Policy Research Working Paper 7332

Latent Trade Diversification and Its Relevance for Macroeconomic Stability

Daniel Lederman
Samuel Pienknagura
Diego Rojas

World Bank Group
Latin America and the Caribbean Region
Office of the Chief Economist
June 2015

Policy Research Working Paper 7332

Abstract

Traditional measures of trade diversification only take into account contemporaneous export baskets. These measures fail to capture a country's ability to respond to shocks by allocating factors of production into activities for which it has already paid the fixed costs associated with exporting. This paper corrects for the shortcoming of traditional measures of diversification by introducing a novel measure of trade diversification—latent diversification—and proposes a proxy to measure latent diversification, which is calculated by taking into account the entire history of a country's exports. The paper shows that the observed gaps between traditional measures of diversification and the proposed proxy of latent diversification are sizeable; countries hold latent export baskets that are, on average, three times as large as their average contemporaneous export basket, and these gaps are largest for poor and small countries. Moreover, latent diversification is an important determinant of volatility—more diversified latent export baskets are associated with lower terms of trade volatility and, subsequently, lower GDP per capita volatility, even after controlling for the degree of contemporaneous export diversification and other trade and country characteristics. The latter result, together with the disproportionately large latent baskets relative to contemporaneous baskets observed in poor and small countries, suggests that latent diversification is an important vehicle toward stability in countries that face barriers in building diversified contemporaneous export baskets.

This paper is a product of the Office of the Chief Economist, Latin America and the Caribbean Region. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The authors may be contacted at at dlederman@worldbank.org and spienknagura@worldbank.org.

Latent Trade Diversification and Its Relevance for Macroeconomic Stability

Daniel Lederman, Samuel Pienknagura, and Diego Rojas[1]

JEL Codes: F1, O11

Keywords: Trade Concentration, Diversification, Extensive Margin of Exports, Volatility

---

[1] Daniel Lederman is Lead Economist and Deputy Chief Economist at the Office of the Regional Chief Economist for Latin America and the Caribbean, The World Bank (dlederman@worldbank.org). Samuel Pienknagura is a Research Economist at the Office of the Regional Chief Economist for Latin America and the Caribbean, The World Bank (spienknagura@worldbank.org). Diego Rojas is a Consultant at the Office of the Regional Chief Economist for Latin America and the Caribbean, The World Bank (lrojasalvarado@worldbank.org).
This paper benefitted from useful comments from Augusto de la Torre, Jaime de Pinies, Juan Carlos Hallak, William Maloney, Marcelo Olarreaga, and participants at DEGIT 2014 Conference, LACEA 2014 Annual Meetings, and the Authors' Workshop for the World Bank's Regional Study on Growth and Volatility in the Caribbean. The authors acknowledge financial support from the Office of the Regional Chief Economist for Latin America and the Caribbean. Nicolás Kohn provided stellar research assistance.

1. Introduction

Specialization is a clear-cut prediction of neo-classical models of international trade. Ricardian models based on technological differences across countries predict export specialization, hence concentration, in products in which countries exhibit relative comparative advantage. Heckscher-Ohlin (HO) models based on factor endowments predict export specialization in products that are relatively intensive in factors of production for which an economy has a relative abundance.

These predictions notwithstanding, policy makers and academics alike have highlighted the perils of export concentration. As de Ferranti et al. (2002) note, ''A recurrent preoccupation of [Latin American] policymakers is that their natural riches produce a highly concentrated structure of export revenues, which then leads to economic volatility and lower growth'' (p. 38). The concern among policy makers regarding the link between export concentration and volatility is confirmed by a growing literature that points to a positive correlation between the two (see for instance Jansen (2004), Bachetta et al. (2007), Lederman and Maloney (2012), among others).

Concentration, as usually measured, is also related to underdevelopment. Countries with more concentrated production and trade structures typically have lower income levels compared to countries that are more diversified (Imbs and Wacziarg (2003), Klinger and Lederman (2004, 2006) and Cadot, Carrère, and Strauss-Kahn (2011)).[2]

Despite the large body of work studying the drivers and benefits of diversification, the literature has overlooked an important aspect of diversification—latent diversification.[3] Traditional measures of diversification typically focus on the composition of contemporaneous (or cross-sectional) export baskets. We argue that, to the extent that countries can shift factors of production to export lines not present in contemporaneous export baskets, traditional measures of diversification underestimate a country's ability to respond to exogenous product-specific shocks. Hence, latent diversification is defined as the full set of export lines to which a country can shift factors of production in response to a shock.

---

[2] To be precise, these papers find a U-shaped relation between concentration and income levels, with diversification peaking at around US$20,000 PPP. This implies that income and concentration exhibit a decreasing relationship for income levels below US$20,000 PPP (approximately the 80th percentile of income distribution in 2010).

