EXPORT POTENTIAL AND DIVERSIFICATION ASSESSMENTS[1]

A methodology to identify export opportunities

Abstract

To assist countries with identifying promising products for export promotion activities, the International Trade Centre (ITC) has developed an export potential assessment methodology. It is based on the decomposition of a country's potential exports into three factors: supply, demand and ease of trade. Depending on a country's particular needs, two approaches are available: (i) The export potential indicator (EPI) helps countries support established export sectors by increasing exports to existing and new markets. Inspired by a gravity-type framework, the EPI identifies products which a country already exports competitively, and which have good prospects of export success in a given target market. (ii) The product diversification indicator (PDI) helps countries diversify and develop new export sectors. Based on Hausmann and Hidalgo's notion of the product space, the PDI identifies products which the exporting country does not yet competitively export but which seem feasible given the country's current export basket and the export baskets of similar countries.

[1] Results of this methodology are disseminated through an online tool, the ITC Export Potential Map, available at: exportpotential.intracen.org. The methodology and tool were initially developed by Yvan Decreux and Julia Spies from the Trade and Market Intelligence (TMI) section of the International Trade Centre (ITC). Subsequent contributions were made by Maria Cantero, Cecilia Heuser, Dumebi Ochem, Sylvain Périllat, Lucas Ouriques Poffo, Julia Seiermann, and Cailey Stevens, among others. The authors wish to thank Lionel Fontagné for reviewing the paper and providing helpful feedback, as well as the users of Export Potential Map, who regularly provide valuable comments, and seminar participants at the ITC and WTO, where the methodology was presented and discussed.

Introduction

The International Trade Centre (ITC) has developed an export potential assessment methodology to support and bring transparency to the selection of products and markets for export promotion activities. Based on detailed trade and market access information, it identifies existing products with high export potential and/or diversification opportunities in a given target market:

- The *export potential indicator (EPI)* serves countries that aim to support established export sectors in increasing their exports to new or existing target markets. It identifies products in which the exporting country has already proven to be internationally competitive, and which have good prospects of export success in specific target markets (intensive product margin).

- The *product diversification indicator (PDI)* serves countries that aim to diversify and develop new export sectors facing promising demand conditions in new or existing target markets. It identifies products which the exporting country does not yet export competitively but which seem feasible based on the country's current export basket and the export baskets of similar countries (extensive product margin).

Conceptually, export potential assessments are based on the assumption that in a world without frictions trade flows can be described as a combination of supply performance, ease of trade and total demand. While a country's capacity to supply existing products (EPI) is captured through projected market shares, its capacity to diversify into new products (PDI) relies on Hausmann and Hidalgo's concept of the product space (Hausmann and Klinger, 2007, Hausmann et al., 2007 and Hidalgo et al., 2007), which establishes links between products through an assessment of how frequently they are found together in the export baskets of countries.

The first approach—the EPI—is based on a structural model that (i) identifies potential export values from supply capacities in the exporting country, demand conditions in the target market and bilateral linkages between the two. This corresponds to an empirical specification with country × product, importer × product and country × importer fixed effects but it avoids computational constraints when working at a detailed product level.[2] Any gap between what countries could export and what they actually do export is then argued to result from factors such as lack of information about rules and regulations in the target market, difficulties complying with regulations or the inability to meet the (quality or price) preferences of consumers. These are considered to be factors that trade advisors can address with local companies.

The second approach—the PDI—is based on the concept of the product space. The export potential assessment methodology develops the purely outcome-based measure of linkages to new products further by accounting for natural endowments that are pivotal for the capacity of a country to produce certain products. It also responds to criticisms of the product space of being entirely supply-side driven (see e.g. Harrison and Rodriguez-Clare, 2011 or Lederman and Maloney, 2012) by combining it with demand and market access information. Since the descriptive nature of the product space does not allow for any meaningful estimate of potential trade values, we present rankings of diversification opportunities in a given country or regional market.

This paper first presents an overview of the trade literature related to export potential. It then describes the methodology used to measure the export and diversification potential of products in given target markets. Next, it details how data is processed and treated to achieve robust and reliable results for all countries. The final section concludes and discusses the limitations of the model.

[2] See Head and Mayer (2014).

Related literature

Export potential assessments are linked to different strands of literature. The EPI is inspired by the gravity model of trade, which has been traditionally used to evaluate trade potential with existing trade partners. Nevertheless, the EPI is also related to the international market selection (IMS) literature, which seeks to provide a practical approach to help firms choose new foreign markets to export. The PDI, meanwhile, explores new products for export diversification and is closely related to the product space methodology.

The gravity model and trade potential

In the traditional gravity framework, the assumption is that trade between two countries is positively associated with the level of supply and demand and negatively associated with trade frictions, as shown in the equation below:

$$X_{ij} = K \times \frac{Y_i^{\alpha} \times Y_j^{\beta}}{T_{ij}^{\theta}}$$

where X_{ij} denotes exports from i to j, K represents a constant, Y_i is the exporter's GDP (supply), Y_j is the importer's GDP (demand), and T_{ij} corresponds to trade impediments between both countries, such as customs duties and geographic distance. The model's fitted values represent the average feasible trade between country pairs—or potential trade—when trade frictions are absent. Potential trade growth is then estimated either as the difference or the ratio between the expected mean trade and actual trade between country pairs (Dadakas et al., 2020; Armstrong, 2007).

Early attempts to use the gravity model for predicting potential trade were focused on East-West integration (Baldwin, 1994; Gros and Gonciarz, 1996; Nilsson, 2000) and have been subject to fundamental critique. For instance, Egger (2002) argues that large deviations of actual from predicted trade indicate a misspecification of the model rather than unrealized opportunities, and advocates for out-of-sample estimations where the model is fitted on a reference group of countries. A few authors have used (out-of-sample) gravity models to determine trade potential at the sector level. While some studies (e.g., Melchior et al., 2009 or Shepotylo, 2009) apply coefficients obtained from country characteristics to infer the potential trade of sectors, others (e.g., Fontagné et al., 2002; Helmers and Pasteels, 2006) suggest that without disaggregated information on the key determinants of sector level trade, results will be inaccurate. Product-level production and consumption data, however, are not available, and inferring these from exports and imports would introduce an endogeneity problem that eventually leads to biased coefficients.

More recently, theory-consistent in-sample estimation methods have gained prominence in the gravity literature (for a summary, see Head and Mayer, 2014) that include for example measures of trade frictions (e.g., multilateral resistance terms) or employ estimators that consider the information present in zero trade flows (Mulabdic and Yasar, 2021; Dadakas et al., 2020). However, in-sample estimations remain problematic as a consistent model should not have systematic differences between predicted and observed trade flows (Egger, 2002). Additionally, some authors argue that the standard gravity approach is inadequate for this type of estimation, as the model's fitted values represent the average instead of the maximum possible trade value (Drysdale et al., 2000; Armstrong, 2007).

Maximum possible trade: the stochastic frontier gravity model

To estimate the maximum possible trade a country-pair can achieve—as opposed to the average estimated by traditional gravity models—some authors have borrowed the "frontier" concept from the production literature and applied it to trade analysis. Specifically, stochastic frontier analysis has been used in combination with gravity models to estimate a "stochastic frontier", under the assumption of no frictions in trade flows. The main difference between the standard gravity model and the stochastic frontier gravity model (SFGM) is that the latter has two error components: one non-negative with a half-normal or exponential distribution to account for trade inefficiencies, and another one normally distributed with mean zero accounting for other random disturbances, such as measurement errors (Drysdale et al., 2000). In addition, the frontier specification generates efficiency scores indicating how far or close bilateral trade is from the trade frontier. A high level of trade efficiency means that two countries operate close to the frontier, while deviations imply that there is potential for trade expansion (Ravishankar and Stack, 2014).

Drysdale et al. (2000) were among the first to apply the SFGM methodology by evaluating China's bilateral trade efficiency with 57 countries. Several other studies focus on regional integration and the impact of trade agreements. For instance, Kumar and Prabhakar (2017) evaluate the impact of India's free trade agreements on its trade efficiency, and Ebaidalla and Ali (2022) assess intra-Arab trade integration and potential. Stack et al. (2018) find that western European countries' trade with new EU members is close to two-thirds of its maximum potential, suggesting high integration. Their results also indicate an increase in trade integration—or a decrease in unrealized potential—in larger EU countries following Brexit. In addition to total trade, studies have focused on outcomes at the sector level (Atif et al., 2017; Barma, 2017; Atif et al., 2019; Obeng, 2022) and trade in services (Nasir and Kalirajan, 2014).

Nevertheless, both traditional and stochastic frontier gravity models impose assumptions about the distribution of the error component that might not always be met, resulting in biased coefficients (Armstrong, 2007; Sotiros et al., 2022). Furthermore, without disaggregated information on the critical determinants of sector-level trade, results for commodity-level specifications will be inaccurate (Helmers and Pasteels, 2006). Nonbiased cross-country and cross-sector specifications would also require estimating many coefficients and become computationally burdensome. Models with country × product, product × market, and country × market fixed effects further bear the disadvantage that the source of the potential export value remains opaque, which makes it challenging to interpret the results and understand the drivers of export potential. Therefore, a different, more practical approach such as the one used in export potential assessments is better suited to guide policymakers and trade advisors.

Identifying new markets: international market selection

While the gravity-framework stems from traditional empirical work in international trade, the international market selection (IMS) literature proposes a different set of empirical strategies to market selection, involving not only economics but also building on the work of engineers and administrators (Deaza et al., 2020). IMS is the process firms adopt to select which new foreign markets to focus on (Papadopoulos et al., 2002; Papadopoulos and Martin, 2011). It involves investigating market potential and classifying it according to pre-established criteria (Andersen and Strandskov, 1998). Research has shown that micro, small, and medium-sized enterprises (MSMEs) implementing a systematic approach (i.e., a methodological approach) to foreign market selection have experienced higher export performance (Brouthers and Nakos, 2005). Nevertheless, non-systematic IMS processes (i.e., relying on experience and intuition) are prevalent, particularly among MSMEs, as they generally lack the expertise and resources to carry out a systematic IMS (Musso and Francioni, 2014).

Several quantitative and qualitative IMS methodologies aiming to identify new potential international trading partners were developed in the last few decades (for a recent review on IMS models, see Deaza et al., 2020). Some trending quantitative empirical strategies include data envelopment analysis (DEA) (Saen, 2011; Shabani and Saen, 2013; Cano et al., 2017a) and machine learning methodologies, such as fuzzy logic and neural networks (Brouthers et al., 2009; Cano et al., 2017b).[3] However, the IMS literature is characterized by fragmented research streams and most of the available models and methodologies are applied on a case-by-case basis, failing to prove their general applicability (Papadopoulos and Martin, 2011). Furthermore, research based on firm-level decisions in developing countries and focused on sectors other than manufacturing and service is scant (Deaza et al., 2020). Besides, IMS methodologies are often too complex to be employed by firms and policymakers as they require prior knowledge of statistical methods (Shabani and Saen, 2013). Finally, a flexible and cost-efficient model which would accommodate the multiplicity of industries is still lacking (Sheng and Mullen, 2011).

