered : The Extensive Product Margin in Theory and Application. *Review of World Economics*, 152(2), pp.351–381.

Minondo, A., 2011. Does comparative advantage explain countries' diversification level? *Review of World Economics*, 147(3), pp.507–526.

Neary, J.P., 2009. Putting the "new" into new trade theory: Paul Krugman's nobel memorial prize in economics. *Scandinavian Journal of Economics*, 111(2), pp.217–250.

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

Parteka, A. & Tamberi, M., 2013a. Product diversification, relative specialisation and economic development: Import-export analysis. *Journal of Macroeconomics*, 38(PA), pp.121–135.

Parteka, A. & Tamberi, M., 2013b. What determines export diversification in the development process? Empirical assessment. *The World Economy*, 36(6), pp.807–826.

Pesaran, M.H., 2004. General Diagnostic Tests for Cross Section Dependence in Panels. *Iza Discussion Papers*, (1240), pp.1–42.

Schott, P.K., 2003. One Size Fits All ? Hecksher-Ohlin Specialization in Global Production One Size Fits All ? Hecksher-Ohlin Specialization in Global Production. *The American Economic Review*, 93(3), pp.686–708.

Windmeijer, F., 2005. A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of Econometrics*, 126(1), pp.25–51.

I. Choropleth maps with alternative measures of export diversification
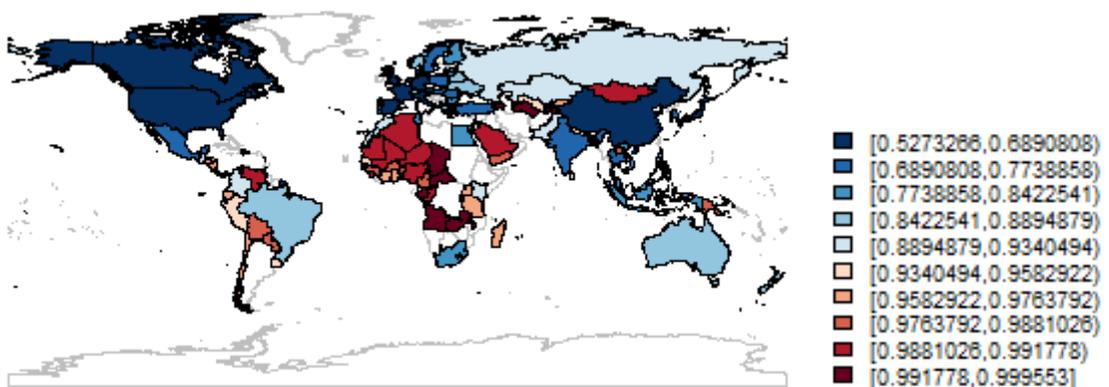
Figure A1. Spatial distribution of export diversification (2)
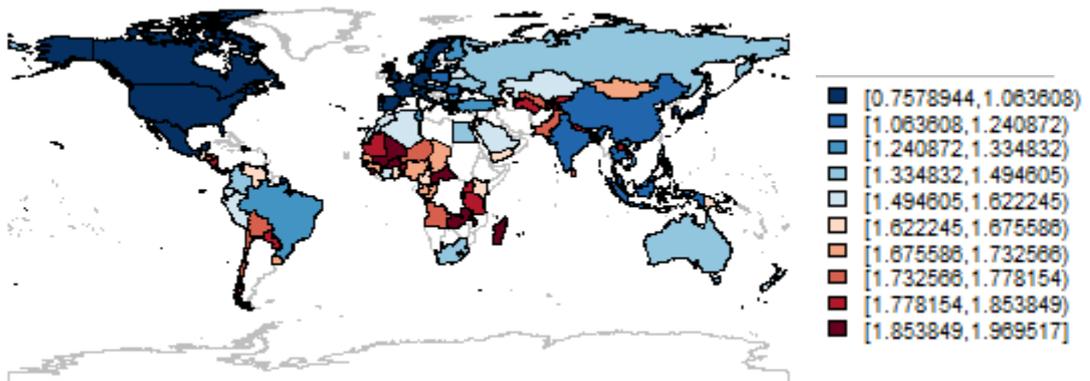
Note: map obtained with *DI* measure of export diversification, 2012

Source: authors' calculations

Figure A2. Spatial distribution of export diversification (3)

Note: map obtained with *RelGini* measure of export diversification, 2012

Source: authors' calculations

II. Alternative method of estimation: System-GMM estimation results

In order to account for potential endogeneity problems, we additionally performed two-step System-GMM robust estimations with Windmeijer's (2005) finite-sample correction: the results are shown in Table A1. The corresponding marginal effects are displayed in Table A2.[20]