[3] It should be noted that the literature has studied alternative measures of diversification related to number of markets (see Bacehtta *et al.* (2007) and Brenton and Newfarmer (2007), among others) and quality (see Schott (2004)).

This paper takes this additional dimension of diversification to the data and proposes a proxy for latent trade diversification. Specifically, we construct a measure of diversification based on the total number of products exported by a country throughout a long period of time and document four empirical regularities regarding this measure of latent diversification. First, differences between the proposed measure of latent trade diversification and traditional measures based on contemporaneous export baskets are large, especially for countries with average yearly export baskets that comprise 20-30 percent of the total number of export lines.

Second, while the size of latent export baskets increases with GDP per capita and the size of the labor force, its gap relative to the contemporaneous export basket is decreasing in both GDP per capita and the size of the labor force. In other words, traditional measures of diversification that rely on contemporaneous export baskets tend to understate the ability of small and poor countries to cope with external product-specific shocks. Moreover, these results are robust to the use of estimators that take into account the bounded nature of the measures of diversification used in the paper. Behind this gap is the relatively high turnover in active export lines observed in smaller and poorer countries. The data indicates that, on average, countries with a small size of the labor force and low levels of GDP per capita add and drop products to their contemporaneous export baskets relative to the previous year at a higher rate than bigger, richer countries. Hence, the flexibility of small and poor countries in opening and closing export lines explains the high number of latent export lines relative to their (low) average number of active export lines per year.[4]

Third, this paper documents a negative relationship between the proposed measure of latent export diversification and economic volatility. In particular, more diversified latent export baskets are associated with lower terms of trade volatility and, subsequently, lower GDP per capita volatility. The observed negative relationship remains statistically significant after controlling for other country characteristics such as traditional measures of concentration, the level of GDP per capita, size of the labor force, and concentration in primary activities. This observation suggests that latent diversification in and of itself is an important determinant of macroeconomic volatility across countries.

[4] This complements the findings of Hanson, Lind, and Muendler (2014) who document large turnover rates in export baskets.

Finally, the paper explores one potential determinant of changes in the composition export baskets, namely factor endowments. In particular, we explore sectoral patterns of latent diversification using Leamer's (1995) industry clusters, as a proxy for products with similar production functions in terms of factor proportions, and its link to the initial distribution of products in the contemporaneous export basket across these industry clusters. Our findings suggests that in the short run the distribution across industry clusters of the products added to a country's latent export basket is tightly linked to the distribution of its contemporaneous export basket. However, this link weakens as time passes, suggesting that the influence of initial factor endowments on sectoral patterns of latent diversification dims in the long-run.

The empirical results presented in this paper have important implications for small and poor countries, which on average have higher levels of volatility than richer and larger countries. In particular, size and GDP per capita appear to influence a country's external volatility only through their effects on latent diversification. This, together with the disproportionately large latent baskets relative to contemporaneous baskets observed in poor and small countries, suggests that latent diversification is an important vehicle towards stability in countries that face barriers to building diversified contemporaneous export baskets.

The paper is related to two strands of the trade literature. It closely follows the growing literature on the determinants of trade concentration. For example, Klinger and Lederman (2004, 2006) and Cadot, Carrère, and Strauss-Kahn (2011) study the link between income levels and export concentration. More recently, Agosín, Alvarez and Bravo-Ortega (2012) explore the link between export concentration and a large set of variables (which include trade openness, exchange rate regimes, and financial depth, among others). This paper is also related to the literature exploring the link between export concentration and volatility. For instance, Jansen (2004) and Bachetta et al. (2007) study the link between export concentration and output volatility and find a positive correlation between the two variables. Similarly, Lederman and Maloney (2012) find a positive relation between terms of trade (ToT) volatility and export concentration.

In contrast to the papers mentioned above, our paper studies a measure of diversification that accounts for a country's entire history of exports instead of looking at measures of trade concentration which consider only contemporaneous export baskets. To the best of our knowledge, this type of study has not been explored in the literature.

The rest of the paper is organized as follows. Section 2 describes the data and methodology used throughout the paper. Section 3 describes some of the patterns of trade concentration over the past four decades as a stepping stone to the introduction of our diversification measure. Section 4 introduces the notion of latent trade diversification, and contrasts it to standard diversification measures. Section 5 discusses the role of economic size (measured by size of the labor force) and the level of development (measured by GDP per capita), as determinants of latent trade diversification. Section 6 studies the role of latent trade diversification in determining income and terms of trade volatility. Section 7 explores the role of factor endowments as a possible determinant of changes in latent export baskets. Section 8 concludes.

