UNITED NATIONS CONFERENCE ON TRADE AND DEVELOPMENT

UNCTAD

Do Differences in the Types of Commodities Exported Matter for Export Concentration?

UNITED NATIONS

© 2019, United Nations

This work is available open access by complying with the Creative Commons licence created for intergovernmental organizations, available at http://creativecommons.org/licenses/by/3.0/igo/.

The findings, interpretations and conclusions expressed herein are those of the authors and do not necessarily reflect the views of the United Nations or its officials or Member States.

The designation employed and the presentation of material on any map in this work do not imply the expression of any opinion whatsoever on the part of the United Nations concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

Photocopies and reproductions of excerpts are allowed with proper credits.

This publication has not been formally edited.

United Nations publication issued by the United Nations Conference on Trade and Development.

UNCTAD/DITC/COM/2019/4

eISBN: 978-92-1-004732-6

Abstract

This paper examines whether the type of commodity dominating a country's exports matters for export concentration. Using a dataset covering 173 countries, including 87 commodity-dependent developing countries, we estimate dynamic panel data models that control for a large set of determinants of export concentration. We find that GDP shares of energy exports and, to a lesser extent, GDP shares of minerals, are important determinants of export concentration. Our results imply that developing countries that are dependent on energy or minerals sectors face challenges that are different from those faced by countries dependent on the export of agriculture and manufacturing products. Economic and export diversification policies need to take these differences into account.

Acknowledgements

This document was prepared by Rodrigo Cárcamo and Janvier. D. Nkurunziza, UNCTAD Commodities Branch, Division on International Trade and Commodities.

The authors would like to thank the following people who provided very valuable comments on earlier versions of this paper: Fernando Cantú, Ramón Pineda, Justin Van de Ven and Yanchun Zhang. The authors also thank Catherine Katongola for administrative assistance and Danièle Boglio for desktop publishing.

The views expressed in this paper are those of the authors and do not necessarily represent the views of UNCTAD. Any remaining errors are exclusively the responsibility of the authors. This document has not been edited.

For further information about this publication, please contact: Commodities Branch, Division on International Trade and Commodities, UNCTAD, Palais des Nations, CH-1211 Geneva 10, Switzerland, tel. +41 22 917 2408, e-mail: commodities@unctad.org.

Material in this publication may be freely quoted or reprinted, but acknowledgement is requested, together with a copy of the publication containing the quotation or reprint, to be sent to the UNCTAD secretariat.

Table of contents

1. Introduction 1

2. Measures of Export Concentration 5

3. The Determinants of Export Concentration 9

   Abundance of factors of production 9
   The level of development 11
   Institutional quality 12
   Export prices 13
   Trade barriers and trade costs 13
   The size of the economy 14
   Foreign direct investment 14

4. Empirical Analysis 15

   Data Description 15
   Empirical Results 18

5. Conclusions 23

6. Bibliography 25

Appendix 29

   Countries in the data sample 29

List of acronyms

LDC	least developed country
LLDC	landlocked developing country
CDDC	commodity-dependent developing country
GDP	gross domestic product
UNCTAD	United Nations Conference on Trade and Development
SIDS	small island developing state

1. Introduction

Many developing economies are characterized by a highly concentrated export sector. In these countries, exports are often limited to a small number of products, and in several cases[1] one primary commodity accounts for more than half of a country's total export earnings (UNCTAD, 2019). In the developing world, export commodity dependence is pervasive. Almost two-thirds of developing countries are commodity-dependent, meaning that at least 60 per cent of their merchandise export revenues come from commodity exports[2]. Poor countries are particularly commodity dependent. Indeed, 85 per cent of Least Developed Countries (LDCs) and 81 per cent of Land-Locked Developing Countries (LLDCs) are commodity-dependent (UNCTAD, 2019). The total number of commodity-dependent countries increased slightly between 1998 and 2017, reaching 102 countries.

Commodity dependence can negatively affect development through different but related channels.[3] Commodity-dependent developing countries are vulnerable to different types of negative shocks. These include terms of trade shocks, and shocks from technological changes that reduce demand for some export commodities. Moreover, high volatility of commodity export prices introduces uncertainty in the economy. These shocks negatively affect the quantity and quality of public and private investment and total factor productivity, ultimately slowing down economic growth. Koren and Tenreyro (2007), Blattman, Hwang, and Williamson (2007), and Van der Ploeg and Poelhekke (2009) link the low and volatile economic growth performance[4] of countries with concentrated commodity exports to commodity price volatility. Additionally, given the importance of good institutions for development (Acemoglu et al., 2002, 2003; Rodrik et al., 2004), poor institutions are another conduit through which commodity dependence negatively affects development (Frankel, 2010; van der Ploeg, 2011).

Strong concentration of developing countries' exports on commodities and the negative link between commodity dependence and economic development have motivated research on the conceptual and empirical determinants of export concentration.[5] While interest in the analysis of export concentration goes back several decades (e.g. Michaely, 1958), only relatively recently have studies attempted to empirically identify the determinants

[1] UNCTAD (2019) shows that among the 50 countries with the highest measure of commodity dependence in the 2013-2017 period, 26 exported one product line (at 3 digits of the SITC classification) that represented more than 51 per cent of the country's total exports (See Table A1, page 43).

[2] Commodity dependence can then be defined as "high" export concentration on commodities. In this paper, we consider "high" to represent situations where more than 60 per cent of total merchandise export revenues come from commodity exports (See UNCTAD (2019) Section 2 for a discussion of data issues in calculating commodity dependence).

[3] The "natural resource curse" literature (Sachs and Warner 2001, Frankel, 2010, Van der Ploeg, 2011) identifies a negative relationship between commodity exports and economic growth. Significant work has been carried out attempting to identify the channels through which such a relationship is mediated.

[4] For a discussion of why macroeconomic volatility affects long-term growth and the level of development, see Aghion et al. (2010) and Barlevy (2007).

[5] The correlation between commodity dependence and the level of development (measured by GDP per capita) has been well established empirically. For a recent example, see UNCTAD (2019), which uses different static panel data models to show the (possibly non-linear) strong relationship between export concentration and income per capita. All the countries below the median income per capita and over the median export concentration index are commodity-dependent developing countries.

of export concentration. Such studies include Bebczuk and Berrettoni (2006), Klinger and Lederman (2009), De Benedictis, Gallegati and Tamberi (2009), Cadot, Carrère, and Strauss-Kahn (2011a,b), Agosín, Alvarez and Bravo-Ortega (2012), Parteka and Tamberi (2013) and Bahar and Santos (2018), among others. Cadot, Carrère, and Strauss-Kahn (2013) provides a recent survey of this literature. Using different datasets, measures of export concentration, and model specifications, these studies found that determinants of export concentration included: the level of development of a country (measured by GDP per capita), the size of the economy (measured by population or GDP size), trade barriers and costs, the terms of trade, export prices, composition of production factors including natural resource endowments, indicators of human and physical capital stock, and institutional quality.

Our paper tests whether the composition of exports matters for export concentration. Specifically, the paper examines whether the types of commodity exports a country depends on could have different impacts on export concentration. Using trade data disaggregated at the three digits of the Standard International Trade Classification (SITC), we classify all exported products into four groups, the first three being commodity groups: *Agricultural Products* (aggregating Food and Agricultural Raw Materials); *Energy*; *Minerals*; and *Non-commodities*. Using dynamic panel data models with data covering 173 countries and the period between 1995 and 2017, we find that countries with large energy, and to a lesser extent, minerals exports have more concentrated exports, even after controlling for many other determinants of concentration. The GDP share of agricultural exports does not seem to be a significant determinant of concentration. This result is robust to the use of different measures of export concentration, namely the Hirschman-Herfindahl Index, the Gini coefficient, and Theil's T index.[6]

This paper contributes to the literature in different ways. First, by decomposing commodity exports into different groups, our analysis goes further than previous findings in explaining how export composition affects concentration. Bebczuk and Berrettoni (2006) for example found that export concentration was positively correlated with the share of fuel exports in total exports, and negatively correlated with the share of manufactures on exports. Cadot et al. (2011a) found that the aggregated share of minerals and energy exports in total exports is an important determinant of export concentration, while controlling for GDP per capita in levels and squares. Unlike previous studies that used the ratios (aggregate or individual groups) of commodity exports over total exports, we use the value of exports of the three individual groups of commodities as a percentage of GDP as covariates. This attempts to address the problem arising from the fact that export shares of total exports are also used in the calculation of measures of concentration like the Hirschman-Herfindahl and Theil's T indexes of concentration. Additionally, our measures may capture better the effects of commodity dependence than trying to include directly difficult-to-quantify determinants like exchange rate overvaluation and exchange rate volatility,[7] while also accounting for the imperfect measurement of institutional quality across countries and time. In this regard, our results complement and extend those of Bahar and Santos (2018), who found that aggregate natural resources exports as a share of total exports are important determinants of the concentration of non-commodity exports.