To the best of our knowledge, the only other comprehensive approach that combines supply, demand, and market access conditions to provide advice on export opportunities at a detailed product level is the TRADE-Decision Support Model (TRADE-DSM) (Cuyvers et al., 2012; Cuyvers et al., 2017). The TRADE-DSM relies on a filtering process that eliminates less realistic export opportunities in a step-by-step process. Trade potential values are then inferred from competitors' market value in remaining product × market combinations. However, contrary to export potential assessments, potential values are not derived from a structural model and do not consider forward-looking market developments, such as import demand growth and scheduled tariff reductions. Furthermore, the source of the export potential is indeterminate, and products in which a country has a low revealed comparative advantage index (RCA) are filtered out, limiting the market analysis to products that are currently an important part of a country's export basket.

Identifying new products: the product space

The PDI makes use of the concept of product space as developed by Hausmann and Klinger (2007), Hausmann et al. (2007), and Hidalgo et al. (2007). According to the authors, export diversification is path-dependent and contingent on the products a country currently exports. They establish a network linking goods based on their likelihood of being exported together, known as the product space, which captures information on the capabilities (i.e., capital, labour, knowledge, institutions) required to produce new goods. More specifically, country A should find it easier to export a new product if that product occurs in the export basket of country B along with products that the country A already exports. The underlying assumption is that the new product requires a similar set of capabilities to those that the country already possesses. Recent studies support this hypothesis: Pinheiro et al. (2018) finds that countries diversify their exports to related products approximately 93% of the time. For a comprehensive review of the literature on relatedness, see Hidalgo et al. (2018).

Several studies have applied this framework to guide developing countries on export diversification.[4] These papers typically focus on indicators established in Hausmann et al. (2014): the product's closeness to a country's current capabilities (distance), the product's level of sophistication (product complexity index), and the product's ability to contribute to economic complexity and development (opportunity gain).[5] Since trade-offs exist between these indicators (e.g., more complex products tend to be further from a country's current capabilities), different indices have been developed based on different policy objectives to help with product selection. For instance, Hausmann et al. (2022) calculate two summary indicators to rank products for Namibia, one giving higher weights to shorter distances (parsimonious transformation) and another one highlighting growth potential by prioritizing

[3] Gupta and Kumar (2021) provide a brief overview of studies using machine learning to predict international trade patterns. While it is not directly linked to IMS and trade potential, this still nascent literature is a promising tool to guide market selection decisions.

[4] Some studies (e.g., Hausmann et al., 2022) also provide suggestions on export promotion at the intensive margin by selecting the goods on which the country has a revealed comparative advantage and has a high complexity index (PCI).

[5] Data on these indices are available at the country level from the Observatory of Economic Complexity website: https://oec.world/ .

higher opportunity gain (strategic bet). Other studies (Hausmann and Chauvin, 2015; Obeng, 2022) use a multistep filtering process based on the feasibility and attractiveness of selected products.

While a few authors (Hausmann and Chauvin, 2015; Obeng, 2022) consider which goods have been intensively imported by neighbouring partners when selecting potential new exports for regional markets, most studies have focused on supply capacities without assessing whether the suggested products have any prospect of export success in the potential target markets. The PDI responds to criticisms of the product space being entirely supply-side driven (Harrison and Rodriguez-Clare, 2011; Lederman and Maloney, 2012) by combining it with demand and market access information.

Methodology

Export potential assessments consist of two indicators: the EPI indicates potential export values of a country's existing export products in any given market (whether new or existing), while the PDI ranks product diversification opportunities in any given market. The two indicators differ in the way supply capacity is captured. Demand and market access conditions are the same.

Export Potential Indicator

The methodology to estimate export potential is inspired by a gravity model specified at the product level. The starting point is the assumption that in a world without frictions, trade flows can be described by a combination of country × product, importer × product and country × importer factors,

$$v_{ijk} = \alpha_{ik}\beta_{ij}\gamma_{jk} \tag{1}$$

where v_{ijk} corresponds to exports from country i of product k to market j. The parameter α_{ik} describes country i's export performance in product k, γ_{jk} reflects market j's demand for product k and β_{ij} the overall ease of trade from i to j.

Equation (1) could be estimated econometrically and the differences between fitted and actual values may be interpreted as unrealized potential. The approach is inconvenient in this context for several reasons: First, the size of a model specified at the detailed product and country level would be very large. Second, the quality of three-dimensional data (v_{ijk}) is often lower than that of two-dimensional data. Third, in the absence of detailed production and consumption statistics, fixed effects would be used to capture ik-, ij- and jk-specific characteristics. A model estimated with fixed effects however makes it difficult to understand the drivers of the resulting export potential values – a key requirement for guiding policy and achieving impact.

Export potential assessments therefore take a different approach and infer potential export values at ijk level from a multiplicative model based on two-dimensional data,

$$\widetilde{v_{ijk}} = \widetilde{\alpha_{ik}}\widetilde{\beta_{ij}}v_{jk} = \frac{v_{ik}}{v_k}\frac{v_{ij}}{\sum_k\left(\frac{v_{ik}}{v_k}v_{jk}\right)}v_{jk} \tag{2}$$

where the first term corresponds to country i's world market share in product k. The second term is a measure of bilateral trade relative to what trade would be if the country held the same share in market j as it holds in the world market for all products. The third term simply reflects total imports. Annex I formally shows that in a world without frictions, potential export values correspond to actual export values under certain assumptions, such as:

- Perfect competition within each country (supply is perfectly elastic but does not feature economies of scale);

- Preferences specifying how imports of a given product are allocated among supplying countries are the same in all markets and take the form of a constant elasticity of substitution (CES) function;
- Bilateral trade costs are ad valorem and do not depend on products;[6]
- The average import price changes from market to market in the same proportion for all products;
- The structure of market demand by product as perceived by any given country does not differ from the structure of world demand.[7]

In reality, hypothetical (or potential) trade flows often deviate from actual trade for a number of reasons:

First, the assumptions listed above, and formally described in Annex I, do not always hold. The sum of estimated shares in a given market may therefore differ from 1, i.e. $\sum_i(\widetilde{\alpha_{ik}}\widetilde{\beta_{ij}}) \neq 1$. To solve this problem, a normalization factor is introduced.

$$\widetilde{v_{ijk}} = \frac{\widetilde{\alpha_{ik}}\widetilde{\beta_{ij}}}{\sum_i(\widetilde{\alpha_{ik}}\widetilde{\beta_{ij}})} v_{jk} = \frac{v_{ik}}{v_k} \frac{v_{ij}}{\sum_k\left(\frac{v_{ik}}{v_k} v_{jk}\right)} \frac{1}{\sum_i\left[\frac{v_{ik}}{v_k} \frac{v_{ij}}{\sum_k\left(\frac{v_{ik}}{v_k} v_{jk}\right)}\right]} v_{jk} \tag{3}$$

Second, equation (3) does not account for all measurable factors that impact supply capacity (global tariff advantages or re-exports may overestimate a country's capacity to export) and demand conditions (tariffs and distance sensitivity of products may influence a market's demand from one specific supplier). These factors will be taken into account to the extent possible.

Third, unmeasurable frictions drive a wedge between potential and actual exports. This final set of factors is often linked to a lack of market research such as difficulty complying with non-tariff measures (NTMs), a mismatch between product characteristics and consumer requirements and the difficulty of finding buyers.

While the export potential values account for factors that are measurable and beyond the control of trade advisers, the gap between potential and actual exports reflects frictions that can be addressed by trade advisers.

Components

The following sections describe the different components that constitute an exporter's potential export value for a given product in a given target market: supply, demand and ease of trade.

Supply

The supply side is based on a dynamic version of market share, corrected for some of the factors that distort the measure of 'true' export performance:

$$Supply_{ik}^{EPI} = Projected\ MS_{ik} \times GTD_{ik}$$

The first factor on the right-hand side is defined as $Projected\ MS_{ik} = \frac{v_{ik} \times \Delta GDP_i}{\sum_i(v_{ik} \times \Delta GDP_i)}$ where v_{ik} is a weighted average of country $i's$ exports of product k over a five-year period, with higher weights

[6] A correction is introduced later to take into account product-specific trade cost variations in the form of (i) tariffs and (ii) sensitivity to distance.
[7] See Annex I for more details.

placed on more recent years,[8] and ΔGDP_i is the expected change in country i's gross domestic product (GDP) six years into the future from the last year of the time series of historical trade data used to calculate current exports (v_{ik}). Including GDP growth in the supply calculation accounts for the fact that economic growth augments an exporter's capacity to export across all sectors.

Because market shares computed using reported trade data may not accurately reflect the real capacity of countries to competitively export products, we introduce a correction for tariff advantages.

Global tariff disadvantage

The last factor on the right-hand side represents the global tariff disadvantage (GTD) the country faces when exporting product k to the world market, $GTD_{ik} = \left(\frac{1+av.\ Tariff_{ik}}{1+av.\ Tariff_k}\right)^{\sigma_k}$, where $\sigma_k > 0$ is the substitution elasticity between suppliers of product k.

The denominator of the global tariff disadvantage is the weighted average tariff applied to product k (markets j imports of product k from suppliers i are used as weights in the calculation of average tariffs). The numerator is the weighted average tariff applied to country i's exports of product k. Country i has a tariff advantage if the denominator exceeds the numerator, i.e. if other suppliers face higher tariffs in the world market than country i for a given product k. The impact of a tariff depends on the sensitivity of trade to changes in prices, measured by σ_k, the substitution elasticity between suppliers of product k.[9]

Intuitively, large tariff advantages allow a country to export more, thus boosting its global market share. Uncorrected global market shares would therefore imply an overestimation of a country's true suppy capacity relative to other suppliers if the country faces a global tariff advantage and an underestimation if it faces a global tariff disadvantage. For this reason, the global market share is downgraded if a country has a global tariff advantage ($GTD_{ik} < 1$) and upgraded if a country has a global tariff disadvantage ($GTD_{ik} > 1$). On the supply side, we attempt to remove the effect of the global tariff advantage or disadvantage to better isolate a country's true supply capacity. On the demand side, however, we re-introduce a market tariff advantage factor to better estimate a market's demand for a good from a given exporter.

Demand

Demand conditions are captured through the combination of projected import values and factors accounting for the openness of the target market to the products exported by a given country:

$$Demand_{ijk} = Projected\ M_{jk} \times MTA_{ijk} \times Distance\ factor_{ijk}$$

where projected imports are given by:

$$Projected\ M_{jk} = v_{jk} \times \left(\frac{\Delta \mathrm{GDP}_j}{\Delta \mathrm{Pop}_j}\right)^{\mathrm{Em}_{dc}} \times \Delta \mathrm{Pop}_j \times \zeta_{dc} \times Tariff\ reduction_{jk}$$

v_{jk} is a weighted average of market $j's$ imports of product k over a five-year period, with higher weights placed on more recent years,[10] and $\left(\frac{\Delta GDP_j}{\Delta Pop_j}\right)$ is the expected change in GDP per capita six years into the future from the last year of the times series of historical trade data used to calculate current imports (v_{jk}). Em_{dc} is the revenue-elasticity of import demand per capita, estimated separately by

[8] Please refer to the data section and Annex VII for more details.

[9] While tariffs come from the ITC Market Access Map database, price elasticities are taken from the GTAP database that distinguishes 43 goods sectors.

[10] Please refer to the data section and Annex VII for more details.

development level d and HS 2-digit chapter c (see Annex III) to account for the possibility that import demand in developing countries reacts differently to a per-capita GDP growth than in developed countries and likewise, that import demand for luxury goods reacts differently than import demand for essential commodities. $\Delta\mathrm{Pop}_j$ is the change in population, impacting projected imports with a unitary elasticity. The factor $\zeta_{dc} = e^{\frac{22}{3}*\alpha_{dc}}$ accounts for a residual trend in import growth, where α_{dc} measures the average annual growth of imports because of factors unrelated to GDP per capita growth, such as inflation.