2. Data and Variables of Interest

Data on trade flows in merchandise goods come from UNCTAD's COMTRADE database. The dataset, which is publically available, reports trade values for exports and imports at the product level for a large set of countries according to eight product classifications. In this paper we use the SITC Revision 1 classification, as it is available for the longest time period (1962-2012).[5] Most of the analysis uses the 5-digit breakdown, which is the most disaggregated data available and defines 944 goods (or export lines).[6]

Data for trade in services, used in section 4, comes from the Consolidated Dataset on International Trade in Services v8.7 described in Francois and Pindyuk (2013). The data is disaggregated at the three-digit Extended Balance of Payments Services (EBOPS) classification. For a detailed discussion about the classification schemes used in the Trade in Services dataset and their associated shortcomings, see Francois and Pindyuk (2013). Finally, data on GDP, population and other country characteristics come from version 8 of the Penn World Tables and the World Bank's World Development Indicators (WDI) dataset.

Throughout the paper we analyze two measures of trade concentration, or equivalently of trade diversification, commonly applied in the literature (see for instance Klinger and Lederman (2006), Cadot, Carrère, and Strauss-Kahn (2009, 2011)). The first measure of concentration we analyze is the number of

[5]SITC Rev. 1 data is chosen over newer revisions or more disaggregated nomenclatures in order to maximize the time span of analysis. This is important since our measure of latent diversification uses a country's history of trade flows. In the analysis, we take all trade values below 1,000 in year 2000 US$ as zero.

[6] Some parts of the paper use more aggregated data to check for robustness.

export (import) lines (the extensive margin of trade). This is our preferred measure of concentration for reasons that are discussed below.

The second measure is the Herfindahl-Hirschman Index (HHI) of concentration, calculated as:

$$HHI_{it} = \sum_k s_{kit}^2 \;;$$

where s_{ikt} is the share of export (import) line k in total export (import) values of country *i* at time t. The index ranges between 0 and 1, with lower values associated with less concentration (more diversification).

Throughout its empirical analysis, this paper uses a sample of countries for which we have data on trade-related variables, ToT volatility, GDP per capita volatility and size of the labor force, and whose ToT growth volatility and GDP per capita growth do not exceed the 95th percentile of the distribution of these volatilities. Hence, the final number of countries used in the empirical analysis in sections 4 through 7 is 95. Table A in Appendix 2 provides summary statistics and data sources for the variables used in the paper.

At the heart of our paper is the idea that measures of concentration/diversification should be calculated not only in a given year, but also for a country's entire history. This dimension of diversification, which we call latent diversification, gives additional information regarding a country's ability to cope with external shocks. Hence, in addition to calculating the number of products for each year, we also calculate the number of products over a period of time, measured as the set of products that a country exports in at least one of the years in the period. We defer the discussion of the conceptual underpinnings of this measure of diversification to section 4.

3. The Evolution of Traditional Measures of Trade Concentration

With a time-invariant trade nomenclature, as explained above, the data reveal three relevant patterns in the evolution of traditional indicators of trade diversification across the world and over time: i) there has been a steady process of diversification in world trade; ii) the process of diversification has been widespread; and iii) despite a move towards more diversification, there is still a large number of countries specialized in relatively few export goods.

Trade diversification stands as one of the most salient patterns of international trade over the last three decades (see Bacchetta et al. (2009)). For instance, the average HHIs for export and import values have

steadily declined since the late 1970s (Figure 1, Panel A).[7] In 2012 the HHI for export values was approximately 30 percent lower than that in 1980; similarly, the HHI for import values declined by approximately 60 percent over the same period.

The same conclusion emerges from the evolution of the number of products exported and imported by the average country. The number of products exported by the average country increased from close to 300 in 1972 to about 500 in 2012. The number of goods imported by the average country rose from approximately 600 in 1972 to about 700 in 2012 (Figure 1, Panel B).

Interestingly, the process of diversification described above is not necessarily concentrated in a handful of countries; on the contrary, it has been extensive and continuous. A look at the cumulative distributions of the number of products exported and imported and the HHI indices highlights the widespread nature of the process of diversification depicted in Figure 1. Approximately 90 percent of the countries in the sample exported less than 600 products per year in 1972 and close to 70 percent of the countries in the sample exported less than 200 goods per year (Figure 2, Panel A). By 2012, the share of countries exporting less than 600 goods per year and less than 200 products per year dropped to 80 percent and 45 percent, respectively. Similarly, the share of countries with HHIs below 0.2 in 1972 was 60 percent; by 2012 that number had increased to 75 percent (Figure 2, Panel B). The same patterns are present when we look at imports (Figure 2, Panels C and D).

Studying the causes of this steady movement towards diversification in detail is an important, albeit challenging, endeavor that stretched beyond the scope of this paper. Moreover, the literature has highlighted a number of potential factors behind trade diversification, ranging from financial and macro variables to trade-related variables.[8] Here we briefly discuss two forces that have received special attention in recent years: a country's level of development, as measured by GDP per capita, and trade reforms.