Secondly, our study uses one of the largest datasets employed so far to study the determinants of export diversification, both in terms of country and time coverage. The inclusion of a large number of developing countries in the sample improves over past studies that were dominated by developed countries, potentially leading to misleading results given that the issue of export concentration is primarily a developing country problem. For example, Bebczuk and Berrettoni (2006) included 56 economies for the

[6] The literature on concentration generally uses these three measures; we follow the same practice.

[7] Agosín et. al. (2012) included a measure of exchange rate overvaluation and of nominal monthly exchange rate volatility, among other variables; they turn out statistically non-significant determinants of concentration. Non-significance may have been due to the challenges of quantifying those variables.

period 1970-2002 into their study, 27 and 4 of which are today classified as developed and high-income developing countries, respectively. Parteka and Tamberi (2013) used a dataset covering manufacturing exports from 60 countries over a 20-year period (1985-2004). Moreover, the time period covered by our study, from 1995 to 2017, is important, given that it represents episodes of sustained low and high commodity prices. It is more representative of the commodity price cycles. Between 1995 and 2003, prices were low, followed by a commodity boom from 2003 to 2011 and then another period of falling prices between 2011 and 2017. Including data for both the commodity price "boom" and "bust" captures the long-term behaviour of commodity prices and their effect on export concentration. Studies like Cadot et al. (2011b) used a dataset that covered a period that only partially includes significant increases in commodity prices. 8 The data used in other studies like Agosín et. al. (2012) does not contain recent structural changes in the patterns of trade such as the industrialization of China and its augmented role in global demand for commodities.[9]

Thirdly, in addition to using the Herfindahl-Hirschman Index (HHI), Gini coefficient, and Theil's T Index as measures of export concentration, we propose an "adjusted" version of Theil's T concentration index that fixes the number of exported product lines at the maximum for all countries, rather than varying it over time according to registered data. This indicator addresses the sensitivity of Theil's T index to the number of lines exported (which for many developing countries show very small values for most products; in some cases, these could even be re-exports). The adjusted Theil's T index accounts for the "jumps" in the number of open product lines in existing trade data, which can potentially be attributed to data collection issues in several developing countries.

Methodologically, our study builds on past work. It accounts for data persistence by applying a dynamic Generalized Method of Moments estimation (Blundell and Bond, 1998), as proposed by Agosín et al. (2012), but using a different and more comprehensive set of determinants that reflect the results of past empirical studies.

The result highlighting the importance of energy and, to a lesser extent, mining but not agriculture commodity exports for a country's export concentration may mean that energy and mining exporting countries are more vulnerable than others to exchange rate overvaluation and Dutch Disease, and to the effects of exchange rate volatility on the quality and the quantity of investment. These effects are difficult to capture directly and are not fully accounted for by the variables included in our model. In terms of policy, the finding suggests that energy and to a lower extent mineral export-dependent countries might face more daunting challenges than countries relying on other sectors as they try to diversify their economies and exports.

The paper is organized as follows. Section 2 discusses different measures of export concentration. Section 3 reviews and discusses the literature on the determinants of export concentration. Section 4 presents the data used in the empirical analysis of the determinants of export concentration and discusses the empirical results from econometric estimations. Section 6 concludes.

[8] They used data from for 87 countries for the period 1990-2004, where only the last two years of the period start to show important commodity price increases.

[9] Agosín et. al (2012) used the trade dataset from Feenstra et. al (2005) in a dynamic panel estimation of concentration determinants, with the HHI, Theil's T index and the Gini coefficient as measures of concentration.

2. Measures of Export Concentration

Export concentration attempts to quantify the degree of dispersion or of lack of homogeneity in a country's exports across products, partners, or both.

Past studies (see Cadot et al., 2013) measured a country's degree of export concentration in different ways. We focus on three measures, namely[10] the normalized Herfindahl-Hirschman Index (HHI), the Gini coefficient and Theil's T Index.

The Herfindahl-Hirschman Index for the exports of country *c* for a given time period is

$$HHI_c = \sum_{k=1}^{N} (S_k)^2$$

where S_k is the export share of product *k* in total exports, calculated as

$$S_k = x_k / \sum_{k=1}^{n} x_k$$

where x_k is the value of product *k* exports, and *N* is the total number of products exported by country *c*.

The normalised version of the HHI, taking values between 0 and 1, is:

$$Normalised \quad HHI_c = \frac{\sum_{k=1}^{n} (S_k)^2 - \frac{1}{N}}{1 - \frac{1}{N}}$$

The HHI adds up the square of the share of each product line in total exports, and in its normalised version, values close to 0 mean low concentration and values close to 1 mean high concentration. As econometric estimation may be problematic when the values of the HHI approach the limit values of 0 and 1 (see Fox, 2016), empirical studies like Agosín et al. (2012) apply a logit transformation to the value of the Normalized HHI, as follows:

$$Logit\ Norm\ HHI_c = log_e \frac{Normalised\ HHI_c}{1 - Normalised\ HHI_c}$$

This removes the upper and lower bounds of the scale, and spread out the tails of the distribution, making the variable symmetric about 0.

The second widely used synthetic concentration measure is Theil's T Index (*TEI*), which can be calculated as

$$TEI = \frac{1}{N} \sum_{k=1}^{N} \frac{x_k}{\mu} ln(\frac{x_k}{\mu}) \quad where \quad \mu = \frac{\sum_{k=1}^{N} x_k}{N}$$

and μ is the average value of exports across all trade lines.

10 Other concentration measures included in past studies on the determinants of concentration such as Cadot et. al (2011a) include the number of active product lines being exported by a country. Palan (2010) discusses in detail the characteristics of different concentration measures.

Theil's T index takes values between 0 (minimum concentration) and *ln(N)* (maximum concentration).

One practical challenge that emerges from the use of Theil's T index concerns the number of trade lines that have zero trade or that are miss-registered and result in a country having very large, implausible changes in the number of exported product lines from one year to the next.[11] This results in the value of the Theil's T index changing abruptly between successive years, as in particular the value of the mean export line μ changes substantially, while also the bounds of the index change from year to year and are different for different countries. The sharp change in the number of active product lines from year to year may be due to measurement errors, especially given the very large jumps from one year to the next in countries like Angola. It may also be due to very high mortality of product lines in developing countries, especially non-commodity product lines.

For this reason, we also calculated for each country and year an adjusted version of the Theil's T index, as follows:

$$TEI_{adj} = \frac{1}{N}\sum_{k=1}^{n}\frac{x_k}{\mu_a} ln(\frac{x_k}{\mu_a}) \quad where \quad \mu_a = \frac{\sum_{k=1}^{n} x_k}{N_T}$$

and μ_a is the average value of exports if the country exported the maximum possible number of product lines (N_T). The maximum possible product lines at 3 digits for the SITC rev. 3 classification in COMTRADE is 261.[12] As it is not possible to know, a priori, whether observed yearly changes in the number of product lines are due to measurement errors, high yearly attrition of exported product lines, or both, the use of both measures of the Theils' T index may increase robustness of empirical results of export concentration. Our data shows that for countries with a diversified trade structure like the United States or Italy, Theil's T index in both its original and adjusted versions are very similar, but for countries with very few export lines and high annual variability in the registered number of lines, as is the case of several commodity-dependent developing countries, the differences are important.

Note that unlike the normalized HHI, Theil's T Index measures export concentration using a weighted average of the ratios of individual export lines to the mean value of total exports, with the mean calculated either using the number of registered product lines (original Theil's T index), or the maximum possible number of product lines (adjusted Theil's T index). Theil's T index is not constrained to values between 0 and 1.

The third widely used measure of export concentration is the Gini Coefficient. To calculate it, we follow Cowell (2011), adjusting for a bias correction:

$$Gini = (\frac{2}{N^2\mu}\sum_{k=1}^{N}(kS_k) - \frac{N+1}{N}) \times \frac{N}{N-1}$$

where N is the maximum number of observed product lines for any country.[13] The bias-corrected Gini coefficient takes values between 0 and 1 (maximum concentration).

The high annual variability in the number of active export product lines, as well as the

[11] For example, Angola registers 198 product lines exported in 2006, 6 lines in 2007, 151 product lines in 2008, and 6 product lines exported in 2009.

[12] The number of 261 lines was obtained from the detailed data at https://unstats.un.org/unsd/tradekb/Knowledgebase/50096/Commodity-Indexes-for-the-Standard-International-Trade-Classification-Revision-3.

[13] The maximum number of observed product lines exported by any country in our data sample is 255 lines at 3 digits of the SITC rev. 3 classification. Note that this is close but not equal to the maximum possible number of product lines at this level of disaggregation, namely 261 lines, as mentioned before.

incidence of re-exports in export data in several developing countries cast doubts on the relevance of using the number of active export product lines as a measure of export diversification by some studies (see Cadot et al., 2011a,b; Bahar and Santos, 2018). Therefore, that measure was not included in this study.

In summary, the Herfindahl-Hirschman Index, Theil's T Index and the Gini coefficient are different but complementary standard measures of export concentration that have been used in different studies on the determinants of export concentration (e.g. Agosín et al., 2012, Bahar and Santos, 2018). A detailed technical discussion of the advantages, disadvantages and characteristics of these and other concentration indices is available in Palan (2010).