The final factor takes into account expected tariff reductions in market j for product k. $Tariff\ reduction_{jk} = \left(\frac{1+av.\ tariff_{jk}}{1+av.\ tariff_{jk}^{FL}}\right)$, with $av.\ tariff_{jk}^{FL}$ equal to the average forward-looking tariff applied by market j to product k.[11] Whenever $av.\ tariff_{jk}^{FL} < av.\ tariff_{jk}$, the tariff reduction factor will be positive, leading to an increase in the market's import potential of the concerned product. The price elasticity of import demand is assumed to be equal to one for every product.[12]

Market tariff advantage

The second factor on the right-hand side of the demand equation represents the market tariff advantage (MTA) or the tariff advantage that a given country enjoys in the target market, $MTA_{ijk} = \left(\frac{1+av.\ tariff_{jk}}{1+tariff_{ijk}}\right)^{\sigma_k}$. Forward-looking tariffs are used here, as we project the expected market tariff advantage that a given country will enjoy in the target market to the same year to which we project global market share and import demand.

The denominator of the tariff advantage in the target market is the tariff applied by market j to imports of product k from country i. The numerator is the weighted average tariff applied by market j to imports of product k from all countries (market j imports of product k from suppliers i are used as weights). Country i has a tariff advantage if the numerator exceeds the denominator. The higher the advantage, the larger the potential to import the product from a given exporter. The magnitude of this impact depends on the substitution elasticity between suppliers of product k (σ_k). As mentioned above, while the global tariff disadvantage has been taken into account on the supply side, a market tariff advantage has been introduced on the demand side to better estimate a market's demand for a good from a given exporter.

Distance factor

The final factor of the demand equation is a product-specific distance indicator which compares market j's distance to country i with its distance to all suppliers of product k:

$$Distance\ factor_{ijk} = e^{-\left|av.\log distance_{jk} - \log distance_{ij}\right|}$$

In empirical estimations of bilateral trade flows, distances are used as a proxy for transportation costs. However, while transportation costs vary across products, bilateral distances are constant and hence, not useful when the objective is to differentiate products according to their export potential.[13] This is different for the average distance over which a product is traded. Perishable products, for instance, are sensitive to distance and therefore typically imported from neighbouring countries, while distance

---

[11] For more information on forward-looking tariffs, please consult the data section.

[12] The price elasticity of import demand lies in an interval between the price elasticity of demand for a product in a country (usually below one) and the substitution elasticity between domestic and imported goods in that country (usually above one). A unitary elasticity thus corresponds to the mid-point between both values.

[13] When available, data on product-specific transport costs are limited to single countries (e.g. inferred from CIF-FOB ratios available for Chile, Brazil, United States and Australia, see Pomfret and Sourdin, 2010) or to a single mode of transportation (as for example in the OECD Maritime Transport Cost database, available at: http://stats.oecd.org/Index.aspx?DataSetCode=MTC).

matters less for durable products. Information embedded in the average distance over which a product is traded can help identify the best products to export to a given market. The closer the match (the lower the absolute difference) between the exporter's distance to the target market and the average distance over which the target market imports the product, the higher will be the export potential of the product to the market.[14]

Ease of trade

Ease of trade is based on actual trade between country i and market j for products with potential relative to their hypothetical trade if exporter i had the same market share for its products in market j as it does in the world market.[15]

$$Ease_{ij} = \frac{v_{ij}}{\sum_k \left(Supply_{ik}^{EPI,Static} \times Demand_{ijk}^{Static}\right)}$$ [16]

v_{ij} is a weighted average of country i's total exports to market j over a five-year period, with higher weights placed on more recent years.[17]

If ease of trade > 1, country i trades more with market j than we would expect based on country $i's$ world market share for all products. This could be because the two countries are located in proximity, share the same language or culture, or have established commercial links. If ease of trade < 1, country i trades less with market j than we would expect based on country $i's$ world market share for all products. The larger the ease, the larger country i's potential to trade any type of product with market j, and vice versa.[18]

Export potential value and unrealized export potential

The export potential value follows directly from the combination of supply, demand and ease of trade:

$$EPI_{ijk} = Supply_{ik}^{EPI} \times Ease_{ij} \times Demand_{ijk}$$

A normalization similar to the one presented in equation (3) ensures that summation over i results in $EPI_{jk} = Projected\ M_{jk}$.

The comparison of potential export values with actual export values reveals unrealized opportunities that trade support institutions can address. It is calculated as:

$$Unrealized\ potential_{ijk} = EPI_{ijk} - \min(v_{ijk}, EPI_{ijk})$$

In case of $v_{ijk} > EPI_{ijk}$, the unrealized potential equals zero.

One part of the unrealized trade potential is explained by the fact that the methodology is forward-looking: it projects supply and demand into the future based on expected GDP and population growth, and forward-looking tariffs. Another part, however, is structural and likely results from a variety of frictions that hinder current trade, including:

[14] Distances are from the Centre des Etudes Prospectives et d'Informationales Internationales' (CEPII) GeoDist database (Mayer and Zignago, 2011). See the data section and Annex II for more details.
[15] Mineral primary products (HS Chapters 25 and 26) are excluded from the ease calculation as trade in raw commodities is usually not controlled by the country itself but by large foreign companies.
[16] See next section for information on static indicators.
[17] Please refer to the data section and Annex VII for more details.
[18] In cases of full trade embargos, we set ease to zero. This is currently the case for trade between Russia and Ukraine, and between the United States and Cuba.

- difficulty complying with non-tariff measures, including rules of origin,
- misalignment of supply with the price or quality preferences of consumers,
- difficulty finding buyers in the target market,
- suboptimal allocation of exports across potential target markets due to lack of market intelligence,
- seasonality.

These frictions affect the country's ability to trade a specific product with a specific market. In order to identify the structural part of the unrealized potential, a static version of the export potential is calculated by removing all forward-looking elements (GDP and population growth, and future tariffs) from the calculation of the supply and demand.

Product Diversification Indicator

The EPI results from a decomposition of potential export values into measures of market share, demand and ease of trade. Market shares can only be computed for products that are already exported. To identify diversification opportunities, linkages from a country's current comparative advantages to potential new ones are established using Hausmann and Hidalgo's concept of the product space (Hausmann and Klinger, 2007, Hausmann et al., 2007 and Hidalgo et al., 2007). In the PDI calculation, the average distance of a product from a country's current export basket replaces market share as a measure of supply capacity, while demand and ease of trade remain identical to the EPI calculation.

Components

The PDI starts from a country's current supply capacities but aims to identify—based on these capacities—products into which the country could diversify. The idea that a country's ability to export one product depends on its ability to export other products originates from Hausmann and Hidalgo's product space. This concept measures the relatedness of products (their "proximity") based on the observation that similar products are more likely to be produced by the same country than dissimilar products. Fundamental to this notion is the assumption that countries possess a set of capabilities that is specific to the production of its exported goods. The higher the overlap between this set of capabilities and another set specific to a new product, the higher the likelihood that the country can diversify its exports. Hausmann and Hidalgo's concept is purely outcome-based: if country i is currently able to export product l that is often found in the export baskets of other countries together with another product k, it will be relatively easy for country i to also start exporting product k.

Current capabilities are measured by Balassa's RCA in Hausmann and Hidalgo's original approach. To account for the shortcomings of reported trade data in reflecting "true" supply capacities, we propose a measure of comparative advantage (CA) that corrects for global tariff disadvantages and re-exports,

$$CA_{ik} = \frac{v'_{ik} \times \sum_{ik} v'_{ik}}{\sum_i v'_{ik} \times \sum_k v'_{ik}}$$

where $v'_{ik} = v_{ik} \times GTD_{ik}$.

The link, φ_{kl}, from product k to product l is then defined based on the conditional probability $P(.|.)$ that a country has a comparative advantage in product k if it already has one in product l:

$$\varphi_{kl} = P(CA'_{ik}|CA'_{il}) \quad \forall\, i$$

where $CA'_{il} = \begin{cases} 1 & if\ CA_{il} \geq 1 \\ 0 & otherwise \end{cases}$.[19] We use information on export bundles from all countries to calculate the average proximity of a country to a new potential export product, which reflects the country's supply capacity:

$$Density_{ik} = \frac{\sum_l (CA'_{il} \varphi_{kl})}{\sum_l \varphi_{kl}}$$

A high density means that the country has developed comparative advantages in many products surrounding product k and that it can therefore start exporting product k relatively easily. By contrast, a low value means that the product is far from the country's current export structure and hence, an unlikely candidate for export diversification.

Even products with very high densities cannot necessarily be exported: first, the trade data-based measure of density does not always reflect the presence of endowments that are necessary to start producing and exporting new products. Boschma and Capone (2014) and Kniahin (2014) have shown that while comparative advantage is positively associated with density, much of its variation is linked to other factors. We therefore modify the original product space approach and narrow the suggested diversification opportunities down based on information about natural resources:[20]

- A Global Trade Analysis Project (GTAP) dataset on land endowments is used to identify the relevant moisture regimes and climate zones necessary for the production of agricultural products. If the PDI suggests a product for a country that does not possess any of the relevant land types, the product will be excluded as a diversification opportunity (for details on this approach, see Annex IV).

- Sea-related products that are suggested as diversification opportunities for landlocked countries are excluded if they have not been exported with comparative advantage and positive trade balance by at least one landlocked country for the three most recent years used in the analysis.

- Transformed products that are (almost) entirely based on excluded raw products from Chapter 71 (pearls, precious stones, metals, coins, etc.) are not considered.[21] Similarly, inorganic chemicals, precious metal compounds and isotopes (Chapter 28) are excluded because their production requires rare earths or other elements that are not available in all countries.

Second, products that are feasible for the country may not be demanded by those markets with which the country finds it easy to trade. Or, even if they are in demand, the tariffs the country faces in this market may be unusually high. Harrison and Rodriguez-Clare (2011) and Lederman and Maloney (2012) argue that demand factors could outweigh the overall low density of a country that is located

[19] Note that the inventors of the product space suggest taking the minimum of both conditional probabilities $P(CA'_{il}|CA'_{ik})$ and $P(CA'_{ik}|CA'_{il})$ to avoid cases where a country is the only exporter of a particular good which would imply "that the conditional probability of exporting any other good given that one would be equal to one for all other goods exported by that country" (Hidalgo et al., 2007, p. 2 of the annex). In the context of this study both versions were computed and differences between them appear minimal. To save computational time and remain closer to the idea that linkages are based on a country's current exports, unidirectional conditional probabilities are used.

[20] Note that manufactured products are not filtered. Even though a country may currently not have the right resources to produce the good, it could invest into their future development (or receive foreign investment), thus bringing it closer to realizing the diversification potential.

[21] Please refer to the data section for a discussion on the types of excluded products and to Annex V for a complete list of excluded products.

in an isolated area of the product space, in particular if there is little competition for the products 'within reach'.[22] This trade-off is not addressed in the original product space literature and its applications. To identify diversification opportunities with good chances of export success, densities are combined with the same demand side and ease of trade indicators that are used in the EPI.