A growing body of literature examines the relationship between GDP per capita and the process of economic diversification. Imbs and Wacziarg (2003) show that the relationship between sectoral

[7] The HHI is a measure of concentration, represented as an index that ranges from 0 to 1. See Section 2 for a description of the index.

[8] For instance, a recent paper by Agosin, Alvarez and Bravo-Ortega (2012) studies the determinants of trade concentration following a dynamic panel GMM approach. A similar approach is used by Longmore, Jaupart, and Cazorla (2013). Cadot, Carrère, and Strauss-Kahn (2009) provide a detailed review of the literature regarding the determinants of trade diversification.

concentration and GDP per capita follows a U-shaped pattern. This same pattern has also been found in the relationship between export concentration and GDP per capita (see Klinger and Lederman (2004, 2006) and Cadot, Carrère and Strauss-Kahn(2011)). This implies that, to the extent that a country's GDP per capita lies to the left of the turning point, growth in GDP per capita could drive a country towards a more diversified trade structure. Indeed, the estimated values of the turning point reported by Klinger and Lederman (2004, 2006) and Cadot, Carrère and Strauss-Kahn (2011), which stand around PPP US$20,000, suggest that economic growth could lead to a more diversified trade structure in approximately 80 percent of the countries in the world.

To be sure, GDP per capita growth alone cannot explain the dynamics of trade diversification depicted in Figures 1 and 2. In particular, there are mismatches in the evolution of export concentration and the recent episode of high growth observed in developing countries in the 2000s. Indeed, the 2000s saw a period of relative stability in export concentration—the HHIs of exports and imports where relatively flat and the pace at which the average number of exported and imported products increased was more modest than in the 1990s.

It is precisely the timing of the trade diversification process that points to trade liberalization as a major driving force towards expanding into new product lines for export. As shown in Figures 1 and 2, the biggest changes in the averages and cumulative distributions of our measures of diversification occur in the period between 1980 and 1995—a period that was characterized by a dramatic shift towards trade liberalization (Sachs and Warner (1995), Wacziarg and Welch (2008)). According to the index of economic reforms, including trade liberalization, compiled by Sachs and Warner (1995) and updated by Wacziarg and Welch (2008), the share of countries with open trade regimes increased dramatically over two decades; from an average of 27 percent in the 1980s to close to 70 percent in the 1990s. There are in fact a significant number of papers documenting a positive effect of trade liberalization on trade diversification (see Cadot, Carrère and Strauss-Kahn (2009) for a review of the papers). More recently Agosin, Alvarez and Bravo-Ortega (2012) show that, on average, export diversification increases (the HHI decreases) around periods of trade liberalization, especially in the first four years following liberalization. While most of the literature studying the link between trade liberalization and trade diversification does not pretend to have uncovered a causal relationship from liberalization to diversification, it does highlight a tight correlation between these two phenomena. We will come back to discuss the effects of trade liberalization on trade diversification in Section 8.

Despite the widespread movement of world trade towards more diversification, there are numerous countries specialized in a relatively small number of exported goods. Figure 1 shows that the average country exports only half the number of total export goods reported in the SITC Rev. 1 classification. Figure 2 presents a more striking pattern—close to 50 percent of the global sample exports less than 30 percent of goods in the SITC Rev. 1 classification. This is consistent with Hanson (2012) and Hanson, Lind, and Muendler (2014) who argue that there is a tendency for countries to "*hyper-specialize*."

These global trends notwithstanding, there are important differences across indicators of trade concentration. First, imports are more diversified than exports, regardless of the indicator of concentration being used. That is, countries tend to import more goods than they export, and the steady decline observed in traditional measures of import concentration over the last four decades are less pronounced that those observed for exports. For this reason, and to avoid repetition, the rest of the paper will focus on export diversification only.[9]

Second, diversification measured by the HHI of export values is much noisier than diversification measured by the number of products exported, as shown in Figure 1. The differences in trends and variability between the HHI and the number of products exported come from three elements affecting trade flows: the number of products exported; the volumes exported; and the relative prices of these goods. In particular, the extensive margin of exports (the number of products exported) is affected by the first of these three elements while the HHI is affected by all three elements. This explains why concentration, as measured by the HHI, increases in the years prior to the global financial crisis (a period of sharp increases in commodity prices) yet falls when measured by the number of products exported.