3. The Determinants of Export Concentration

As pointed out in the introduction, several studies have attempted to identify what factors determine export concentration. While several determinants have been proposed, seven of them are particularly important.

Abundance of factors of production

Trade theory suggests that relative abundance of factors of production combined with heterogenous factor mobility across productive sectors due to technological reasons, can influence export concentration. This implies that countries with natural resource abundance would have higher export concentration and a high component of commodities in their exports, while abundance of human capital would suggest the opposite. Indeed, natural resources are not very mobile across productive sectors whereas human capital is very mobile. Bebczuk and Berrettoni (2006) found the share of fuel exports over total exports to be positively associated with export concentration. Cadot et. al. (2011a) found that the export share of minerals plus energy exports is positively and significantly associated with export concentration, controlling for GDP per capita in levels and squares. Bahar and Santos (2018) find that the share of natural resource exports in total exports is positively correlated with different forms of export concentration of non-commodity exports, controlling for GDP per capita and the size of total exports.

Agosín et al. (2012) found that a measure of human capital, the Barro and Lee (2001) measure of average years of schooling, was negatively associated with the Gini coefficient,the HHI, and Theil's T index, all measuring export concentration, but the relationship was not significant for the latter. Cadot et. al. (2011b) also found years of schooling and Theil's T index to be negatively correlated.

Past empirical studies have generally not explored the differentiated effect on concentration of economies dependent on the export of agricultural, mineral and energy resource products. An important conceptual reason for doing so is that certain types of natural resources can be used in different production processes making them more flexible as inputs than others. For example, while human capital is highly mobile in the production process across sectors of the economy petroleum, gas or minerals have more specific uses. Therefore, energy and mineral-abundant countries would a priori be expected to be more export concentrated than land-abundant countries. This paper looks into this issue.

The abundance of certain types of natural resources, especially energy and minerals, may also be associated with other explanatory variables that are difficult to measure directly or which have complex relationships with export concentration. First, the covariate *export prices* not only has a direct relationship with export concentration, as we explain below, but also one that operates indirectly via incentives to invest in the sector with high relative prices. This indirect relationship may also operate with a lag on export concentration for commodity-dependent countries, as persistently high commodity prices (e.g. of petroleum) lead to increases in production and export of those commodities, and hence of export concentration. This indirect relationship between persistently high commodity prices and export concentration has been analysed extensively under the umbrella of the *Dutch Disease* phenomenon. High commodity prices result in real effective exchange rate overvaluation,[14] curtailing economic and export diversification (Corden, 1984). Bahar and Santos (2018) present a theoretical model that shows how export concentration may be affected by commodity prices, via the impact on wages in a heterogeneous-firm model. The authors propose that high commodity prices require higher productivity for firms exporting non-commodities, resulting in reduced export diversification (i.e. higher concentration). Moreover, as labour-intensive sectors are more affected by higher wages induced by changes in the terms-of-trade, higher commodity prices result in higher concentration of non-commodity exports into capital-intensive goods.

While one could include a measure of exchange rate overvaluation to control for this indirect and lagged effect of export prices, especially for commodity-dependent countries, as is the case with Agosín et al. (2012), it is empirically challenging to assess the existence and the degree of exchange rate misalignment across countries and time (Aguirre and Calderón 2005). It is also difficult to determine the correct number of lags with which exchange rate overvaluation affects export concentration.

Second, it is challenging to determine *a priori* the relevant time period to use for the measure of exchange rate volatility when assessing its impact on concentration[15], making it difficult to use the variable *exchange rate volatility* as a covariate. However, the effects on export concentration of both the Dutch disease and exchange rate volatility are particularly important for commodity-dependent developing countries.[16] Including as an explanatory variable a measure of natural resource abundance of each type of commodity exported can help to account for these issues because different types of commodities display differences in terms of price volatility and magnitude of price shocks. During the last commodity price boom, energy and mineral prices grew much more than agricultural prices (UNCTAD, 2019), so *ceteris paribus*, countries dependent on energy and minerals would have been much more at risk from Dutch Disease and relative price distortions in general than those exporting agricultural or non-commodity products. Similarly, energy and mineral prices (especially, precious metals) have been shown to be much more volatile than agricultural prices (UNCTAD, 2018). This justifies the relevance of distinguishing between different types of natural resources a country exports in empirical studies of export concentration.

Third, different authors have suggested that institutional quality and natural resource abundance are correlated, and that the type of commodity produced and exported is important for this relationship[17]. This relationship may be stronger in the case of geographically concentrated and easier to control resources, such as petroleum and certain high-value minerals like gold, with the direction of causality possibly running both ways (see Frankel, 2010; Van der Ploeg, 2011).

Therefore, including indicators of *different types* of natural resource abundance as explanatory variables may indirectly capture the effects of: i) exchange rate misalignment resulting from commodity-dependence; ii) the effects of vulnerability to commodity price volatility; and iii) weaknesses in institutions and governance of the natural resource sector. This is particularly important for empirical studies that cover a relatively large number of commodity-dependent countries.

In practice, measuring the impact of factor abundance on export concentration is challenging, in particular for developing countries where data availability and quality are

[14] This is particularly so in economies with open capital accounts where capital inflows are positively correlated with commodity price movements. Of course, overvaluation may also be due to the combination of high domestic inflation with an inflexible nominal exchange rate.

[15] For example, one question is if yearly (or multi-year) volatility is more or less relevant than monthly volatility (as used by Agosín et al., 2012) for long-term investors introducing new products into an economy or exporting them. More work is needed on a conceptual framework of the relationship between these variables.

[16] Harding and Venables (2016) found a very large negative effect of natural resource exports on non-resource exports.

[17] Measurement of institutional quality is, however, difficult, especially in developing countries.

problematic. There are different measures of natural resources abundance. However, they all are associated with either conceptual or empirical challenges. For example, the measure of the size of natural resource "rents" is associated with conceptual issues about what the concept actually measures, suffers from data quality and availability problems in developing countries, and raises the issue of comparability among countries and different types of resources (such as agricultural "rents"). Using a measure of the size of deposits of non-renewable resources such as mining, or energy commodities also raises similar issues. For this reason and in view of the limitations of using export shares of each commodity group over total exports as discussed above, this paper uses the export values of each commodity group as a percentage of GDP. This measure is not only widely available for developing countries but also it is easily comparable across countries.

For the physical capital stock variable, it is *a priori* difficult to determine its effect on export concentration as this depends on the type of physical capital considered. For example, one could expect that physical capital indicators like infrastructure that are not sector-specific[18] would be negatively correlated with concentration. With aggregate measures like physical capital to GDP ratios it is more difficult to determine *a priori* the sign of the relationship.

Therefore, the specific measures of physical capital and the country sample used may strongly influence the results. For example, Bebczuk and Berrettoni (2006) found the fixed capital-to-GDP ratio and the number of telephone lines per 1000 people (an indicator of infrastructure) to be significantly and positively correlated with concentration, measured by the HHI. Cadot et. al. (2011b) found instead a negative correlation between Theil's T Index and an aggregate infrastructure index containing also the number of telephone lines per 1000 people, the length of the road and railway network, and the share of paved roads in total roads. This ambiguity and the measurement issues may explain why several studies on export concentration have not included this variable among explanatory variables.

The level of development

The level of development has been found to be an important determinant of export concentration. This is because it represents a country's productive structure, including the product space where domestic production takes place and the extant "capabilities" that help to determine that product space, as pointed out by Hausmann and Hidalgo (2011). The capacity of a country to produce a product, as indicated by Hidalgo et al. (2007), is conditional on the availability of the specific inputs associated with its production, labelled "capabilities", while products also differ in terms of the number and sophistication of the capabilities they require to be produced. This includes inputs like intermediate goods and services, know-how, capital goods and specialised human capital, a variety of public goods both at the microeconomic level (e.g. infrastructure services, regulatory capacity, etc) and macroeconomic level (e.g. price and currency stability, security, property rights). All these elements are usually more readily available in countries with a higher level of development, usually proxied by GDP per capita. Countries with few "capabilities" can produce and export only a small number of products, usually not very sophisticated, leading to export concentration on basic goods such as commodities. Hausmann and Hidalgo (2011) also propose that "capabilities" benefit from positive network externalities: the usefulness of each additional capability in terms of the number of products (of growing complexity) that can be produced, is related to the number (and sophistication) of existing capabilities.

Therefore, the value of the marginal capability is much higher in a country with many (and advanced) capabilities than in one with few (and basic) ones. If the number and level of sophistication of "capabilities" available in a country increases with the level of

[18] An oil or gas pipeline is sector-specific and would a priori increase concentration. A port or airport is not sector-specific and would *a priori* not increase export concentration.

development, one could expect that production and export concentration in certain products would fall in a non-linear way, due to the mentioned network externalities in capabilities. Koren and Tenreyro (2013) showed that as a country develops, new varieties of inputs are produced in the economy, leading to productive (and possibly export) diversification. Therefore, one would expect that concentration decreases with the level of development.