Product diversification rank

Densities differ from market shares in various ways and therefore have to be normalized before being combined with demand and ease of trade factors. First, the PDI supply side varies on a much smaller scale across products for a given country than the corrected market shares used in EPI. Without adjustment, the importance of the supply relative to the demand component would differ from that in the computation of the EPI. The transformation $Density'_{ik} = Density_{ik}^{\frac{\log Supply^{EPI}_{ik_1} - \log Supply^{EPI}_{ik_N}}{\log Density_{ik_1} - \log Density_{ik_N}}}$ ensures that $\frac{Density'_{ik_1}}{Density'_{ik_N}} = \frac{MS_{ik_1}}{MS_{ik_N}}$ where N is the rank of the last product for which comparative advantage is computed and MS_{ik} is the share of country i in the world market of product k.

While market share is a relative concept, density is an absolute measure, meaning that many countries can have a high density around the same product. Yet, a rough correspondence with total expected export capacity by country is desirable. Densities are therefore rescaled as follows:

$Density''_{ik} = Density'_{ik} \times \frac{\sum_{jk} Exp\ v_{ijk}}{\sum_{jk} PDI_{ijk}}$ so that $PDI'^{23}_{i} = Exp\ v_i$

where $Exp\ v$ corresponds to trade augmented by expected GDP growth:

$$Exp\ v_{ijk} = v_{ijk} \times \Delta GDP_i$$

Assuming that the new supply indicator is given by normalized density, $Supply^{PDI}_{ik} = Density''_{ik}$ and combining it with ease of trade and demand gives:

$$PDI_{ijk} = Supply^{PDI}_{ik} \times Ease_{ij} \times Demand_{ijk}$$

A final normalization is applied to $Supply^{PDI}_{ik} \times Ease_{ij}$ to ensure that market shares sum up to 1 in every market and $PDI_{jk} = Projected\ M_{jk}$. PDI values are then ranked by country, market, or product to suggest diversification opportunities.

Aggregation and presentation of results

Export and diversification potential values are calculated for every country × product × market combination. To aggregate at the level of sectors or regions, potential values can be simply summed up,

$$EPI(PDI)_{IJK} = \sum_{i \in I, j \in J, k \in K} EPI(PDI)_{ijk}$$

where I is a set of countries, J a set of markets, and K a set of products.

In the EPI, the extent of unrealized potential can be aggregated as well:

---

[22] If a country is located in an isolated area of the product space, its density is generally lower than the density of a country in a populated area, regardless of the product.

[23] $PDI'_i = \sum_{jk} (Density''_{ik} \times Ease_{ij} \times Demand_{ijk})$

$$Unrealized\ potential_{IJK} = \sum_{i \in I, j \in J, k \in K} EPI_{ijk} - \sum_{i \in I, j \in J, k \in K} \min\left(v_{ijk}, EPI_{ijk}\right)$$

Taking the minimum ensures that situations in which actual exports exceed export potential do not offset unrealized potential elsewhere. For a group of countries, markets or products, the unrealized potential therefore often exceeds the difference between EPI and current exports.

Data

Export potential assessments can be implemented for any country for which trade data is available at a detailed product level. The following sections explain which data is used and how it is processed in order to obtain reliable results.

Sources

Trade data comes from the ITC Trade Map database, which in turn sources statistics from UN Comtrade, the world's largest database of trade statistics, maintained by the United Nations Statistics Division (UNSD), and complemented with national sources.

Tariffs come from the ITC Market Access Map. Forward-looking tariffs account for the tariff reduction schedules of 281 agreements that apply to 122 markets. Whenever forward-looking tariffs are not available, current tariffs are used.

Product-level substitution elasticities are taken from the GTAP database, and distances from the CEPII GeoDist database.

GDP and population growth projections come from the World Economic Outlook (IMF) and the Population Division of the United Nations.

Coverage

Export potential assessments currently cover 226 countries and territories. The analyses start from the full list of products as defined at the 6-digit level of the Harmonized System (HS) classification. Products that are not compatible with international conventions (e.g., products that have damaging effects on the environment),[24] are based on mineral fuels or mineral oils, or are not in line with ITC's work programme (e.g. wastes, antiques, etc.) are excluded. Please refer to Annex V for a full list of excluded products. In light of revisions to the HS classification, consistency across years and countries is achieved by introducing product groups that encompass all items that are linked to a revised HS code.[25] There is one exception to this rule: product groups that spread over different HS2 chapters are split into sub-groups so that every group is encompassed by a single chapter.[26] Export potential assessments therefore distinguish 4,490 HS6-based product groups instead of the 5,387 HS6 products that form the HS 2017 revision.[27]

Data treatment

Re-exports

Declared exports often comprise re-exports that are not linked to any capacity of the country to produce the good. The aim of export potential assessments is to give guidance on trade support for sectors that have a significant production base in the country. In an effort to identify and exclude re-exports from the analysis, a filter is applied to products in chapter 73 and chapters 84 to 90 of the Harmonized System (HS). A product is excluded from the list of potential export products for a given country if it meets one of the following conditions:

---

[24] The World Customs Organization (WCO) publishes correlation tables between the HS and several international conventions, available at: http://www.wcoomd.org/en/topics/nomenclature/instrument-and-tools/interconnection-table.aspx.

[25] Note that HS 1992 (revision 0) and HS 1996 (revision 1) are no longer in use. Product groups are therefore built on redefined codes between HS 2002 and HS 2017.

[26] Manual checks have indicated that misallocations of trade in these cases are negligible.

[27] A correspondence table between HS 6-digit codes and ITC product groups is available here: https://exportpotential.intracen.org/en/resources/correspondences.

- Imports are at least twice as large as exports, and the product's share in the country's imports is at least 20 times the product's share in world imports;
- Imports are at least 20 times as large as exports;
- Imports are at least twice as large as exports, and the world market share of the country in the product is less than 0.01%.

Returned merchandise

Reported trade statistics often include returned merchandise (e.g. as a result of non-compliance with safety regulations). Even if these trade flows are small, they could distort the interpretation of results. For instance, exporters of some products may appear as significant importers of these same products although they are only re-importing their own products. To eliminate returned merchandise from trade data, bilateral imports (exports) of an HS 6 product from a given partner in a given year are removed when they are only a small fraction (< 1/15) of the country's exports (imports).

Reliability of trade flows

Four data reports are normally available for every country pair: country A's declarations of exports to and imports from country B, and country B's declarations of exports to and imports from country A. Ideally, what A exports to B (as declared by country A) should match what B imports from A (as declared by country B) with only a small difference stemming from the fact that import values include insurance and freight costs (CIF) while export values are reported free on board (FOB). In reality, however, differences between direct reports and the so-called mirror reports are often substantial and it is not straightforward to identify the best source of data.

Export potential assessments aim to rank products according to their export or diversification potential. As these rankings could eventually determine to which projects public funds will be allocated, a major concern is the identification of false positives, or products that appear interesting from a data perspective but are not very promising in reality. Export potential assessments rely on a set of measures to ensure that misreported data do not affect the robustness of results. First, trade flows use direct data, mirror data, or a combination of both, depending on the reporting reliability of a country and its trade partner (for details, please see Annex VI). Second, to even out incorrect data reports in single years, a weighted average of exports and imports over the past five years is used in the calculation of indicators, with higher weights placed on more recent years. To remain in the analysis, a product must be exported in the three most recent years and imported in the five most recent years.

Missing data

When data on tariffs, distances or GDP are missing, the following treatments are adopted:

- Tariffs: We assume Most Favoured Nation (MFN) treatment when countries do not report preferential tariffs applied on specific products.

- Distances: Gaps in the CEPII database are filled with great circle distances calculated from latitudes and longitudes of the country's capitals. In very rare cases, the distance indicator cannot be calculated because markets import the product only from areas not elsewhere specified. These products are assigned the average distance correction factor for other products exported by the country to those markets.

- Economic growth: If GDP forecasts for specific countries are missing, the predicted world GDP growth rate is used. If GDP forecasts are available for at least one year while missing for other years, a growth rate of 0% is used in the missing years.[28]

Thresholds

A product must have at least $200,000 in export potential or be among the products that account for 95% of the total export potential of a given exporter. Likewise, we apply a filter to only show the top 95% of products in terms of product diversification.

Empirical assessment[29]

The ability of the EPI to predict future trade outcomes can be tested empirically. To this end, trade at the extensive and intensive margins is regressed on historical EPIs, computed for 13 years, from 2006 to 2018. To focus beyond growth-driven patterns of EPI, the regressions test the static—or current—version of EPI.[30] Regressions with the total version of EPI are also performed, as robustness checks. If the indicator performs as expected, countries should increase their exports more in sectors and markets with high (unrealized) export potential than in sectors and markets with low (unrealized) export potential.

Empirical strategy

The performance of the indicator is assessed differently depending on whether trade already exists between country i and market j in sub-sector s.[31] The dataset is thus divided into two different samples.

Intensive margin

The first part of the empirical analysis restricts the dataset to entries for which trade is strictly positive between countries i and j in sub-sector s at time t. This restriction allows us to determine, when trade already exists, whether the indicator predicts future levels of trade, controlling for past export growth. The model is specified as follows:

$$\dot{X}_{ijs,t+3} = \beta_0 + \beta_1 D2P_{ijs,t} + \beta_2 \dot{X}^{p}_{ijs,t} + \gamma_i + \theta_j + \delta_t \mu_s + \epsilon_{ijs} \qquad (4)$$

with:

$$\dot{X}_{ijs,t+3} = log(\frac{(X_{ijs,t+1} + X_{ijs,t+2} + X_{ijs,t+3})/3}{X_{ijs,t}})$$ [32]

and

[28] When economic growth forecasts are missing for some years while being available in other years, it is often the result of a big economic shock, such as the civil war in Syria or the economic crisis in Lebanon. Because of the uncertainty created by the shock, no GDP forecast can be safely computed.

[29] This section is based on Périllat (2018) and summarizes the main results of this study.

[30] As described in the section "Export potential value and unrealized export potential" of this document.

[31] Products are classified into 76 sub-sectors and 17 sectors. See exportpotential.intracen.org/en/resources/data/correspondences for full correspondence table. All empirical models are also computed using sector and HS2 chapter aggregations. Results are robust to these different aggregations. Tables are available upon request.

[32] Note that using a log specification implies that observations with no exports for the following three years but positive exports at time t are not taken into account. They represent about 2% of all observations and are thus unlikely to substantially alter the results.

$$\dot{X}^{p}_{ijs,t} = log(\frac{X_{ijs,t}}{(X_{ijs,t-1} + X_{ijs,t-2} + X_{ijs,t-3})/3})$$

$X_{ijs,t}$ are exports of products in sub-sector s from country i to market j at time t. The dependent variable thus measures the growth rate of exports between year t and the average of the following three years.[33] It is regressed on $D2P_{ijs,t}$, the share of static unrealized in total export potential of products in sub-sector s from i to j at time t, and on $\dot{X}^{p}_{ijs,t}$, a measure of the past growth rate of exports. Country (market) fixed effects account for differences in growth rates in a given i (j) for any s, t and j (i).[34] . Time × sub-sector fixed effects account for sector trends affecting all countries.

The main coefficient of interest in this model is β_1, which measures the predictive power of the share of current unrealized potential on the future growth rate of exports. Note that the coefficient does not reveal any causality: It is not directly the share of unrealized potential that triggers exports, but efforts by countries to facilitate trade or by businesses to tap into opportunities. However, what may be concluded from this model is whether the indicator is a better predictor than a 'naive' one, $\dot{X}^{p}_{ijs,t}$, measuring a linear extrapolation of exports. β_1 and β_2 can be compared to observe which of the two variables has the strongest effect, and whether the indicator explains any variation of trade on top of what is explained by past exports.