Which measure of diversification is a better proxy for a country's ability to buffer shocks to relative prices: the HHI or the number of products? This paper argues that, from an economic point of view, the latter is a better proxy because a country's ability to cope with external shocks is closely related to its ability to

[9] Recent papers on trade diversification and its implications also focus on export diversification only. For example, Klinger and Lederman (2006) study the relation between the HHI of export values and GDP per capita. Similarly, Cadot, Carrère, and Strauss-Kahn (2011) construct three measures of export concentration—the HHI, the Theil entropy index, and the Gini coefficient—and study how they relate to GDP per capita. Hummels and Klenow (2005) study the relationship between the extensive margin of exports, one of our measures of export concentration, and GDP and population. Finally, Bacchetta et al. (2007) and Agosin, Alvarez, and Ortega-Bravo (2012) study the determinants of several measures of export concentration. This does not, however, imply that the potential welfare implications of import diversification should be overlooked (see Cadot, Carrère, and Strauss-Kahn (2009) for a review of the literature on import diversification and its benefits). Rather, from an empirical point of view, concentration is more acute in exports than imports.

shift resources from one export line to another.[10] Clearly, the lower the number of export lines a country has, the lower its ability to adjust to shocks.[11] Hence, most of the sections that follow will focus on studying the extensive margin of exports, namely the number of exported products.

The discussion thus far has focused on measures of diversification that rely on a country's cross-sectional export basket. However, to understand a country's ability to cope with external shocks, one should take into account a country's capacity to shift resources between sectors and products. In contrast to the existing literature which focuses on the contemporaneous export basket, the analysis that follows focuses on a measure of diversification—latent diversification—that has this feature and that is the object of interest of the rest of the paper. We begin in the next section by defining latent diversification conceptually and by studying how it differs from traditional measures of diversification. Sections 5 and 6 discuss how latent diversification relates to economic size and income, and its link to economic volatility.

4. Latent Trade Diversification: Definition, Measurement, and Differences with Traditional Measures of Diversification

Conceptually, latent export lines are goods that have been exported by a country at some point in time or that could be opened in response to external shocks. The former are goods for which exporters in the country have already paid the fixed costs associated with exporting them.[12] These fixed costs range from investments necessary to establish foreign business partners and clients to costs associated with the establishment of appropriate production plants and adapting (training) the labor force. Presumably, a country that has exported a good at some point in time (and hence has paid the fixed cost of exporting) will be more capable of shifting resources towards the production of that good—even if it did not export it in the past several years—than a country that has not exported the good in the past. The other portion

[10] A country's ability to shift to one export line from another is related to two components. On the one hand, the country has to be able to shift production factors across sectors quickly. For example, countries with lower labor mobility costs will be more agile in responding to external shocks (see Artuc, Lederman and Porto (2013) for a detailed analysis and country level estimates of labor mobility costs). On the other hand, a country's ability to shift resources across export lines will depend on the number of export lines available. This paper focuses mainly on the latter component.

[11] Acemoglu and Zilibotti (1997) make a similar argument regarding the number of sectors in an economy. The authors propose a model that explains why poorer countries have more volatile convergence paths. In their model, a country's income level is positively related to the number of sectors operating in the economy. This implies that richer countries are better equipped to cope with idiosyncratic sectorial risks.

[12] The literature analyzes the export decision as a function of both variable and fixed costs. Variable costs, such as transport costs and trade barriers, vary with export volumes. Fixed costs are independent of export volumes and are typically paid before or upon entry into export markets.

of latent export products is comprised of export lines that have not been exported in the past, but could become profitable to export under certain shocks.

The above discussion suggests that countries may have concentrated export baskets in a cross-sectional data sample, as documented by Hanson (2012) and Hanson, Lind, and Muendler (2014), but at the same time can have a diversified latent export basket. This implies that the introduction of the concept of latent export lines poses important challenges to traditional measures of diversification. More precisely, standard measures of export concentration tend to overstate concentration (understate diversification), especially in countries with few export products per year.

To illustrate these points, consider an economy with N latent export lines, but which only exports a subset of these goods each year, with the subset being comprised of N'<N goods. Notice that the subset of goods exported each year could vary; we only assume that the number of goods in the subsets is the same in each year. In this simple example, standard measures of concentration would fail to capture N-N' export lines, which an economy could tap when relative prices are right. The problem of under-estimating diversification could be more acute when the set of latent export lines changes over time. To the extent that N', the number of goods exported by a country per year, grows at a slower pace than N, there is an increasing discrepancy between standard measures of export diversification and measures based on our notion of latent export lines.

Recent literature on international trade presents evidence that supports the view that standard measures of diversification indeed understate export diversification. For instance Hanson and Muendler (2014) estimate a standard gravity model to obtain a measure of export capability. The authors find that there is hyper-specialization across countries (that is, a large degree of concentration in few export goods) but also large turnover among top exported goods. This result would be consistent with our example, i.e. a large gap between N and N'.