Different empirical studies confirm the existence of a statistically significant negative relationship between a country's export concentration and economic development, but the relationship may be non-monotonic. Cadot et al. (2011a) found that the relationship between export concentration and the level of development is not only statistically significant but also non-linear and non-monotonic: while diversification increases as income grows, beyond some level of income, diversification decreases again. Cadot et al. (2011a) explain why countries first diversify and then concentrate as they develop by invoking the Heckscher-Ohlin specialization model proposed by Schott (2003). The argument is that as countries change their production factor endowment by accumulating capital, they move to production cones producing products that are more capital-intensive, so that low-capital and high natural-resource endowment countries concentrate on commodities, while high-capital countries concentrate on non-commodities.

Cadot et al. (2011a) propose that the product lines associated with the old specialization cone (i.e. natural-resource-intensive) are slow to disappear for different reasons. Capital accumulation leads to diversification of production and exports, until the time when such activities slowly wane and re-concentration (probably, on non-commodities, as indicated by UNCTAD, 2019) occurs. Bahar (2016) also found empirical evidence of re-concentration at high income per capita levels, even after controlling for commodity dependence and using trade data to measure concentration with different levels of disaggregation. Other empirical studies that relate export concentration negatively to the level of development include Cadot et al., (2011b; 2013), Parteka (2010), and Parteka and Tamberi (2013). Therefore, one could think that both the comparative advantage and accumulation of capabilities arguments might complement each other in explaining the pattern of export concentration19. However, a consistent economic model using both arguments is still needed.

One well-known challenge of empirically estimating the impact of GDP per capita on concentration is the possibility that the direction of causality runs in both ways: while the level of development of a country affects export concentration, the latter also affects the level of development, especially when concentration is on commodity production and export, as we discuss in Section 2.3.

Institutional quality

Institutional quality can also influence export concentration, because productive and export diversification require investment into higher-productivity sectors of the economy that lead to new products being produced and exported competitively. In turn, good institutions foster investment, so one would expect a priori that institutional quality is negatively correlated with export concentration.

Institutional quality has many different dimensions, such as the degree of political stability, government effectiveness, respect for the rule of law, the degree of corruption, the quality of regulation, and others. Among the essential elements associated with investment in new products is a functioning public sector that does not hinder private investment into new sectors through for example inadequate regulation, or lack of capacity to deliver public goods (e.g. security, rule of law).

[19] Additionally, both these arguments were applied only to merchandise trade. It is also possible that countries also diversify into the production of new (and higher value-added) exportable services intensive in human capital as they become richer. The analysis of trade in services, however, is beyond the scope of this paper.

The importance of institutions for development has been amply analysed (Acemoglu et al., 2002, 2003; Rodrik et al., 2004). Cadot et al. (2011b) found that an indicator of the quality of government was negatively and significantly correlated with Theil's T Index.

Controlling for institutional quality in empirical work is fraught with challenges.[20] First, the variable is highly correlated with the level of income of a country, another explanatory variable. Second, there are different indicators that measure different aspects/dimensions of institutional quality, such as government effectiveness, political stability, regulatory quality, rule of law, and others. These are jointly important even though they are also usually highly correlated. In turn, this raises challenges about heterogeneous measurement issues, aggregation of these variables or choosing which variable to use. In this paper, we construct a summary indicator of institutional quality combining all these different dimensions.

Export prices

Export prices affect export concentration both directly via their impact on the value of exports and indirectly via their impact on resource allocation. High (low) export prices of the products that a country exports, *ceteris paribus*, increase (decrease) export concentration in the short run, while increasing incentives to channel investment to sectors where prices are high, especially when such high export prices are persistent[21]. In natural resource-abundant countries, high commodity prices often result in increased production of the commodities with favourable relative prices. Higher concentration on the commodity sector might also result from price shocks that lead to persistent relative price distortions, negatively affecting the non-resource sector (i.e. *Dutch Disease*). Therefore, *a priori* one would expect a weighted measure of export prices to be positively correlated with export concentration. Agosín et al. (2012) found that the terms of trade variable (the ratio of import prices to export prices) has a significant and positive relationship with the Gini coefficient and Theil's T index but not with the HHI index of export concentration.

Trade barriers and trade costs

Modern theories of trade that account for heterogenous firms in differentiated product markets (Melitz and Reading, 2014) brought to attention the importance of factors that affect the capacity of individual firms to become exporters. Among these, the capacity and cost of accessing foreign markets is an important determinant of export concentration. Following Melitz (2003), it would be expected that *a priori* higher costs of access to foreign markets, either via trade barriers or transport costs, would result in a narrower variety of products being exported, and thereby higher export concentration.

Several empirical studies, including Bebczuk and Berrettoni (2006), Cadot et al. (2011b), Agosín et al. (2012), Parteka and Tamberi (2013) (the latter, for manufacturing exports only), and Dennis and Shepherd (2011), attempt to test this relationship differently. The variables used in panel studies by these authors to account for trade costs include: the availability of transport infrastructure, distance to foreign markets weighted by the economic size of such markets, indirect measures of transport costs, participation of the country in preferential trade agreements weighted by the relative size of each trade partner with whom the preferential trade agreement exists (see Cadot et al., 2011b), the degree of "openness" of the economy, and others. As a group, these studies found that higher trade costs, economic distance and less participation in regional trade agreements increase export concentration.

[20] This, in addition to the challenges posed by comparability between countries, especially developing ones, of variables that are essentially qualitative and measured using opinion surveys, subjective information from analysts and other non-quantitative sources.

[21] In particular, while volatile in the short and medium run, commodity prices have in the past followed persistent price cycles over time (see Jacks, 2013).

The size of the economy

Another well-known determinant of export concentration is the size of the economy. A large domestic economy is expected to provide domestic firms with a larger market for their production and allow more firms that otherwise would have not been competitive internationally to survive. In the presence of economies of scale, and/or positive spillovers in terms of labour and input markets, these firms would then be internationally competitive, decreasing export concentration. Therefore, the size of the economy, which is usually proxied by the size of the population, would be *a priori* negatively correlated with export concentration. Empirical studies like Cadot et al. (2011b) and Parteka and Tamberi (2013) found a significant negative relationship between the size of the population and export concentration.

In this study we use the size of the urban population instead of the total population as a proxy for the size of the economy of a country. The reason is that urban population may be a better indicator of the size of the domestic market in developing countries where the rural population, often large, has substantially less buying power and access to goods (and services) than the urban population.

Foreign direct investment

Concentration could also be influenced by the size and composition of foreign direct investment (FDI). FDI may increase domestic production and export capabilities by direct investment into new productive sectors, as well as by raising productivity at the firm level in existing sectors. FDI can also increase capital accumulation in credit-constrained countries or sectors, raise sectoral availability of R&D and technology (e.g. embedded in capital goods and human capital), thereby raising the probability of positive productivity spillovers in the economy. FDI also increases the likelihood[22] that a country can integrate an international value chain (Harrison and Rodríguez-Clare, 2010).

In particular, as value addition has become increasingly disaggregated across countries, FDI can also contribute to fostering diversification if firms that are active in international value chains leverage their existing supplier and customer networks to produce domestically for some segments of the chain. Alternatively, FDI into resource extraction in the mining and energy sectors may result in more concentration if it helps the creation of “enclave” industries with little value added domestically, and limited employment generation (especially, of domestic human capital), technology transfer and integration into domestic value chains. Therefore, *a priori* the sign of the relationship between foreign direct investment and export concentration is unclear, depending on the type of FDI, the destination sector, and other factors.

However, the empirical evidence relating FDI and export concentration in past studies is underwhelming. For example, Cadot et al. (2011b) found the FDI-to-GDP ratio to be a determinant of concentration, albeit with statistical significance at 10 per cent and a very small coefficient. Agosin et al. (2012), Parteka and Tamberi (2013) and Bahar and Santos (2018) did not include FDI as a determinant of export concentration in their empirical studies.

22 This may occur in different ways. For example, search and transaction costs may be reduced in cases of within-firm trade.

4. Empirical Analysis

Data Description

The dataset used in this paper has 3956 observations covering 173 countries over the period 1995-2017. Our dataset includes 38 developed countries, 40 least developed countries and 95 Other developing and transition countries. Out of the 173 countries in the sample, 92 are commodity dependent (87 of which are developing and transition countries). The Appendix details the countries included in the dataset.

This dataset has wider coverage than most past studies, and it includes large and small countries from every region, and at different levels of development. This addresses the problem of selection besetting some studies where small developing countries and transition economies are not properly represented. As many of these countries are among the most export concentrated, their inclusion makes the analysis and its policy implications more robust.

The dependent variables are the widely used measures of export concentration discussed earlier: the logit-transformed[23] normalized Herfindahl-Hirschman Index (HHI), the logit-transformed Gini coefficient and Theil's T index. These concentration indexes were calculated using UNCTADStat data for gross[24] merchandise exports, disaggregated to three digits of the SITC, third revision. We use total allocated[25] export data, including exports of commodities, to calculate export concentration measures, following convention.[26] For 104 countries in our sample, commodity exports in 2017 were more than half of registered exports, and for 61 and 30 of those countries, commodities represented more than 80 per cent and 90 per cent of exports, respectively. This highlights the importance of analysing commodity exports to understand export concentration.