As an extension, the total version of the unrealized potential (including the dynamic component) is used. Even if intuitively we can expect the effect of both indicators to be of the same sign and magnitude, it remains interesting to see whether there are differences between the two models. Since the total version of the unrealized potential takes into account expected GDP growth and expected population growth, there is no more need for country and market fixed effects as they are captured within the new components of the indicator. The revenue elasticity contained in the total indicator being computed separately for each HS2 chapter c, the time × sector fixed effects are also not necessary anymore. A specification with total unrealized potential and fixed effects is also run, as robustness check.[35]

Extensive margin

The second part of the empirical analysis focuses on sub-sectors in which no product is exported from i to j at time t (considering only products for which EPI is computed, i.e. that are already exported to other markets at time t). First, the likelihood of the emergence of trade in a particular sub-sector within the following three years is regressed on the log of the current EPI at time t:

$$T_{ijs,t+3} = \beta_0 + \beta_1 log(EPI_{ijs,t}) + \gamma_i + \theta_j + \delta_t \mu_s + \epsilon_{ijs} \quad (5)$$

with:

$$T_{ijs,t+3} = \begin{cases} 1 \ if \ X_{ijs,t+1} + X_{ijs,t+2} + X_{ijs,t+3} > 0 \\ 0 \ if \ X_{ijs,t+1} + X_{ijs,t+2} + X_{ijs,t+3} = 0 \end{cases}$$

[33] We have tested alternative time periods, such as growth over one or two years instead of three. Although coefficients are smaller, they remain positive and significant. Tables are available and will be shared upon request.

[34] A specification with ij fixed effects instead of separate i and j fixed effects has been tested on a random 10% sample of exporter-market pairs. Results remain positive and highly significant. Tables are available and will be shared upon request.

[35] Results are very similar. Detailed tables are available upon request.

This regression is estimated with a linear probability model (LPM). However, the LPM has drawbacks, such as its predicted probabilities that are not bounded between zero and one. A probit model is thus used as well in order to take into account the binary nature of the dependent variable. The same fixed effects are added.

Second, the value of exports of a sub-sector between i and j in the next three years with no trade at time t is regressed on the EPI, under the condition that trade is strictly positive in at least one of these following years:

$$log(X_{ijs,t+1} + X_{ijs,t+2} + X_{ijs,t+3}) = \beta_0 + \beta_1 log(EPI_{ijs,t}) + \gamma_i + \theta_j + \delta_t \mu_s + \epsilon_{ijs} \quad (6)$$

Conditioning on

$$X_{ijs,t+1} + X_{ijs,t+2} + X_{ijs,t+3} > 0 \text{ and } X_{ijs,t} = 0$$

While equation (5) assesses whether export potential can predict the emergence of a new bilateral trade flow, equation (6) assesses the extent of the latter.[36]

Results

Intensive margin

Results of the regressions at the intensive margin are shown in Table 1. Regressing $\dot{X}_{ijs,t+3}$ only on the current share of unrealized potential and fixed effects result in β_1 at 0.59. When controlling for the past growth rate, β_1 increases slightly to about 0.62, meaning that a sub-sector—in a given country i and market j at time t—with a 10 percentage points higher share of current unrealized potential will see the exports of its products increase by about 6.2%[37] in the next three years compared to another sub-sector with the same past growth rate. β_2 is negative and significant, suggesting that past growth of exports—taking the last three years into account—is negatively correlated with future growth of exports. The share of unrealized potential thus explains variation on top of what is explained by past growth rates. Using total instead of current untapped potential leads to similar coefficients. β_1 is slightly higher, suggesting a better performance of the total indicator.

Table 1: Effect of the unrealized potential on the future export growth rate

	(1)	(2)	(3)	(4)
	$\dot{X}_{ijs,t+3}$	$\dot{X}_{ijs,t+3}$	$\dot{X}_{ijs,t+3}$	$\dot{X}_{ijs,t+3}$
Current D2P	0.586***	0.615***		
	(0.003)	(0.003)		
$\dot{X}^{p}_{ijs,t}$		–0.386***		–0.371***
		(0.001)		(0.001)

[36] As a robustness check, we have re-estimated all models removing outliers. Coefficients change slightly but remain highly significant. Results are available upon request.

[37] D2P can take on values between zero and one. A unit increase in D2P thus represents an increase of 100 percentage points of the share of unrealized potential.

Total D2P			0.669***	0.690***
			(0.003)	(0.003)
Constant	-0.043	0.087	−0.334***	−0.295***
	(0.058)	(0.060)	(0.002)	(0.002)
N	3,754,273	3,567,521	3,754,273	3,567,521
Fixed effects	Yes	Yes	No	No

Note: $* \, p < 0.10$, $** \; p < 0.05$, $*** \, p < 0.01$. ijs-level clustered standard errors in parentheses.
Values rounded to three digits, when non zero.

Extensive margin

Table 2 shows the results of regression (5). The LPM, presented as a benchmark, leads to coefficients of 0.023 and 0.028 for respectively the static and the total EPI. When the log of the static (total) version of the EPI increases by 1%, the future probability of trade increases by 2.3 (2.8) percentage points. Both indicators thus appear to have a significant positive impact on the future probability of trade.

Once again, the total EPI outperforms the static EPI: Since the total EPI usually exceeds the static EPI, for both indicators to have a similar performance, the coefficient of the static version should be higher. Here however, the coefficient for the static indicator is lower, suggesting a better performance of the total indicator.

Table 2: Effect of the EPI on the probability of future trade, LPM

	(1)	(2)
	$T_{ijs,t+3}$	$T_{ijs,t+3}$
Current EPI	0.023***	
	(0.0001)	
Total EPI		0.028***
		(0.0001)
Constant	0.043***	0.133***
	(0.006)	(0.0002)
N	9,955,827	9,955,827
Fixed effects	Yes	No

Note: $* \, p < 0.10$, $** \; p < 0.05$, $*** \, p < 0.01$. ijs-level clustered standard errors in parentheses.
Values rounded to three digits, when non zero. Static and total EPI specified in log terms.

Table 3 presents the average marginal effect of the EPI (in logs) on the trade dummy in the first and third columns, and the marginal effect at the mean value of the log of the EPI in the second column.[38]

[38] The marginal effect at the mean value of the log of the EPI is not reported for the static indicator as its computation is too demanding.

The probit model confirms the previous result. The EPI predicts well the emergence of future trade. A unit increase in the log of the total version of the EPI (i.e. a multiplication of EPI by e, or about 2.718) leads to an average increase of about 3.3 percentage points in the probability of future trade. At the average value of the EPI, it is slightly lower, at around 3 percentage points. The static EPI yields similar coefficients.

The last empirical model restricts the sample to sub-sectors in which trade is non-existent at time t between i and j but strictly positive during the following three years. Regression (6) looks at whether the initial export potential value has an impact on future exports. The results are shown in Table 4. Both versions of the indicator have a positive and significant impact on future trade. A 1% increase in the static (total) EPI will increase future exports by about 26% (24%).

Table 3: Marginal effects of the EPI on the probability of future trade, probit model

	(1) AME	(2) MEM	(3) AME
Total EPI	0.033***	0.030***	
	(0.0001)	(0.0001)	
Current EPI			0.030***
			(0.0001)
N	9,955,827	9,955,827	9,955,819
Fixed effects	No	No	Yes

Note: $* \, p < 0.10$, $** \, p < 0.05$, $*** \, p < 0.01$. ijs-level clustered standard errors in parentheses. Values rounded to three digits, when non zero. Total and static EPI specified in log terms. AME refers to average marginal effects, MEM to marginal effects at the mean.

Table 4: Impact of the present EPI on the extent of future trade

	(1) Future exports	(2) Future exports
Total EPI	0.258***	
	(0.001)	
Current EPI		0.243***
		(0.001)
Constant	1.839***	4.539***
	(0.005)	(0.116)
N	1,380,692	1,380,692
Fixed effects	No	Yes

Note: $* \, p < 0.10$, $** \, p < 0.05$, $*** \, p < 0.01$. ijs-level clustered standard errors in parentheses. Values rounded to three digits, when non zero. All variables specified in log terms.

Conclusion

In response to the need of countries to learn about export opportunities, ITC has developed an export potential and diversification assessment methodology. Based on supply conditions in the exporting country, and demand and market access conditions in the target market, it supports and brings transparency to the selection of projects for export promotion activities. The methodology focuses on products which are already competitively exported by a country (through the EPI) or on products which a country does not yet competitively export but which are likely candidates for diversification (through the PDI).

Conceptually, export potential assessments are based on a decomposition of potential export values into country × product, product × market and country × market factors. They resemble a gravity equation specified at the product level but allow for more detailed interpretation and therefore more adequate policy recommendations. All analyses are based on data that have undergone a thorough treatment to ensure that results are not driven by misreported trade flows, measurement errors or misattribution.

The empirical analysis confirms the performance of the indicator. For existing trade, the share of unrealized potential can predict to some extent the future growth rate of exports. When there is no current trade, the EPI indicates whether trade will occur in the future and how much of it. Overall, the performance of the EPI confirms that the indicator successfully points at markets and sectors in which additional exports are feasible.

While export potential assessments are a quantitative approach to identify promising export sectors, suppliers and markets on a global scale, other, often intangible factors may also influence the choice of where to focus trade support measures. These include for example, the willingness and ability to attract foreign direct investment, the possibilities of marketing and branding and the existence of synergic development plans or sector strategies. Export potential assessments also abstract from the costs related to export promotion activities. All these factors may however influence the feasibility of exporting certain products. Trade advisors should therefore look at results with a critical eye and engage in further investigations, notably regarding the suggested options for diversification.

In summary, export potential assessments provide a starting point in an export promotion decision-making process that should be complemented with desk research and consultations with public and private sector stakeholders in the country.

References

Andersen, P. H., Strandskov, J. (1998), *International market selection: A cognitive mapping perspective.* Journal of Global Marketing, 11(3): 65-84.

Armstrong, S. P. (2007), *Measuring trade and trade potential: A survey*. Crawford School Asia Pacific Economic Paper, (368).

Atif, R. M., Haiyun, L., Mahmood, H. (2017), *Pakistan's agricultural exports, determinants and its potential: an application of stochastic frontier gravity model.* The Journal of International Trade & Economic Development, 26(3): 257-276.

Atif, R. M., Mahmood, H., Haiyun, L., Mao, H. (2019), *Determinants and efficiency of Pakistan's chemical products' exports: An application of stochastic frontier gravity model*. PloS one, 14(5), e0217210.

Baldwin, R. (1994), *Towards an Integrated Europe*, Centre for Economic Policy Research, London.

Barma, T. (2017), *Efficiency of India's agricultural exports: A stochastic panel analysis*. South Asia Economic Journal, 18(2): 276-295.

Berthou, A., Emlinger, C. (2011), *The Trade Unit Values Database*, CEPII Working Paper 2011-10.Boschma, R., Capone., G. (2014), *Relatedness and Diversification in the EU-27 and ENP countries*, Utrecht University Papers in Evolutionary Economic Geography 14.07.

Brouthers, L. E., Nakos, G. (2005), *The role of systematic international market selection on small firms' export performance,* Journal of Small Business Management, 43(4): 363-381.

Brouthers, L. E., Mukhopadhyay, S., Wilkinson, T. J., Brouthers, K. D. (2009), *International market selection and subsidiary performance: A neural network approach,* Journal of World Business, 44(3): 262-273.