Similarly, there is a growing literature which documents changes in relative comparative advantage over time. Levchenko and Zhang (2011) find, for a sample of 75 countries, that revealed comparative advantage (RCA), measured by total factor productivity (TFP) differences, has become weaker over the past five decades. Proudman and Redding (2000) also study RCA dynamics. The authors use Balassa's index of RCA for G-5 economies and find that four of the five economies exhibit a pattern consistent with dynamic

changes in RCA.[13] One implication of these results is that cross-sectional export baskets change as revealed comparative advantage changes and countries start exporting goods that were previously absent in their export baskets. Moreover, to the extent that the loss in comparative advantage in the goods in the "old" set of exported goods is not too large (i.e. these are goods for which the country could regain its comparative advantage in the future), these goods would still be part of the latent export basket. In such case, this would imply N growing faster than N'.

Our empirical approximation of latent export lines takes the total number of goods exported by a country in at least one year over a long period of time (in the case of this paper the 1962-2012, period). Formally, the empirical approximation is

$$N_{c,T}^{L}=\left|\bigcup_{t=t_0}^{T}\{X^{c,t}\}\right| \qquad (1)$$

where $X^{c,t}$ is the set of goods exported by country c in year t, t_0 is the initial year, T is the final year, and $|.|$ is the operator measuring the cardinality of the set.

As mentioned earlier, there are two components of latent export lines which are relevant for analyzing the relationship between our proposed proxy of and the concept of latent diversification. The first comprises export lines already open but not necessarily active (i.e. with positive exports in a given period). The second component comprises export lines that are not open in period *t* but that could be opened in period *t+1*. The latter component is not observed ex-ante, and hence cannot be incorporated in our simple proxy of latent diversification. Moreover, the first component is also difficult to measure accurately, partly because of the finite time dimension of available data, and also because there may be export lines that are obsolete. In other words we have that

$$N_{c,T}^{*L} > N_{c,T} = |\{X^{c,T}\}|$$

$$N_{c,T}^{L}=\left|\bigcup_{t=t0}^{T}\{X^{c,t}\}\right|>N_{c,T}=|\{X^{c,T}\}| \qquad (2)$$

$$N_{c,T}^{*L} \lesseqgtr N_{c,T}^{L} = \left|\bigcup_{t=t0}^{T}\{X^{c,t}\}\right|$$

[13] More precisely, the authors study an econometric model of transition probabilities in the RCA index. They find that France, the US, the UK and Germany have high transition probabilities. Proudman and Redding (2000) also argue that the evidence they find is inconsistent with standard models of international trade with endogenous growth as this model tends to reinforce RCA patterns. To address this problem, Redding (1999) proposes a modified version of a Ricardian model that incorporates learning-by-doing across goods. His model allows for endogenous changes in RCA over time.

where $N_{c,T}^{*L}$ is the conceptual measure of latent export lines and $N_{c,T}$ is the standard cross-sectional indicator of the extensive margin of trade (exports) at time T. The third inequality comes from the fact that $N_{c,T}^{L}$ includes some export lines that are obsolete at time *T* (either due to worldwide technological changes or due to changes in comparative advantage at the country level), hence are outside of the conceptual latent basket, but it excludes inactive, non-obsolete export lines that have not been exported since 1962, and, more importantly, it excludes export lines that have not been exported up to *T* but that could be exported under certain shock realizations. Under the reasonable assumption that the latter export lines exceed the former, something we will assume in what follows, the proposed approximation of the number of latent export lines becomes a lower bound of a country's true latent export basket, which is less biased than the cross-sectional estimate.[14]

Despite the potential bias towards concentration of our measure of latent diversification relative to the theoretical counterpart, the set of products in the former accounts for a large share of the products introduced from year to year by the average country. Out of all the products exported by the average country in year *t* that were not present in the contemporaneous export basket in period *t-1*, between 80 and 90 percent belong to the set of observed latent products in year *t*. Countries tend to re-open observed export lines as external conditions change.

The rest of this section turns to the analysis of the properties of the proposed proxy of latent diversification. First, it compares our proxy to other measures of diversification that rely on cross-sectional export baskets. Then it analyzes the effect on our proxy for the latent export basket of product aggregation and the omission of trade in services data.

4.1. Comparison between Latent Diversification and Traditional Measures of Diversification

The above discussion suggests that there are conceptual reasons why one would expect latent diversification to differ from traditional diversification measures. Moreover, as is argued later in the paper, this difference could have important implications for a country's ability to respond to external shocks. But, is this conceptual possibility present in the data?

[14] Appendix 1 analyzes the composition of the latent export basket by status of activity. That is, it analyzes the fraction of the latent export basket that has been inactive for more than 5 years and 15 years, respectively. The results show that only 20 percent of the average latent export basket in the world does not appear in the cross-sectional export basket for more than 15 consecutive years.

Figure 3 provides a first attempt at answering this question. In particular, it looks at the discrepancy between standard measures of diversification, based on cross-sectional export baskets, and our proposed empirical approximation of latent export baskets by plotting the proposed approximation of the number of latent export lines, $N_{c,T}^{L}$, on the y axis and the average number of goods exported by a country from *t*o until *T*, $\overline{N}_{c,T}$, on the x axis, where we choose *t*o=1972 and *T*=2012.[15] The 45 degree line benchmarks the expected relationship between the two indicators of diversification if a country holds a constant export basket at each point in time.