We measure Theil's T index for every country-year in two ways: i) the "standard" way, using the number of registered trade lines as *N*; and ii) using the maximum number of possible trade lines as *N* for every country, which for the SITC revision 3 classification is 261 product lines. This "adjusted" Theil T index addresses the volatility of Theil's T index for countries with highly volatile numbers of export lines from year to year, and it may be a better measure of concentration if the year-to-year variation observed in the number of exported product lines is mostly due to measurement error. On the other hand, if such year-to-year changes are due to real product line attrition, the original Theil's T index is a better measure of concentration. The use of different measures of concentration increases the robustness of our findings. In any case, even though these indicators of concentration are calculated differently, they are highly correlated, as shown by Table A1. The coefficient of correlation is between 0.92 and 0.98.

[23] The normalized HHI index and Gini coefficient are used in this paper after a logistic transformation, in order to take into account the bounded (i.e. between 0 and 1) nature of both variables, as indicated by Baum and others (2008) and Fox (2016), among others, and following Agosín et al. (2012).

[24] This data does not correct for those transactions registered as "exports" that are effectively "re-exports". Total exports in each country and year corresponds to the addition of exports attributed to each product and identified partner.

[25] Allocated export data are data entries that can be allocated to a specific product line and trading partner in a specific time period.

[26] One exception is Bahar and Santos (2018), who use non-commodity export data to calculate export concentration.

The covariates used are those discussed in Section 3. The level of development is measured by GDP per capita, calculated in logs with base 10,[27] and its square, in order to test for the existence of a non-linear relationship between GDP per capita and export concentration, as found by Cadot et al. (2011). As a result of the conceptual discussion in Section 3 about possible endogeneity of the income variable, we treat GDP per capita and its square as predetermined variables and use the first lag of these variables in our dynamic empirical models below. GDP per capita data was obtained from UNCTADStat.

The importance of factors of production abundance was tested by including the value of commodity exports as a percentage of GDP[28], grouping such commodity exports into three categories: agricultural exports; minerals; and energy exports.[29]

Following UNCTAD (2019), we considered the following SITC revision 3 codes to construct each group: the agricultural products group includes agricultural raw materials (codes 21X, 23X, 24X, 25X, 26X, and 29X) and also food, tropical beverages and vegetable oils and fats (codes 0XX, 1XX, 4XX and 22X); the minerals group includes codes 27X, 28X, 667, 68X and 971; while the energy group contains codes 3XX. Importantly, using SITC revision 3 at three digits instead of other classification systems such as HS with higher individual product disaggregation used by other studies on concentration, we are able to obtain relatively disaggregated trade data for a very large number of developing countries, which is often not possible using other data sources with more disaggregated trade data. Three digits is the highest degree of disaggregation available in UNCTADStat database.

Urban population, measured in logs, is used to proxy for the size of the economy, as discussed earlier. Urban population is calculated using data from UNCTADStat.

In order to control for institutional quality, we construct an indicator of institutional quality using data from the Worldwide Governance Indicators of the World Bank. Specifically, we construct a weighted average of six different indicators that incorporate different dimensions of institutional quality, namely *Control of Corruption, Government Effectiveness, Political Stability and Absence of Violence/Terrorism, Regulatory Quality, Rule of Law,* and *Voice and Accountability*. The weights of these variables in the aggregate indicator were calculated from the data and correspond to the contributions to the first principal component, which explains 55 per cent of the variance, by each of these variables.[30] We include the first lag of our institutional quality variable as a covariate instead of the current value because the effects of institutional quality on concentration come with a lag. This also helps address the issue of potential endogeneity between institutional quality and export concentration discussed in Section 2.

The effects of export prices are included in our model by an Index of yearly export prices.

[27] Note that for all logged variables we use base 10.

[28] By using exports of each commodity type as a percentage of GDP instead of as a percentage of total exports we avoid problems that may arise from the fact that the latter is clearly correlated with concentration measures like the HHI or Theil's T index, by construction.

[29] In the category "abundance of factors of production" covariates, and due to the conceptual and practical issues discussed in Section 3 and following previous studies like Agosín et al. (2012) and Bahar and Santos (2018), we do not include physical capital abundance indicators in our estimations. Also, data from the Penn World Tables Human Capital Index was used in estimations but not included in our results. The reason is that the data on this variable was available until only 2014 and the dataset has a large number of missing observations (1209 out of 3956 observations in total, or 30.5 per cent), which especially affects developing countries (1075 missing observations out of 3082, or 35 per cent), calling into question the representativity of the data for our purposes. Additionally, the Human Capital Index data had very high collinearity with GDP per capita data (78 per cent, as shown by Table A1).

[30] The specific weights are: *Control of Corruption* (18.6 per cent), *Government Effectiveness* (18.04 per cent), *Political Stability and Absence of Violence/Terrorism* (12.12 per cent), *Regulatory Quality* (18.78 per cent), *Rule of Law* (17.08 per cent), and *Voice and Accountability* (15.39 per cent).

This index, included in log form, was constructed using a weighted sum[31] of the Index prices for Food, Tropical Beverages, and Vegetable Oils, Agricultural Raw Materials; Minerals, Ores and Metals; the price of petroleum and the World Bank's Manufactures' Unit Value Index, sourced from UNCTADStat. The weights used are the shares of total merchandise exports of agricultural products, mining, energy and non-commodities, calculated for each country from UNCTADStat trade data.

Trade costs measures were included by using two different variables. First, we include the "economic distance" of a country, calculated using the weighted distance between countries obtained from the CEPII *GeoDist* database. The weights are the shares of each country's GDP obtained from UNCTADStat:

$$wdist_{it} = \sum \frac{GDP_{jt}}{GDP_{Wt}} log(D_{ij}),$$

where GDP_{jt} is the GDP of each of country i j=1,2,...,J trading partners in year t; GDP_{Wt} is the World's GDP in year t; and D_{ij} is the distance (weighted by city population)[32] between countries i and j.

Second, we include a Preferential Market Access variable, calculated using the methodology in Cadot et al. (2011b):

$$pta_{ijt} = \sum_{j=1}^{J} \frac{GDP_{jt}}{GDP_{Wt}} PBA_{ijt}$$

where GDP_{jt} is the GDP of each country i j=1,2,...,J trade partners in year t; GDP_{Wt} is the World's GDP in year t; and PBA_{ijt} is an indicator variable that takes value 1 if there is a preferential trade agreement between countries i and j. Data on existing bilateral trade agreements was obtained from Jeffrey Bergstrand's webpage at the University of Notre Dame.[33]

We also include Foreign Direct Investment as a percentage of GDP in our model, using data from UNCTADStat. One challenging empirical issue associated with the use of this variable is the determination of the lag structure. While it is likely that the effects of FDI on concentration operate with a lag,[34] the lag length is difficult to identify and may vary across countries, sectors and over time. This paper uses one lag, but other lag lengths were used with similar qualitative results. Table 1 below shows some descriptive statistics of the variables used in the econometric model.

[31] The methodology of weighting some form of commodity prices according to country export products is common in the literature, with differences in the actual prices and weights used (e.g. see the calculation of "world commodity prices" in Bahar and Santos, 2018).

[32] Specifically, the *distw* variable from CEPII was used, as it includes data from cities outside the main one or the capital too. See http://www.cepii.fr/PDF_PUB/wp/2011/wp2011-25.pdf.

[33] See https://www3.nd.edu/~jbergstr/DataEIAsApril2017/EIADatabaseApril2017.zip. One limitation of this data is that it is only available yearly until 2012, so 2013 to 2017 repeated the observation of 2012. This may be one of the reasons behind the low explanatory power of this variable in our estimations below.

[34] Among other things, this may be due to "time to build" consideration for greenfield investment, the delayed impact of FDI when associated investments in human resources are necessary, and others.

Table 1:
Summary Statistics

Variable	1st Qu.	3rd Qu.	Max.	Mean	Median	Min.	NAs
Theil's T Index	1.523	2.981	5.005	2.341	2.2335	0.638	0
Theil's T Adj. Index	1.585	3.352	5.209	2.546	2.421	0.646	0
Logit HHI	-3.196	-1.206	2.000	-2.108	-2.214	-4.949	0
Logit Gini	1.512	3.341	6.705	2.513	2.385	0.395	0
Log GDP p.c.	3.087	4.143	5.054	3.632	3.610	2.205	0
Agricult. Expo/GDP	0.017	0.074	0.468	0.054	0.035	0.000	1
Energy Expo/GDP	0.002	0.050	0.778	0.063	0.009	0.000	1
Minerals Expo/GDP	0.003	0.029	0.470	0.032	0.009	0.000	1
Log Urban Population	3.137	4.087	5.912	3.558	3.605	1.160	0
Institut. Quality	-0.722	0.606	1.995	-0.060	-0.251	-2.072	33
Log Export Price Index	2.030	2.272	2.623	2.163	2.150	1.711	0
Economic Distance	3.823	3.967	4.143	3.904	3.911	3.746	0
Prefer. Market Access	0.445	0.727	0.893	0.580	0.621	0.000	79
FDI/GDP	1.072	5.376	499.6	4.821	2.586	-58.326	85

Note: NA stands for missing observations.