Bussière, M., Callegari, G., Ghironi, F., Sestieri, G., Yamano, N. (2011), *Estimating Trade Elasticities: Demand Composition and the Trade Collapse of 2008-09,* American Economic Journal: Macroeconomics, American Economic Association, 5(3), 118-51.

Cano, J. A., Campo, E. A., Baena, J. J. (2017a*), Application of DEA in international market selection for the export of goods*. Dyna, 84(200): 376-382.

Cano, J., Campo, E., Gómez-Montoya, R. (2017b), *International market selection using fuzzy weighing and Monte Carlo simulation,* Polish Journal of Management Studies, 16(2), 40.

Clavijo, F., Faini, R. (1989), *Differentiating cyclical and long-term income elasticities of import demand.* World Bank Policy Research Working Paper, (197).

Cuyvers, L., Steenkamp, E., Viviers, W. (2012), *The methodology of the Decision Support Model (DSM)*, in: Cuyvers, L., Viviers, W. (eds.) Export Promotion: a Decision Support Model Approach. South Africa: Sun Media Metro.

Cuyvers, L., Steenkamp, E., Viviers, W., Rossouw, R., Cameron, M. (2017), *Identifying Thailand's high-potential export opportunities in ASEAN+ 3 countries.* Journal of International Trade Law and Policy, 16(1): 2-33.

Dadakas, D., Ghazvini Kor, S., Fargher, S. (2020), *Examining the trade potential of the UAE using a gravity model and a Poisson pseudo maximum likelihood estimator*. The Journal of International Trade & Economic Development, 29(5): 619-646.

Deaza, J. A., Díaz, N. F., Castiblanco, S. E., Barbosa, M. I. (2020), *International market selection models: a literature review*. Tendencias, 21(2): 191-217.

Donges, J., Riedel, J. (1977), *The expansion of manufactured exports in developing countries: An empirical assessment of supply and demand issues*, Review of World Economics (Weltwirtschaftliches Archiv), 113(1), 58-87.

Drysdale, P., Huang, Y., Kalirajan, K. P. (2000), *China's trade efficiency: measurement and determinants*, in: Drysdale, P., Yuling, Z., Song, L. (eds.) APEC and liberalisation of the Chinese economy. Canberra: Asia Pacific Press.

Ebaidalla, E. M., Ali, M. E. M. (2022), *Assessing Intra-Arab Trade Integration and Potential: Evidence from the Stochastic Frontier Gravity Model*. The International Trade Journal.

Egger, P. (2002), *An Econometric View of the Estimation of Gravity Models and the Calculation of Trade Potentials*, World Economy 25(2), 297-312.

Fontagné, L., Guimbard, H., Orefice, G. (2022). Tariff-based product-level trade elasticities. Journal of International Economics, 137: 103593.

Fontagné, L., Pajot, M., Pateels, J. (2002), *Potentiels de commerce entre économies hétérogènes : un petit mode d'emploi des modèles de gravité*, Economie internationale, 1(152-153), 115-139.

Fortunato, P., Razo, C., Vrolijk, K. (2015), *Operationalizing the Product Space: A Road Map to Export Diversification*, UNCTAD Discussion Papers No. 219.

Gros, D., Gonciarz, A. (1996), *A note of the Trade Potential of Central and Eastern Europe*, European Journal of Political Economy, 12(4): 709-721.

Gupta, V., Kumar, E. (2021), *Review on Machine Learning Techniques for International Trade Trends Prediction,* 3rd International Conference on Advances in Computing, Communication Control and Networking (544-547), December 2021, *IEEE, India*.

Harrison, A., Rodríguez-Clare, A. (2011), *Trade, Foreign Investment, and Industrial Policy for Developing Countries*, NBER Working Paper No. 15261.

Hausmann, R., Chauvin, J. (2015). *Moving to the adjacent possible: Discovering paths for export diversification in Rwanda*. CID Working Paper Series, (294).

Hausmann, R., Hidalgo, C. A., Bustos, S., Coscia, M., Simoes, A. (2014), The Atlas of Economic Complexity: Mapping Paths to Prosperity. MIT Press.

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

Hausmann, R., Klinger, B. (2007), *Structural Transformation and Patterns of Comparative Advantage in the Product Space*, Harvard Centre for International Development Working Paper No. 128.

Hausmann, R., Rodrik, D. (2006), *Doomed to choose: Industrial policy as predicament*, Center for International Development Blue Sky Conference, September 9 2006*, Cambridge, MA.*

Hausmann, R., Santos, M., Barrios, D., Taniparti, N., Tudela, J., Lu, J. (2022). *The Economic Complexity of Namibia: A Roadmap for Productive Diversification*. CID Faculty Working Paper Series, (410).

Head, K., Mayer, T. (2014), *Gravity Equations: Workhorse, Toolkit, and Cookbook*, in: Gopinath, G., Helpman, E., Rogoff, K. (eds.) Handbook of International Economics vol. 4. Elsevier, 131-195.

Helmers. C., Pasteels, J. (2006), *Assessing Bilateral Trade Potential at the Commodity Level: An Operational Approach*, ITC Working Paper.

Hertel, T.H., Hummels, D., Ivanic, M., Keeney, R. (2004), *How Confident Can We Be in CGE-Based Assessments of Free Trade Agreements?,* GTAP Working Paper No. 26.

Hidalgo, C. A., Balland, P. A., Boschma, R., Delgado, M., Feldman, M., Frenken, K., ... Zhu, S. (2018), *The principle of relatedness*. In International conference on complex systems (pp. 451-457). Springer, Cham.

Hidalgo, C., Klinger, B., Barabasi, A.L., Hausmann, R. (2007), *The product space conditions the development of nations*, Science 317, 482-487.

Houthakker, H. S., Magee, P. (1969), Income and Price Elasticities in World Trade, The Review of Economics and Statistics, 51(2), 111-125.

Kniahin, D. (2014), *Modeling International Trade Patterns with the Product Space Methodology*, Master thesis.

Kumar, S., Prabhakar, P. (2017), *India's trade potential and free trade agreements: a stochastic frontier gravity approach.* Global economy journal, 17(1): 20160074.

Lederman, D., Maloney, W. (2012), *Does What You Export Matter? In Search of Empirical Guidance for Industrial Policies*, Washington, D.C.: The World Bank.

Mayer, T., Zignago, S. (2011), *Notes on CEPII's distances measures: the GeoDist Database*, CEPII Working Paper 2011-25.

Melchior, A., Zheng, J., Johnsen, A. (2009), *Trade barriers and export potential: Gravity estimates for Norway's exports*, Paper written for the Ministry of Trade and Industry, Norway.

Mulabdic, A., Yasar, P. (2021), *Gravity Model-Based Export Potential: An Application to Turkey*. World Bank Policy Research Working Paper, (9557).

Musso, F., Francioni, B. (2014), *International strategy for SMEs: Criteria for foreign markets and entry modes selection*. Journal of Small Business and Enterprise Development, 21(2): 301-312.

Nasir, S., Kalirajan, K. (2014), *Modern Services Export Performances among Emerging and Developed Asian Economies*. Asian Development Bank Working Paper Series on Regional Economic Integration, (143).

Nilsson, L. (2000), *Trade integration and the EU economic membership criteria.* European Journal of Political Economy, 16(4): 807-827.

Obeng, C. (2022), *Export Efficiency and Diversification in Ghana*. African Economic Research Consortium Policy Brief, (771).

Papadopoulos, N., Chen, H., Thomas, D. R. (2002), *Toward a tradeoff model for international market selection*. International business review, 11(2): 165-192.

Papadopoulos, N., Martín, O. M. (2011), *International market selection and segmentation: perspectives and challenges*. International Marketing Review, 28(2): 132-149.

Périllat, S. (2018), *The Export Potential Indicator: an Assessment of its Predictive Power*, Master thesis.

Pinheiro, F. L., Alshamsi, A., Hartmann, D., Boschma, R., Hidalgo, C. (2018), *Shooting low or high: do countries benefit from entering unrelated activities?*. Papers in Evolutionary Economic Geography, 18(07).

Pomfret, R., Sourdin, P. (2010), *Trade Facilitation and the Measurement of Trade Costs*, Journal of International Commerce, Economics and Policy, 1(1), 145-163.

Ravishankar, G., Stack, M. M. (2014), *The Gravity Model and Trade Efficiency: A Stochastic Frontier Analysis of Eastern European Countries' Potential Trade*. The World Economy, 37(5): 690-704.

Saen, R. F. (2011), *International market selection using advanced data envelopment analysis.* IMA Journal of Management Mathematics, 22(4): 371-386.

Shabani, A., Saen, R. F., Vazifehdoost, H. (2013), *The use of data envelopment analysis for international market selection in the presence of multiple dual-role factors*. International Journal of Business Information Systems, 13(4): 471-489.

Sheng, S. Y., Mullen, M. R. (2011), *A hybrid model for export market opportunity analysis. International Marketing Review*, 28(2): 163.

Shepotylo, O. (2009), *Gravity with Zeros: Estimating Trade Potential of CIS countries*, Kyiv School of Economics Discussion Paper No. 16.

Sotiros, D., Rodrigues, V., Silva, M. C. (2022), *Analysing the export potentials of the Portuguese footwear industry by data envelopment analysis.* Omega, 108(C).

Stack, M. M., Pentecost, E. J., Ravishankar, G. (2018), *A stochastic frontier analysis of trade efficiency for the new EU member states: Implications of Brexit*. Economic Issues, 23(1): 35-53.

Vollrath, T. (1991), *A Theoretical Evaluation of Alternative Trade Intensity Measures of Revealed Comparative Advantage*, Review of World Economics (Weltwirtschaftliches Archiv), 127(2).

Annex I: Theoretical underpinnings of the decomposition of potential export shares

The export potential and diversification assessment methodology is based on the assumption that actual world trade differs from what it would be in a world without frictions (such as non-tariff trade measures). Its objective is to provide a means to estimate how trade flows would be distributed in this idealized world.

The paragraphs below show that under certain assumptions world trade values would take the following simple multiplicative form:

$$v_{ijk} = \alpha_{ik}\beta_{ij}\gamma_{jk}$$

With:

- i Exporting country
- j Market (i.e., importing country)
- k Product (4,490 products based on the HS6 nomenclature)
- v_{ijk} Trade value

The three factors can be interpreted as follows:

- α_{ik} Relative supply performance, in terms of cost and quality
- β_{ij} Relative ease of trade from country i to market j
- γ_{jk} Total demand

Potential trade is defined as what trade would be in this theoretical world. Potential trade can depart significantly from actual trade for a number of reasons: unrealized opportunities, non-tariff measures, etc.

Model

Demand

Assumption 1: ***differentiation by supplier***

Products exported by different countries are perceived as different by consumers and companies purchasing them. For a given product, demand behaviour results from the optimization of a constant elasticity of substitution (CES) function (Armington assumption).

CES functions are supposed to be the same in all markets. This assumption is necessary to derive a country's capacity to export to a new market from its observed performance in other markets.

$$\frac{v_{ijk}}{v_{jk}} = \alpha_{ik}\left(\frac{p_{jk}}{p_{ijk}}\right)^{\sigma_k - 1}$$

With $\sum_i \alpha_{ik} = 1$ (this assumption sets the price-volume split of the "composite good" but does not imply any additional restriction on demand behaviour).

Assumption 2: ***demand elasticity***

While v_{jk} can be influenced by income growth in market j, it is assumed not to depend on prices (the elasticity of aggregate import to aggregate price is assumed to equal 1).