Figure 3 shows that there is indeed a sizeable difference between traditional measures of diversification and latent diversification, especially for countries with few export products per year. Moreover, the gap between the two measures of diversification is largest for countries with an average of 200 to 300 active export lines per year, numbers that are not far from the number of products exported by the average country in the world identified in Figure 1. In contrast, the bias in favor of concentration present in standard measures of diversification is less severe for countries with large export baskets, giving the concave shape to the curve as we move to the northeast on the graph. The latter observation should not come as a surprise given the nature of the trade data—any empirical measure of the extensive margin of exports will have an upper bound set by the number of product lines in the SITC classification.

The differences between the proposed empirical approximation of latent export lines and standard measures of the extensive margin of exports are further illustrated by the dynamics of the two measures of diversification. Figure 4 plots the coefficients on the year fixed effects obtained from the following regression:

$$y_{c,t} = \theta + \alpha_c + \tau_t + \varepsilon_{c,t} \qquad (3)$$

where $y_{c,t} = \{N_{c,t}^{L}, N_{c,t}\}$, θ is a constant, α_c is a country fixed effect, τ_t is a year effect, and $\varepsilon_{c,t}$ is an error term. The year fixed effects are relative to the averages in 1972.

The evidence presented in Figure 4 is clear—the process of trade diversification depicted by the number of latent export lines is much more dramatic than that presented by standard measures of the extensive margin of exports. On average countries accumulated approximately 350 export lines in their basket of

[15] We decided to start our analysis in 1972, instead of 1962 when trade data is first available, because coverage improves over time. Starting in 1972 gives a sufficiently long time period while ensuring better cross-country coverage.

latent export lines from 1972 to 2012. In contrast, the number of products exported per year by the average country increased by 150 goods from 1972 to 2012.

4.1.1. Broad versus Narrow Sectoral Categories

The results presented thus far rely on export data from COMTRADE using the SITC Rev.1 classification at the 5 digit level. The level of aggregation might affect the difference between standard measures of diversification and our proposed measure of latent diversification. In the extreme, at the highest possible level of aggregation, all countries' export count would be one and thus the variance of both indicators across countries would be wiped out. The question then is, which of the two measures is more sensitive to the level of aggregation in the data?

One can perform two exercises to assess the extent to which this variance-attenuation bias is a function of the level of aggregation of the trade data. First, Figure 3 can be replicated with data at the 1, 2, 3, 4, and 5 digit level of disaggregation to construct a difference between the curve linking latent diversification and cross-sectional diversification and the 45 degree line. This exercise assesses the effect of aggregation on the difference between the two measures of diversification. To make the exercises comparable across aggregation levels, one needs to rescale the two measures of diversification by the maximum number of export goods at each level of disaggregation.

The exercise requires an empirical approximation of the curve linking the two measures of diversification. We obtain this approximation by estimating the polynomial of order 5 that best fits the relationship between the average number of products exported per year and the number of latent export lines in 2012. We then use the predicted polynomial to calculate (twice) the area between the curve and the 45 degree line (the distance of the polynomial with respect to the 45 degree line). Mathematically, the exercise can be represented as:

$$D_{SITC-k} = 2\left(\int_0^1 (p_k(x) - x)dx\right) \qquad (4)$$

where $p_k(x)$ is the predicted polynomial for each level of aggregation k.

The results of the exercise, presented in Panel A of Figure 5, show that the difference in the area below the five polynomials predicted for each level of aggregation is larger when we move from the 1-digit to the 2-digit level of aggregation, and from the 4-digit to the 5-digit level of aggregation. The difference, however, does not vary much for intermediate levels of aggregation. This shows that disaggregation does

appear to diminish latent diversification more than cross-sectional diversification, although the strength of this attenuation problem varies at different levels of aggregation.

A second way to address the variance-attenuation bias induced by aggregation is to compare the cross-country variation of the two measures of diversification—i.e. latent diversification and cross-sectional diversification. This exercise is presented in Figure 5, Panel B.

The evidence presented in Panel B of Figure 5 supports the result found in Panel A of Figure 5—aggregated data reduces the cross-country variation of both measures of diversification. However, the variance-attenuation bias induced by aggregation appears to be stronger for latent diversification than for cross-sectional diversification. These observations have important implications when it comes to empirically assessing the importance of diversification in explaining volatility, something we turn to in Section 6. In particular, the lower variance of latent diversification when using more aggregated data implies that its estimated effect on volatility is expected to decrease with the level of disaggregation. On the other hand, the significance of latent diversification is expected to increase with the level of disaggregation (see Section 6 for more on this).