Empirical Results

Inspection of our data shows that all measures of concentration are persistent, which leads us to estimate a dynamic panel data model using the Generalized Method of Moments estimator following Blundell and Bond (1998). Variables in levels and differences are used as instruments (i.e. "system" GMM). Fixed country effects and period dummies are also included. In order to address the concerns raised by Roodman (2009) regarding the weakness of Hansen-Sargan joint tests of instrument validity in the face of multiple instruments, we use three lags of the dependent variable as instruments starting from the second lag.

The model estimated is:

$$y_{it} = \beta_0 y_{it-1} + \beta_1 X_{it} + \mu_i + \gamma_t + \epsilon_{it}$$

with y_{it} *representing* $normhhi_{it}, adjtheil_{it}, theil_{it}, gini_{it}$, X_{it} being a matrix containing the explanatory variables presented above, μ_i the country fixed effects, γ_t the period fixed effects and ϵ_{it} the errors.

The results are presented in Table 2 below for the four measures of concentration, namely, the logit transformed normalized HHI, the logit transformed Gini coefficient, Theil's T index and the adjusted Theil's T Index, respectively. Standard errors in Table 2 are calculated using robust estimates of the coefficient covariance matrix, as proposed by Windmeijer (2005).

We observe that all the coefficients in Table 2 that are statistically significant have the expected signs as discussed in Section 3 above. Three covariates were found to be non-significant, namely, agricultural exports as a share of GDP, the preferential market access variable, and foreign direct investment as a share of GDP.

To answer our research question about the need to account for the type of export in explaining export concentration, we find in all models that energy exports as a share of GDP is positively and significantly associated with all measures of export concentration.

Table 2:
Generalized Method of Moments – Level and Differences Instruments

	1-HHI	2-Gini	3-Theil's T	4-Adj.Theil's T
Logit Norm.HHI (1 lag)	0.727***			
	(0.038)			
Logit Gini Coeff. (1 lag)		0.674***		
		(0.039)		
Theil's T (1 lag)			0.688***	
			(0.028)	
Adj.Theil's T (1 lag)				0.713***
				(0.028)
Energy Expo/GDP share	1.313***	1.299***	1.198***	1.093***
	(0.246)	(0.216)	(0.164)	(0.166)
Minerals Expo/GDP share	0.750*	0.598*	0.846***	0.594***
	(0.392)	(0.315)	(0.239)	(0.221)
Agricult. Expo/GDP share	-0.344	-0.406	-0.227	-0.357
	(0.318)	(0.338)	(0.247)	(0.268)
Log GDP per capita (1 lag)	-0.761**	-0.947***	-0.661***	-0.704***
	(0.330)	(0.288)	(0.240)	(0.228)
Square Log GDP per capita (1 lag)	0.097**	0.120***	0.085**	0.091***
	(0.048)	(0.042)	(0.035)	(0.034)
Institut. Quality (1 lag)	-0.117***	-0.131***	-0.100***	-0.116***
	(0.045)	(0.038)	(0.035)	(0.036)
Log Urban Population	-0.152***	-0.189***	-0.104***	-0.152***
	(0.030)	(0.032)	(0.021)	(0.024)
Export Prices	0.589***	0.513***	0.459***	0.475***
	(0.190)	(0.157)	(0.152)	(0.140)
Econ. Distance	0.530**	0.667***	0.519***	0.536***
	(0.220)	(0.223)	(0.170)	(0.177)
Preferential Market Access	-0.008	0.012	0.038	0.016
	(0.096)	(0.082)	(0.072)	(0.067)
FDI/GDP (1 lag)	0.000	0.000	0.000	0.000
	(0.001)	(0.000)	(0.000)	(0.000)
n	173	173	173	173
T	23	23	23	23
Num. obs.	3956	3956	3956	3956
Sargan Test: chisq	113.261	119.937	127.823	114.884
Sargan Test: p-value	0.057	0.023	0.007	0.046
Autocorrelation test (1): p-value	0.000	0.000	0.000	0.000
Autocorrelation test (2): p-value	0.579	0.822	0.126	0.282
Wald Test Coefficients: chisq	4848.658	5146.726	4750.458	6530.676
Wald Test Coefficients: p-value	0.000	0.000	0.000	0.000
Wald Test Time Dummies: chisq	101.866	94.018	105.592	120.437
Wald Test Time Dummies: p-value	0.000	0.000	0.000	0.000

***p < 0.01, **p < 0.05, *p < 0.1

For the four models in Table 2, the variable mineral exports as a share of GDP is also positively associated with export concentration, but with lower significance levels for both the HHI and Gini coefficients – at 5.5 per cent and 5.7 per cent significance levels, respectively. The effect of energy exports on export concentration is much stronger than that of minerals exports, as evidenced by the coefficient of energy exports being between 42 per cent and 117 per cent larger than that of mineral exports, depending on the dependent variable.

Combined with the observation that agricultural exports as a percentage of GDP do not significantly affect export concentration in any model, our results lend support to the hypothesis that different types of commodity exports affect concentration in different ways. Note that in addition to the direct effect of the variable, there is also an indirect one whereby concentration increases with a lag due to increased production as discussed in Section 3. Moreover, as minerals and energy commodity prices are more volatile than food prices (see UNCTAD, 2018), this may have impacted export concentration via the effects of volatility and real effective exchange rate appreciation on investment.

As we discussed in Section 3, the export shares on GDP of different types of commodities may be proxying for complex effects that affect the composition of exports (and hence, concentration) which are difficult to control for in a linear panel data model, due to different measurement problems mentioned above, especially for developing countries. Agosín et al (2012) tried to measure such effects in their dynamic panel data model by using a measure of exchange rate volatility and one of overvaluation, which they found not to be significant for the Gini coefficient or Theil's T index, and only significant at 10 per cent for the HHI. Additionally, the increases in energy and mineral commodity prices during the last commodity price boom (2003-2011) were much higher than those of agricultural products, possibly resulting in very different direct and indirect effects on export concentration *ex post*.

Table 2 shows that lagged GDP per capita is negatively correlated with export concentration and statistically significant, while its square is positively correlated with export concentration. This confirms the findings of other studies (e.g. Cadot et al., 2011b, Parteka and Tamberi, 2013) about the non-monotonic relationship between the level of development and export concentration that was discussed in Section 3 above.

Similarly, a large domestic economy, proxied by the log of urban population, is associated with lower concentration, as expected and matching the results of Cadot et al. (2011b) and Parteka and Tamberi (2013).

Institutional quality is negatively and significantly correlated with export concentration for all measures of concentration. This matches the *a priori* expectation following the discussion in section 3, and also confirms the results of previous studies. For example, Cadot et al. (2011b), using a static OLS model with fixed effects with 87 countries and data from 1990 to 2004 found a negative correlation between two measures of institutional quality: the ICRG Quality of Government indicator and the Revised Combined Polity Score calculated by the Quality of Government Institute. Our results then reinforce the idea that the effects of energy or (to a lesser extent) mineral-export dependence on concentration are separate from the effects of weak institutions.

We also observe that our estimate of export prices for each country, controlling for their trade structure, are significantly and positively correlated with export concentration, for all measures of concentration. This matches the discussion in Section 3 and with the significant and positive relationship found by Agosín et. al (2012) between the terms of trade and the Gini coefficient as well as the Theil T index. Our results suggest that export prices are an important factor of export concentration

Table 2 also shows that the economic distance of a country is positively correlated with export concentration: the longer the distance to markets, the more concentrated exports are, confirming the findings of Cadot et. al. (2011b), Agosín et al. (2012) and Parteka and Tamberi (2013). This suggests that trade costs are relevant variables explaining concentration.

The preferential market access variable is not significant for any measure of export concentration. Using a static model with 87 countries for the period 1990-2004, Cadot et al. (2011b) find that the variable is significant. One possible explanation for the lack of significance of the preferential market access variable in our results may relate to problems with the data used to construct the preferential market access variable; our constructed variable may not be a good proxy for preferential market access. Another possibility is related to the fact that our data contains a very large number of developing countries, and it might be the case that the positive effects of trade openness on diversification, as posited in the Melitz heterogeneous firm trade model (Melitz, 2003), may weaker in commodity-dependent developing countries such as the LDCs that have weak productive capacities to build upon.

Finally, Table 2 suggests that lagged FDI as a percentage of GDP is not a significant explanatory variable of export concentration. Cadot et. al. (2011b) found a significant impact but with a coefficient close to zero, so our results are qualitatively similar to their findings. A look at the individual country relationships between concentration and FDI suggests heterogeneity across countries. The question is, therefore, whether our model sufficiently controls for the other factors that may explain that heterogeneity (such as economic structure, here proxied by share of commodities in GDP). In particular, commodity exports by type[35] as a share of GDP, economic distance, and the other control variables included in our model may not be capturing all the elements relating concentration and FDI. For example, not only aggregate data of FDI as a percentage of GDP fails to account for the type and quality of FDI (e.g. sectoral, "green field" vs privatization, etc), but also the relationship even between the size of FDI and concentration may be non-linear on variables like the level of development or the economic structure.