Supply

<u>Assumption 3</u>: ***product homogeneity***

Products exported by a given country to different markets are assumed to be the same.

<u>Assumption 4</u>: ***supply elasticity***

Supply is assumed to be perfectly elastic: export prices do not depend on exported quantities.

<u>Assumption 5</u>: ***trade costs***

p_{ijk} can be decomposed into a factor reflecting the production cost in country i and a bilateral factor,

$$p_{ijk}{}^{1-\sigma_k} = \varepsilon_{ik}\theta_{ij}$$

For simplicity, the volume of product k exported by country i is defined so that $\varepsilon_{ik} = 1$. This does not imply any additional restriction. Any difference in cost or quality between two suppliers is reflected in the α_{ik} Armington parameter. Product-specific trade costs and barriers are introduced later.

Calculation

The optimization of the CES function leads to:

$$p_{jk}{}^{1-\sigma_k} = \sum_i \left(\alpha_{ik} p_{ijk}{}^{1-\sigma_k}\right) = \sum_i \left(\alpha_{ik}\theta_{ij}\right)$$

The price of composite good k in market j is higher if the trade costs of the most competitive suppliers (characterized by larger α_{ik}) to that market are large (i.e., θ_{ij} is small; for instance as a result of physical or cultural distance between i and j).

It follows:

$$v_{ijk} = \frac{\alpha_{ik}\theta_{ij}}{\sum_i \left(\alpha_{ik}\theta_{ij}\right)} v_{jk} \tag{A1}$$

Interpretation of the denominator: for a given demand v_{jk}, Armington parameter α_{ik} and the bilateral factor θ_{ij}, the potential value v_{ijk} is larger if the most competitive suppliers (i.e. the ones with large α_{ik}) are "far" from market j (i.e. low θ_{ij})

Approximation 1:

$p_{jk}{}^{1-\sigma_k} = \sum_i \left(\alpha_{ik}\theta_{ij}\right)$ is assumed to depend only on the market: if a market is "far" from competitive suppliers, it is assumed as "far" for all products.

$$\sum_i \left(\alpha_{ik}\theta_{ij}\right) = \vartheta_j$$

We define

$$\beta_{ij} = \frac{\theta_{ij}}{\vartheta_j}$$

Equation (A1) becomes

$$v_{ijk} = \alpha_{ik}\beta_{ij}v_{jk} \tag{A2}$$

Estimation

Strategy

Econometric estimation

One approach would be to estimate α_{ik} and β_{ij} econometrically. The size of such model, especially if zero flows are treated correctly, would be very large. It would also require the use of three-dimensional data whose quality is low because trade partners are not always identified accurately.

Calibration

Another approach is to assume that actual trade patterns at a more aggregate level (i.e., two dimensions instead of three) are accurately reported and allow the determination of the parameters α_{ik} and β_{ij}.

Implementation

Supply performance

From equation (2), it follows:

$$\frac{v_{ik}}{v_k} = \alpha_{ik} \frac{\sum_j (\beta_{ij} v_{jk})}{\sum_i [\alpha_{ik} \sum_j (\beta_{ij} v_{jk})]}$$

Because ease of trade to market j may be correlated with demand by this market, $\sum_j (\beta_{ij} v_{jk})$ in principle depends on i. To simplify the analysis however, we assume that it does not.

Approximation 2:

$$\frac{v_{ik}}{v_k} = \alpha_{ik}$$

Trade costs

From equation (A2), we obtain

$$v_{ij} = \beta_{ij} \sum_k (\alpha_{ik} v_{jk})$$

$$\beta_{ij} = \frac{v_{ij}}{\sum_k (\alpha_{ik} v_{jk})}$$

The denominator reflects the complementarity between country i's supply capacity and market j's demand for product k. β_{ij} reflects the ease of trade to a particular market, relative to the ease of trade of other countries to supply this market.

Normalization

From calculations above, it follows:

$$\widetilde{v_{ijk}} = \widetilde{\alpha_{ik}} \widetilde{\beta_{ij}} v_{jk} = \frac{v_{ik}}{v_k} \frac{v_{ij}}{\sum_k \left(\frac{v_{ik}}{v_k} v_{jk} \right)} v_{jk}$$

$\widetilde{\alpha_{ik}}\widetilde{\beta_{ij}}$ is the expected market share of country i in market j for product k. As a result of approximations 1 and 2, the sum of expected market shares in a given market generally differs from 1.

$$\sum_i \left(\widetilde{\alpha_{ik}}\widetilde{\beta_{ij}}\right) \neq 1$$

To solve this problem, market shares are normalized.

$$\widetilde{v_{ijk}} = \frac{\widetilde{\alpha_{ik}}\widetilde{\beta_{ij}}}{\sum_i\left(\widetilde{\alpha_{ik}}\widetilde{\beta_{ij}}\right)} v_{jk} = \frac{v_{ik}}{v_k} \frac{v_{ij}}{\sum_k\left(\frac{v_{ik}}{v_k} v_{jk}\right)} \frac{1}{\sum_i\left[\frac{v_{ik}}{v_k}\frac{v_{ij}}{\sum_k\left(\frac{v_{ik}}{v_k} v_{jk}\right)}\right]} v_{jk}$$

Market access

Beyond the overall ease of trade to market j (as a result of distance, etc.), access of country i to market j is also product-specific:

- Tariff preferences are product specific;
- Impact of distance depends on the product.

To solve this issue, a correction factor is added. The equation becomes

$$v_{ijk} = \alpha_{ik}\beta_{ij}\delta_{ijk}v_{jk}$$

Approximation 2 is modified. It now states that $\sum_j\left(\beta_{ij}\delta_{ijk}v_{jk}\right)$ does not depend on i. It follows:

$$\widetilde{\alpha_{ik}} = \frac{v_{ik}}{v_k}$$

$$\widetilde{\beta_{ij}} = \frac{v_{ij}}{\sum_k\left(\alpha_{ik}\delta_{ijk}v_{jk}\right)}$$

$$\widetilde{v_{ijk}} = \frac{\widetilde{\alpha_{ik}}\widetilde{\beta_{ij}}\delta_{ijk}}{\sum_i\left(\widetilde{\alpha_{ik}}\widetilde{\beta_{ij}}\delta_{ijk}\right)} v_{jk} = \frac{v_{ik}}{v_k} \frac{v_{ij}}{\sum_k\left(\frac{v_{ik}}{v_k}\delta_{ijk}v_{jk}\right)} \frac{1}{\sum_i\left[\frac{v_{ik}}{v_k}\frac{v_{ij}}{\sum_k\left(\frac{v_{ik}}{v_k}\delta_{ijk}v_{jk}\right)}\delta_{ijk}\right]} \delta_{ijk}v_{jk}$$

In practice, in order to better interpret the meaning of the δ_{ijk} parameter, it is split into two components: the first one enters into the supply performance factor, while the other one is part of the demand indicator. This split does not have an impact on export potential values.

Annex II: Data sources

Variable	Source	Link
Export and import values	ITC Trade Map	www.trademap.org
Ad-valorem tariffs	ITC Market Access Map	www.macmap.org
Price elasticities	GTAP (Hertel et al., 2004)	https://www.nber.org/system/files/working_papers/w10477/w10477.pdf
Distances	CEPII GeoDist (Mayer and Zignago, 2011)	www.cepii.fr/CEPII/fr/bdd_modele/presentation.asp?id=6
GDP growth projections	World Economic Outlook (IMF)	https://www.imf.org/en/Publications/SPROLLs/world-economic-outlook-databases#sort=%40imfdate%20descending
Population	UN Department of Economic and Social Affairs, population division	https://population.un.org/wpp/Download/Standard/Population/
Land endowment	GTAP "Land Use" database, version 10 (2020)	https://www.gtap.agecon.purdue.edu/resources/download/11922.pdf

Annex III: Estimation of revenue elasticities

Demand is projected based on expected GDP growth, expected population growth, expected tariff changes, and revenue elasticities. Revenue elasticities measure the percentage change in imports that is caused by a 1% change in real GDP. These elasticities are calculated for each of the 96 HS chapters via simple OLS econometric regressions, which are run separately for two groups of countries: developed and developing countries. These regressions are of the form:

$$\dot{m}_{c,d,t} = \alpha_{c,d}\dot{y}_{d,t} + \beta_{c,d} + \varepsilon_{c,d,t}$$

Where

- c is the HS chapter
- d is the country group: developing or developed
- t is the time index
- $\dot{m}$ is the average annual growth rate of imports per capita m_t:
$$\dot{m}_t = \frac{1}{n}\ln\left(\frac{m_{t+n}}{m_t}\right)$$
- $\dot{y}$ is the average annual growth rate of real GDP per capita y_t:
$$\dot{y}_t = \frac{1}{n}\ln\left(\frac{y_{t+n}}{y_t}\right)$$
- $n \in \{1, \dots, 7\}$ is the number of years over which annual growth rates are computed
- $\alpha_{c,d}$ is the income elasticity
- $\beta_{c,d}$ is a trend, that includes inflation in US dollars: GDP is expressed in real terms, but imports are nominal
- $\varepsilon_{c,d,t}$ is the error term

We use data between 2001 and 2021 to estimate medium-to-long term income elasticities that match the time horizon of our export potential projections. More weight is given to more recent years. Other specifications than the one above were tested and led to less reliable estimators. These alternative specifications include using logs for the dependent and independent variables instead of differences in logs, a log-linear specification, a quadratic specification, and lags of GDP for the regressions. Results are available upon request.

Some import growth rates are quite large for small countries. With equal weights, these fluctuations would be given as much importance in the regression as more stable growth rates. To tackle this issue, we would ideally like to use weights proportionate to the inverse of the variance for each observation, to obtain the best linear unbiased estimates (BLUE) estimates. We do not have access to the true variance of each individual observation, but we can use the observed variance of the dependent variable across time, for every chapter and development level, as an estimate of this variance.

We tried three weighting schemes:

1) Equal weights
2) Weights equal to the inverse of the observed variance of the dependent variable across time:
$$W_{c,d,t} = \frac{1}{var_{c,d}(\dot{m}_{c,d,t})}$$
3) Linearly increasing weights, based on the formula
$$W_{c,d,t} = \frac{t}{var_{c,d}(\dot{m}_{c,d,t})}$$
With $t \in \{1, \dots, 18 - n\}$

For every chapter and development level, the method provides estimates for the income elasticity and the trend, along with an estimate of the standard error of these two parameters.

We define the following score for every chapter and development level:

$$\text{Score}_{c,d} = \sigma_{\beta_{c,d}} + \delta_y \sigma_{\alpha_{c,d}}$$

Where:

- $\sigma_{\beta_{c,d}}$ is the standard error of $\beta_{c,d}$
- $\sigma_{\alpha_{c,d}}$ is the standard error of $\alpha_{c,d}$
- δ_y is the standard deviation of the average annual growth rate of GDP per capita, for all countries within a development group

To compare the different strategies, we consider two criteria: the average score of a strategy, and the maximum score for a chapter and development level (where a small score is better).[39]

The best strategy overall is the strategy with increasing weights and 6-year growth rates (strategy 6-3).

Table 5: Average elasticities

Country group	Average elasticity
Developing countries	1.201
Developed countries	0.928

As table 5 shows, our estimates of import elasticities average at 1.201 for developing countries and 0.928 for developed countries.