Table A1. Dynamic spatial model – two-step Robust System GMM estimates. Sample: all countries (114). Period: 1992-2012

	Dep. Variable: *RelTheil*		
	W1	**W2**	**W3**
$\mathbf{Y}_{t-1}$	0.668***	0.702***	0.660***
	(7.930)	(8.460)	(7.970)
$\mathbf{WY}_t$	0.444***	0.638***	0.860***
	(2.830)	(3.660)	(2.610)
$\mathbf{WY}_{t-1}$	-0.407***	-0.608**	-0.644***
	(-2.970)	(-3.900)	(-2.500)
$\mathbf{lnGDPpc}_t$	-0.287***	-0.245***	-0.341***
	(-3.620)	(-3.290)	(-3.780)
$\mathbf{lnPop}_t$	-0.213***	-0.192***	-0.166***
	(-2.960)	(-2.830)	(-3.510)
$\mathbf{ln(Oil}_t + 0.01)$	-0.160	-0.121	-0.150
	(-1.660)	(-1.240)	(-1.280)
AR(1)	[0.003]	[0.002]	[0.002]
AR(2)	[0.374]	[0.422]	[0.390]
Hansen *J*	[0.197]	[0.407]	[0.113]
C-stat. Level-eq.	[0.315]	[0.393]	[0.607]

Notes: **W1** is an inverse distance matrix; **W2** is an exponential distance matrix; **W3** is a bilateral trade matrix; *t* statistics in parenthesis; *, ** and *** denote significance at the 1, 5 and 10 per cent levels respectively. Source: authors' calculations. *p*-values in brackets.

---

[20] The number instruments is quadratic in T. A large number of instruments can over-fit the instrumented variables and leads to inaccurate estimation of the optimal weight matrix, downward-biased two-step standard errors and wrong inference in the Hansen test (Roodman, 2009). To avoid these problems, we use a restricted set of instruments for GMM estimates. The number of instruments is set to two for estimations in differenced equations: we use two lagged levels in time periods t-2 and t-3 as instruments, while we use one-period lagged first differences for GMM in levels equations. We also include the spatial lags of $\mathbf{ln(Oil}_t + 0.01)$ and of $\mathbf{lnPop}_t$ as additional instruments.

Table A2. Marginal effects – system GMM estimates

	Dep. Variable: *RelTheil*		
	W1	**W2**	**W3**
	Short-run effects		
ADE - **lnGDPpc**$_t$	-0.308***	-0.288***	-0.349***
	(-4.639)	(-3.297)	(-4.555)
AIE - **lnGDPpc**$_t$	-0.730***	-1.069***	-0.213***
	(-4.037)	(-4.330)	(-3.003)
ATE- **lnGDPpc**$_t$	-1.038***	-1.357***	-0.562***
	(-4.231)	(-4.247)	(-4.008)
ADE - **lnPop**$_t$	-0.228***	-0.226***	-0.169***
	(-4.114)	(-3.912)	(-3.562)
AIE - **lnPop**$_t$	-0.540***	-0.840***	-0.056***
	(-4.889)	(-4.192)	(-3.002)
ATE - **lnPop**$_t$	-0.768***	-1.069***	-0.225***
	(-4.234)	(-4.045)	(-3.007)
	Long-run effects		
ADE - **lnGDPpc**$_t$	-0.879***	-0.834***	-0.948***
	(-4.125)	(-4.782)	(-4.301)
AIE - **lnGDPpc**$_t$	-1.126***	-1.065***	-1.063***
	(-3.796)	(-2.320)	(-3.018)
ATE- **lnGDPpc**$_t$	-2.005***	-1.899***	-2.011***
	(-4.574)	(-3.997)	(-3.035)
ADE - **lnPop**$_t$	-0.651***	-0.656***	-0.461***
	(-3.303)	(-3.370)	(-3.312)
AIE - **lnPop**$_t$	-0.834***	-0.837***	-0.506***
	(-2.742)	(-2.310)	(-3.018)
ATE - **lnPop**$_t$	-1.485***	-1.493***	-0.969***
	(-3.318)	(-3.915)	(-3.034)

Notes: **W1** is an inverse distance matrix; **W2** is an exponential distance matrix; **W3** is a bilateral trade matrix; *t* statistics in parenthesis; *, ** and *** denote significance at the 1, 5 and 10 per cent levels respectively. ADE – direct marginal effect, AIE – indirect marginal effect, ATE – average total effect (ADE+AIE).
Source: authors' calculations.