[35] An empirical attempt (not included in our results) to test whether the effect of FDI on concentration varied according to the share of each commodity type on exports (a proxy for the type of commodity dependence) did not find the relationship significant across different measures of concentration and empirical models.

5. Conclusions

Studying export concentration is important for development, as high export concentration is associated with low levels of development. In this paper, using a dynamic panel data model for 173 countries covering the period 1995 – 2017, we find that the type of commodities that a country exports is an important determinant of its degree of export concentration. This result holds after controlling for several important determinants of concentration previously identified in the economic literature, such as the level of income of the country, the size of the economy, institutional quality, export prices and the economic distance to markets. In particular, we find that high shares of energy and, to a lesser extent, mineral exports (i.e. non-renewables) are important determinants of export concentration. Our findings are robust to the use of different measures of export concentration, namely the HHI index, the Gini coefficient and Theil's T index, measured in two different ways.

It is possible that different determinants of concentration act through indirect impacts, via joint effects due to interactions among themselves in complex ways. Even though these processes are not specifically studied in this paper, such variables may include the real exchange rate and its misalignment, exchange rate volatility, and other potentially important determinants of concentration. Therefore, our inclusion of different types of commodity exports as a share of GDP may proxy for these indirect effects. The reason for this, as we discussed above, is that energy and mineral commodity prices during the period of analysis were more volatile and experienced stronger shocks than agricultural commodities. Therefore, the effects of volatility and real exchange rate appreciation associated with different types of commodities are different. This study may have indirectly captured them at least partially through the decomposition of commodity exports by commodity types. In this regard, our study complements previous work and may offer an alternative way of addressing the challenges of past studies that tried to measure directly the effects of those variables on concentration.

Using a higher number of control variables and a larger dataset, our work also confirms and extends the findings of Cadot et al. (2011a) that the relationship between concentration and GDP per capita is better characterized by a quadratic functional form.[36]

Our empirical findings have important policy implications for developing countries. In particular, they highlight the relevance of the policy prescription that commodity-dependent developing countries, particularly those dependent on the export of minerals and, especially, energy, need to diversify their economies and export sectors.

Finally, the results presented in the paper open several areas for future work, some of which remain empirically challenging like better inclusion of different types of physical and human capital in the model, foreign direct investment, transport costs and trade barriers, among others. Extending research in these areas should be mindful of the need for representativity, especially the inclusion of developing countries, the ones most concerned by the issue of export concentration.

[36] Cadot et al. (2011a) use a dataset with 2497 observations, corresponding to 141 countries over the 1988-2006 period, while our dataset has 3956 observations for 173 countries for the 1995-2017 period. Note that our dataset includes the commodity price boom of the 2000s, which is important given that export prices are significantly and positively correlated with export concentration, as we show.

6. Bibliography

Acemoglu, D. Johnson, S. and Robinson, J.A. (2002), "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution." *The Quarterly Journal of Economics* 117 (4), pp. 1231–94.

Acemoglu, D., Johnson, S., Robinson, J.A. and Thaicharoen, Y. (2003), "Institutional Causes, Macroeconomic Symptoms: Volatility, Crises and Growth." *Journal of Monetary Economics* 50 (1), pp. 49–123.

Aghion, P., Angeletos, G.M., Banerjee, A. and Manova, K. (2010), "Volatility and Growth: Credit Constraints and the Composition of Investment", *Journal of Monetary Economics* 57 (3), pp. 246–65.

Agosín, M.R., Alvarez, R. and Bravo-Ortega, C. (2012), "Determinants of Export Diversification Around the World: 1962–2000", *The World Economy* 35 (3), pp. 295–315.

Aguirre, A., and Calderón, C. (2005), "Real Exchange Rate Misalignments and Economic Performance", *Documentos de Trabajo* (Banco Central de Chile), no. 315, pp. 1–49.

Bahar, D. and Santos, M.A. (2018), "One more resource curse: Dutch disease and export concentration", *Journal of Development Economics*, Vol. 132, pp. 102 – 114.

Bahar, D. (2016), "Diversification or Specialization: What it the Path to Growth and Development?", Global Economic and Development at Brookings Policy Brief, November 2016.

Barlevy, G. (2007), "On the Cyclicality of Research and Development", *American Economic Review* 97 (4), pp. 1131–64.

Barro, R.J. and Lee, J.W., 2001. "International data on educational attainment: updates and implications", *Oxford Economic Papers*, 53(3), pp. 541-563.

Baum, C.F, and others. (2008), "Stata Tip 63: Modeling Proportions." *Stata Journal* 8 (2): 299.

Bebczuk, R.N. and Berrettoni, D. (2006), "Explaining Export Diversification: An Empirical Analysis", Universidad Nacional de la Plata, Departamento de Economía Documentos de Trabajo 65, October 2006.

Blattman, C., Hwang, J. and Williamson, J.G. (2007), "Winners and Losers in the Commodity Lottery: The Impact of Terms of Trade Growth and Volatility in the Periphery 1870–1939", *Journal of Development Economics* 82 (1), pp. 156–79.

Blundell, R., and Bond, S. (1998), "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models." *Journal of Econometrics* 87 (1), pp. 115–43.

Cadot, O., Carrère, C. and Strauss-Kahn, V. (2011a), "Export Diversification: What's Behind the Hump?" *The Review of Economics and Statistics* 93 (2), pp. 590–605.

———. 2011b. "Trade Diversification: Drivers and Impacts." In *Trade and Employment. from Myths to Facts*, Jansen, Marion, Ralf Peters, and José Manuel Salazar-Xirinachs., pp. 253–96. Geneva: International Labour Office.

———. 2013. "Trade Diversification, Income, and Growth: What Do We Know?" *Journal of Economic Surveys* 27 (4), pp. 790–812.

Corden, W. M. 1984. "Booming Sector and Dutch Disease Economics: Survey and Consolidation." *Oxford Economic Papers* 36 (3): 359–80.

Cowell, F. (2011), *Measuring Inequality*, Oxford University Press, Oxford: 2011.

De Benedictis, L., Gallegati, M. and Tamberi, M. (2009), "Overall Trade Specialization and Economic Development: Countries Diversify." *Review of World Economics* 145 (1), pp. 37–55.

Dennis, A., and Shepherd, B. (2011), "Trade Facilitation and Export Diversification." *The World Economy* 34 (1), pp. 101–22.

Feenstra, R.C., Lipsey, R.E., Deng, H., Ma A.C. and Mo, H. (2005), "World Trade Flows: 1962-2000", NBER Working Paper No. 11040, January.

Fox, J. (2016), *Applied Regression Analysis and Generalized Linear Models*. Sage Publications.

Frankel, J. A. (2010) "The Natural Resource Curse: A Survey." Working Paper 15836. National Bureau of Economic Research.

Harding, Torfinn, and Anthony J. Venables. 2016. "The Implications of Natural Resource Exports for Nonresource Trade." *IMF Economic Review* 64 (2): 268–302.

Harrison, Ann, and Andrés Rodríguez-clare. 2010. "Trade, Foreign Investment, and Industrial Policy,", *Handbook of Development Economics Volume 5*, edited by D. Rodrik and M. Rosenzweig, 2010.

Hausmann, Ricardo, and César A. Hidalgo. 2011. "The Network Structure of Economic Output." *Journal of Economic Growth* 16 (4): 309–42.

Hidalgo, C. A., B. Klinger, A.-L. Barabási, and R. Hausmann. 2007. "The Product Space Conditions the Development of Nations." *Science* 317 (5837): 482–87.

Jacks, D.S. (2013), "From Boom to Bust: A Typology of Real Commodity Prices in the Long Run", Working Paper No. 18874, National Bureau of Economic Research, 2013.

Klinger, B. and Lederman, D., 2009. "Diversification, innovation, and imitation of the global technological frontier", in Newfarmer, R., Shaw, W. and Walkenhorst, P., editors: *Breaking Into New Markets: Emerging Lessons for Export Diversification*, Washington D.C.: World Bank, pp.101-10.

Koren, Miklós, and Silvana Tenreyro. 2007. "Volatility and Development." *The Quarterly Journal of Economics* 122 (1): 243–87.

———. 2013. "Technological Diversification." *American Economic Review* 103 (1): 378–414.

Melitz, Mark J., and Redding, Stephen J. (2014), "Heterogeneous Firms and Trade", in Handbook of International Economics Vol. 4, edited by Helpman, E. Rogoff, K. and Gopinath, G., Amsterdam: Elsevier North Holland, pp. 1 – 54.

Melitz, Marc J. 2003. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica* 71 (6): 1695–1725.

Michaely, Michael. 1958. "Concentration of Exports and Imports: An International Comparison." *The Economic Journal* 68 (272): 722–36.