Figure 1: Import elasticities by HS chapter

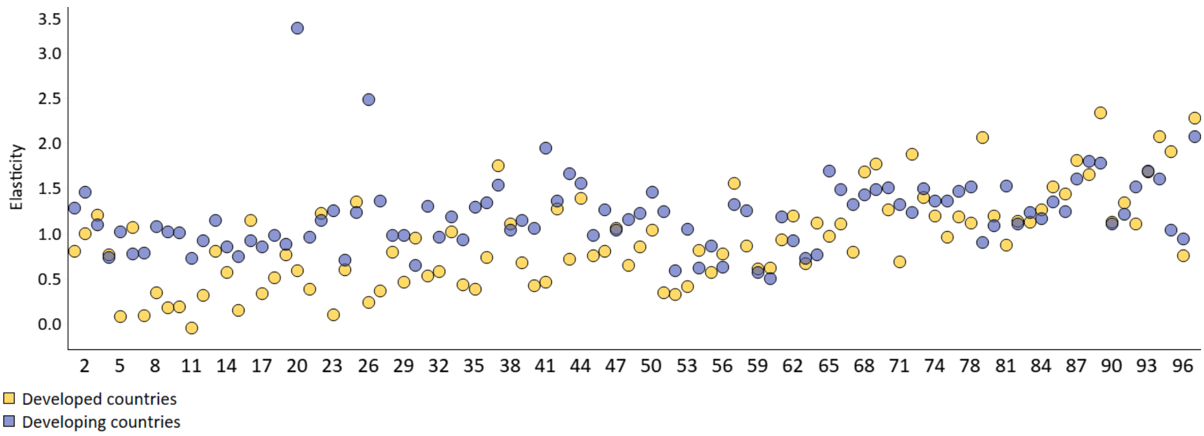

Figure 1 shows the value of the import elasticities by chapter. There are significant differences between country groups for most of the calculated elasticities, yet the estimated coefficients are generally close to one for developing countries, and below one for developed countries. In general, manufactured products (HS Chapters 28 to 97), especially in developed countries, have higher elasticities than other types of products.

[39] The selection of the best estimation strategy could be done for every chapter and development level. However, different estimation strategies correspond to slightly different concepts. Therefore, we decided to select the strategy that performed best overall.

Annex IV: Land endowments

The product space approach relies on linkages between products inferred from trade data. But these linkages could result from a statistical coincidence without being truly related to production capacities (e.g., the mere fact that products l and k are jointly exported by many countries does not necessarily mean that any country currently exporting l could also diversify into k). In order to refine the list of suggested products to a given country, information on actual land endowments from a GTAP dataset that distinguishes 18 agro-ecological zones (AEZs) is used. AEZs are available for 121 individual countries and for all other countries at the level of their respective region. A country or region is considered as producing a product if it has exported it with comparative advantage and a strictly positive trade balance in the three most recent years of the analysis.

For each agricultural or forestry product (product groups 060110 to 140490, 1701, 1801, 1802, 4001, 4401 to 4403, 4407, 4409, 4501, 5201 to 5203, 5301 to 5303, 5305), an iterative process identifies all land types suitable for its production. We first compute an estimation of the exports by product, region and land type:

$$Est\ v_{rak} = v_{rk} \times \frac{R_{ra}}{\sum_a R_{ra}}$$

v_{rk} are the total exports of product k from region r. R_{ra} is the rent associated with region r and land type (AEZ) a. Then, we rank the land types in descending order of $Est\ v_{rak}$ by product. Starting with the region with the highest $Est\ v_{rak}$, each region is added to the list of suitable land types until all exporters in the top 95% of world exports of a given product are included. Finally, for a given region, we only keep products for which at least one of its associated land types is in the top 95% in terms of rent.

With this method, between one and five AEZs are considered as relevant for the production of a particular good. If the product space approach identifies strong links from the country's current export basket to one of these products but the country does not possess any of the relevant AEZs, the product is removed from the list of feasible diversification opportunities for this country.

Annex V: Excluded products

HS code	Description	Main Reason
24	Tobacco & manufactured tobacco substitutes	ITC (Risk for human beings)
252410, 252490, 252530	Crocidolite asbestos; asbestos; mica waste	ITC (Extractive industries)
25* (except above-mentioned products)	Salt; sulphur; earths & stone; plastering materials	ITC (Extractive industries)
262019, 262021, 262029, 262030, 262040, 262060, 262091, 262099, 262110, 262190	Slag, ash & residues; leaded gasoline & anti-knock compound sludges	ITC (Risk for human beings)
26* (except above-mentioned products)	Ores	ITC (Extractive industries)
27 (except 2712 & 2716)	Mineral fuels, mineral oils & products of their distillation; bituminous substances (except petroleum jelly and electrical energy)	ITC (Extractive industries)
28*	Inorganic chemicals; organic or inorganic compounds of precious metals, of rare-earth metals, of radioactive elements or of isotopes	ITC (natural resource based)
2903 (most products)	Halogenated derivatives of hydrocarbons	WCO (Pollutants)
291010	Oxirane "ethylene oxide"	WCO (Pollutants)
300692	Waste pharmaceuticals	WCO (Waste)
3825	Residual products of the chemical or allied industries, n.e.s; municipal waste; sewage sludge; clinical waste	WCO (Waste)
3915	Waste, parings & scrap, of plastics	WCO (Waste)
4004	Waste, parings & scrap of soft rubber & powder & granule obtained therefrom	WCO (Waste)
401220	Used pneumatic tyres of rubber	WCO (Waste)
411520	Parings & other waste of leather or of composition leather, not suitable for the manufacture of leather articles; leather dust, powder & flour	WCO (Waste)
450190	Cork waste; crushed, powdered or ground cork	WCO (Waste)
4707	Recovered "waste & scrap" paper or paperboard (excluding paper wool)	WCO (Waste)
4813	Cigarette paper	ITC (Risk for human beings)
5003	Silk waste, incl. cocoons unsuitable for reeling, yarn waste & garnetted stock	WCO (Waste)

* Only excluded from the product diversification indicator.

510320 & 510330	Waste of wool or of fine or coarse animal hair, incl. yarn waste (excluding noils & garnetted stock)	WCO (Waste)
520210 & 520299	Cotton waste, including yarn and thread waste but excluding garnetted stock	WCO (Waste)
530130	Flax tow and waste	WCO (Waste)
5505	Waste of man-made fibres, incl. noils, yarn waste & garnetted stock	WCO (Waste)
6309	Worn clothing, footwear & textile articles (excluding floor coverings & tapestries)	WCO (Waste)
6310	Used or new rags, scrap twine, cordage, rope and cables and worn-out articles thereof, of textile materials	WCO (Waste)
710820, 711230, 711291, 711292, 711299, 711810, 711890	Monetary gold; coins; waste and scrap of gold, silver & platinum	ITC (Extractive industries)
71 (except 710221, 710229, 710410, 710420, 710490, 711311, 711319, 711320, 711411, 711419, 711420, 711510, 711590, 711610, 711620, 711711, 711719, 711790)*	Precious stones and semi-precious stones, unworked; precious metals, unwrought or in powder form; pearls; semi-manufactured precious metals	ITC (Extractive industries)
7404	Copper waste & scrap	WCO (Waste)
7802	Lead waste & scrap	WCO (Waste)
810420, 810730, 811020, 811213, & 811252	Waste and scrap of magnesium, cadmium, antimony, beryllium, chromium, and thallium	WCO (Waste)
8478	Machinery for preparing or making up tobacco	ITC (Risk for human beings)
854810	Waste & scrap of primary cells, primary batteries & electric accumulators	WCO (Waste)
8710	Tanks & other armoured fighting vehicles & parts, n.e.s.	ITC (Risk for human beings)
880521	Air combat simulators and parts thereof	ITC (Risk for human beings)

* Only excluded from the product diversification indicator.

890610	Warships of all kinds	ITC (Risk for human beings)
8908	Vessels & other floating structures for breaking up	WCO (Waste)
93	Arms & ammunition, parts & accessories thereof	ITC (Risk for human beings)
9614	Smoking pipes, incl. pipe bowls, cigar or cigarette holders, & parts thereof, n.e.s.	ITC (Risk for human beings)
9704	Used postage/revenue stamps	ITC (no relevance for trade promotion)
9706	Antiques of > 100 years old	ITC (no relevance for trade promotion)

Annex VI: Report reliability

While direct reports of import data include the cost of insurance and freight (CIF), mirror reports are free on board (FOB) and do not include these costs. This explains part of the deviation between direct and mirror data. Another part is explained by misallocations or measurement errors. Assuming that the CIF/FOB ratio is small, the reliability of reported trade can be detected as follows:

For each country, trade direction and year, a reliability score is calculated based on the absolute difference between direct and mirror reports of bilateral trade of each product.[40] The initial score is based on the weighted sum of absolute values of differences, $\delta_{ijt} = \sum_k \left|X_{ijkt} - M_{ijkt}\right|$, divided by the weighted sum of reports by both sides, $\sigma_{ijt} = \sum_k \left(X_{ijkt} + M_{ijkt}\right)$, which implies that larger partners are given more importance. Bilateral scores are then weighted with the square of the partner's reliability score to ensure that non-reliable partners are given little importance:

$$S_{it} = 1 - \frac{\sum_j \left[S_{jt}{}^2 \delta_{ijt}\right]}{\sum_j \left[S_{jt}{}^2 \sigma_{ijt}\right]}$$

$$S_{jt} = 1 - \frac{\sum_i \left[S_{it}{}^2 \delta_{ijt}\right]}{\sum_i \left[S_{it}{}^2 \sigma_{ijt}\right]}$$

The computation is iterated until scores converge.

Although the scores are continuous, they are transformed into a binary variable with a threshold of 0.5. A reporter is assessed as "reliable" if its final score is greater than 0.5 and as "unreliable" if its final score is less than or equal to 0.5.

If one reporter is reliable and the other reporter is not, only trade flows reported by the reliable reporter are used. If both reporters are reliable, a geometric average of the trade flows reported by both countries is used. If both reporters are unreliable, no trade data is used.[41] The trade data used to compute market shares and import demand is therefore given by

$$Estimated\ trade_{i(j)k} = \sqrt{\sum_{j(i)\in RC} Dir_{ijk} \times \sum_{j(i)\in RC} Mir_{ijk}} + \sum_{j(i)\in NRC} Dir_{ijk}$$

Where RC refers to a reliable reporting country and NRC to a non-reliable or non-reporting country.

[40] Products in the exclusion list are not considered.

[41] There is one exception: bilateral trade data X_{ijk} used to compute unrealized potential and total bilateral trade rely on arithmetic averages of direct and mirror data because partners are often misidentified, such that using a geometric average would result in a significant reduction of trade flows. Furthermore, the concern of presenting false positives does not apply for these values. As bilateral trade data is used to rescale potential shares and compute potential values, the overall magnitude of trade potential is not affected by the choice of relying on geometric averages to compute potential shares.

Annex VII: Detailed data treatments

Trade data

Export potential estimates aim to predict a baseline level of trade in the next five years. To make sure that a one-year fluctuation of trade does not completely distort the indicator, average across five years are taken. As we wish to reflect a recent trade situation, we place higher weights on more recent years: the most recent year is given a weight of 5, the second most recent year a weight of 4, etc until the oldest year, given a weight of 1.

Tariff data

To make sure that prohibitively large tariffs—often part of tariff-rate quotas—do not completely distort our results, we get rid of all tariffs which have an ad-valorem equivalent higher than 80%.