Palan, Nicole. 2010. "Measurement of Specialization. the Choice of Indices." Working Paper 62. FIW Working Paper.

Parteka, Aleksandra. 2010. "Employment and Export Specialisation Along the Development Path: Some Robust Evidence." *Review of World Economics* 145 (4): 615–40.

Parteka, Aleksandra, and Massimo Tamberi. 2013. "What Determines Export Diversification in the Development Process? Empirical Assessment." *The World Economy* 36 (6): 807–26.

Rodrik, Dani, Arvind Subramanian, and Francesco Trebbi. 2004. "Institutions Rule: The Primacy of Institutions over Geography and Integration in Economic Development." *Journal of Economic Growth* 9 (2): 131–65.

Roodman, David. 2009. "A Note on the Theme of Too Many Instruments." *Oxford Bulletin of Economics and Statistics* 71 (1): 135–58.

Sachs, Jeffrey D., and Andrew M. Warner (2001), "The Curse of Natural Resources." *European Economic Review*, 15th annual congress of the European economic association, 45 (4): 827–38.

Schott, P.K. (2003), "One Size Fits All? Heckscher-Ohlin Specialization in Global

Production", *The American Economic Review*, Vol. 93 (3), pp. 696 – 708.

UNCTAD (2019), "Commodity Dependence: A Twenty-Year Perspective." Geneva, Switzerland: UNCTAD.

UNCTAD (2018), "Diversification and Value Addition. Note by the UNCTAD Secretariat", available at https://unctad.org/meetings/en/SessionalDocuments/cimem2d42_en.pdf.

Van der Ploeg, Frederick. 2011. "Natural Resources: Curse or Blessing?" *Journal of Economic Literature* 49 (2): 366–420.

Van der Ploeg, Frederick, and Steven Poelhekke. 2009. "Volatility and the Natural Resource Curse." *Oxford Economic Papers* 61 (4): 727–60.

Windmeijer, F. 2005. "A Finite Sample Correction for the Variance of Linear Efficient Two–Steps GMM Estimators." *Journal of Econometrics* Vol 126, pp. 25–51.

Appendix

Countries in the data sample

Our data sample contains 173 countries. Montenegro has 10 yearly observations, Timor-Leste has 15, Kazakhstan and Sudan have 22, and the other 169 countries have 23 observations. The number of NAs in each variable is shown in Table 1 above. The list of countries in our dataset is the following:

Afghanistan
Albania
Algeria
Andorra
Argentina
Armenia
Australia
Austria
Azerbaijan
Bahrain
Bangladesh
Barbados
Belarus
Belgium
Belize
Benin
Bhutan
Bolivia (Plurinational State of)
Bosnia and Herzegovina
Botswana
Brazil
Brunei Darussalam
Bulgaria
Burkina Faso
Burundi
Cabo Verde
Cambodia
Cameroon
Canada
Central African Republic
Chad
Chile
China
Colombia
Comoros
Congo
Costa Rica
Côte d'Ivoire
Croatia
Cuba
Cyprus
Czechia
Dem. Rep. of the Congo
Denmark
Djibouti
Dominica
Dominican Republic
Ecuador
Egypt
El Salvador
Equatorial Guinea
Eritrea
Estonia
Eswatini
Ethiopia
Fiji
Finland
France
Gabon
Gambia
Georgia
Germany
Ghana
Greece
Grenada
Guatemala
Guinea
Guinea-Bissau
Guyana
Haiti
Honduras
Hungary
Iceland
India
Indonesia
Iran (Islamic Republic of)
Ireland
Israel
Italy
Jamaica
Japan
Jordan
Kazakhstan
Kenya
Kiribati
Korea
Dem. People's Rep. of
Korea
Republic of
Kuwait
Kyrgyzstan
Lao People's Dem. Rep.
Latvia
Lebanon
Lesotho
Libya
Lithuania
Luxembourg
Madagascar
Malawi
Malaysia
Maldives
Mali
Malta
Mauritania
Mauritius
Mexico
Mongolia
Montenegro
Morocco
Mozambique
Myanmar
Namibia
Nepal
Netherlands
New Zealand
Nicaragua
Niger
Nigeria
Norway
Oman
Pakistan
Panama
Papua New Guinea
Paraguay
Peru
Philippines
Poland
Portugal
Qatar
Republic of Moldova
Romania
Russian Federation
Rwanda
Saint Kitts and Nevis
Saint Lucia
Saudi Arabia
Senegal
Serbia
Seychelles
Sierra Leone
Singapore
Slovakia
Slovenia
Solomon Islands
South Africa
Spain
Sri Lanka
Sudan
Sweden
Switzerland-Liechtenstein
Syrian Arab Republic
Tajikistan
TFYR of Macedonia
Thailand
Timor-Leste
Togo
Tonga
Trinidad and Tobago
Tunisia
Turkey
Turkmenistan
Uganda
Ukraine
United Arab Emirates
United Kingdom
United Republic of Tanzania
United States of America
Uruguay
Uzbekistan
Venezuela (Bolivarian Rep. Of)
Viet Nam
Yemen
Zambia
Zimbabwe.

Table A1:
Correlations Table – All Variables

	theil	theil_adj	hhi_t	gini_t	lgdp_pc	pwt_hci	agri_gdp	min_gdp	nrg_gdp	lpop	lpop_urb	inst_pca	expo_p	w_dist	pta	fdi_gdp
theil	1.00	0.97	0.96	0.92	-0.33	-0.50	0.04	0.24	0.53	-0.20	-0.27	-0.48	0.26	0.33	0.26	0.01
theil_adj	0.97	1.00	0.98	0.98	-0.39	-0.55	0.06	0.19	0.46	-0.32	-0.39	-0.50	0.21	0.37	0.28	0.01
hhi_t	0.96	0.98	1.00	0.93	-0.35	-0.50	0.04	0.19	0.48	-0.28	-0.34	-0.47	0.21	0.34	0.26	0.02
gini_t	0.92	0.98	0.93	1.00	-0.42	-0.58	0.08	0.15	0.44	-0.35	-0.43	-0.52	0.20	0.38	0.30	0.01
lgdp_pc	-0.33	-0.39	-0.35	-0.42	1.00	0.78	-0.20	-0.10	0.22	-0.15	0.05	0.81	-0.05	-0.37	-0.51	0.09
pwt_hci	-0.50	-0.55	-0.50	-0.58	0.78	1.00	-0.11	0.02	0.00	-0.10	0.10	0.70	0.01	-0.43	-0.53	0.08
agri_gdp	0.04	0.06	0.04	0.08	-0.20	-0.11	1.00	0.08	-0.16	-0.20	-0.24	-0.07	-0.03	0.26	0.11	0.00
min_gdp	0.24	0.19	0.19	0.15	-0.10	0.02	0.08	1.00	-0.05	-0.04	-0.05	-0.04	0.34	0.14	0.15	0.02
nrg_gdp	0.53	0.46	0.48	0.44	0.22	0.00	-0.16	-0.05	1.00	-0.07	-0.01	-0.14	0.18	-0.01	-0.09	0.01
lpop	-0.20	-0.32	-0.28	-0.35	-0.15	-0.10	-0.20	-0.04	-0.07	1.00	0.97	-0.20	0.03	-0.07	-0.07	-0.17
lpop_urb	-0.27	-0.39	-0.34	-0.43	0.05	0.10	-0.24	-0.05	-0.01	0.97	1.00	-0.06	0.03	-0.16	-0.15	-0.14
inst_pca	-0.48	-0.50	-0.47	-0.52	0.81	0.70	-0.07	-0.04	-0.14	-0.20	-0.06	1.00	-0.15	-0.30	-0.43	0.11
expo_p	0.26	0.21	0.21	0.20	-0.05	0.01	-0.03	0.34	0.18	0.03	0.03	-0.15	1.00	0.12	0.11	0.03
w_dist	0.33	0.37	0.34	0.38	-0.37	-0.43	0.26	0.14	-0.01	-0.07	-0.16	-0.30	0.12	1.00	0.38	-0.06
pta	0.26	0.28	0.26	0.30	-0.51	-0.53	0.11	0.15	-0.09	-0.07	-0.15	-0.43	0.11	0.38	1.00	-0.04
fdi_gdp	0.01	0.01	0.02	0.01	0.09	0.08	0.00	0.02	0.01	-0.17	-0.14	0.11	0.03	-0.06	-0.04	1.00

Key: Theil's T index (theil), Theil's T index adjusted (theil_adj), logit-transformed Hirschman-Herfindahl index (hhi_t), logit-tranformed Gini coefficient (gini_t), Log GDP per capita (lgdp_pc), Penn World Tables Human Capital Index (pwt_hci), Agricultural exports/GDP (agri_gdp), Mineral exports/GDP (min_gdp), Energy exports/GDP (nrg_gdp), Log population (lpop), Log urban population (lpop_urb), Institutional Quality (inst_pca), Export prices (expo_p), Economic distance (w_dist), Preferencial Trade Access (pta), Foreign Direct Investment/GDP (fdi_gdp).