In sum, the two exercises seem to point in the same direction—more disaggregation attenuates diversification with latent diversification suffering more than cross-sectional diversification. However, it appears that the problem is not necessarily "linear" in the sense that the attenuation problem appears to smaller at intermediate levels of aggregation.

4.1.2. Trade in Services and the Extensive Margin of Trade

A glaring omission in most of the literature on international trade is the exclusion of analysis related to trade in services. In the past there were at least two reasons to exclude these flows. First, with some exceptions, services, which account for approximately 70 percent of world GDP and employment, were typically considered non-tradable goods. However, the rise of multinationals, outsourcing and the reduction of communication costs have made some services highly tradable, thus making this assumption unrealistic. A second reason was the lack of data on trade in services, especially at the sectorial level.

Recent efforts to record international transactions in services have produced new data on international trade in services.[16] One of such dataset appears in Francois and Pindyuk (2013). The authors compiled

[16] There have also been a number of papers written in recent years studying the patterns of international trade in services. See for example Gervais and Jensen (2013) and van Merel and Shepherd (2013).

data on cross-border trade in services from various sources. These data report bilateral trade flows for more than 200 countries in a broad set of activities (according to BOP classifications). The data are not flawless, however. First, not all modalities of trade in services are covered—see Francois and Pindyuk (2013) for a description of all the modalities of trade in services. Second, coverage at the country-product level has improved dramatically over time, which implies that the data quality in the early years is not as strong as in later years.

These caveats aside, we use the data in Francois and Pindyuk (2013) to examine how sensitive the relationship between the two indicators of diversification is to the inclusion of services. To do so, we calculate the measure of distance defined above (i) for the COMTRADE data (merchandise trade) at the 5 digit level of disaggregation; and (ii) for the COMTRADE data merged with the Francois and Pindyuk (2013) data for trade in services. The latter will have the 944 5-digit product classifications defined in the SITC Rev. 1 classification plus the 63 services defined in Francois and Pindyuk (2013). Similar to when we analyzed the effect of the different levels of aggregation on our measures of diversification, we rescale the variables that measure the number of latent products and the average number of products per year so that both variables are comparable between the two datasets. The calculations are made using the period 1995-2012, where the coverage of the trade in services dataset improves substantially.

Figure 6 presents results for the relationship between the number of latent export lines and the average number of export lines per year for both the data including and excluding services. There does not seem to be a disproportionate attenuation problem between the two measures of diversification when using data with and without services; D is less than 5 percent higher when we include services versus when we exclude services (Figure 6). This suggests that leaving out services, as measured by Francois and Pindyuk (2013), does not seem to change the main conclusion of the paper.

5. Latent Diversification, Economic Size, and Economic Development

The previous section proposed a new measure of trade diversification and studied how it compares to traditional diversification measures, such as the number of products exported in a given year. The results suggest that there are indeed large differences between our proposed measure, which we call latent trade diversification, and standard measures. This section expands the comparison between the two diversification measures by analyzing the extent to which country characteristics, such as the size of the

labor force and GDP per capita, correlate with the two measures of diversification and the gap between them.

As alluded to earlier, the influential work by Acemoglu and Zilibotti (1997) and Imbs and Wacziarg (2003) sparked interest in exploring the determinants of diversification. Acemoglu and Zilibotti (1997) studied a dynamic model of country-wide portfolio choice between a safe asset and a set of risky projects. In their model, the full set of risky projects is not available to all countries, as these investment projects are subject to indivisibilities in the form of minimum size requirements. The number of risky projects chosen by a country (hence, its diversification) is positively related to GDP per capita, as richer countries have the capital to pay for the minimum requirements of more projects.[17]

Imbs and Wacziarg (2003) estimate empirically the relationship between sectorial diversification—both in value added and employment—and development. They find a U-shaped pattern between these variables. Countries with levels of GDP per capita below $20,000 (in PPP 2000 US$) diversify as they get richer while countries above this threshold re-concentrate. Interestingly, a similar U-shaped pattern as that found by Imbs and Wacziarg (2003) has been found for the relationship between export concentration and GDP per capita. Moreover, as shown in Klinger and Lederman (2004, 2006) and Cadot, Carrère, and Strauss-Kahn (2011), the threshold level of GDP per capita after which there is export re-concentration is very close to that found in Imbs and Wacziarg (2003). There is also evidence of a positive relationship between firm level concentration (concentration of sales in few firms) and GDP (di Giovanni and Levchenko (2012)).

A second economic characteristic that can affect patterns of diversification is the size of the labor force. Easterly and Kraay (2000) document that small countries are more open than large countries (in terms of the ratio of exports plus imports to GDP), which, according to di Giovanni and Levchenko (2009) and Agosin, Alvarez and Bravo-Ortega (2012), is associated with lower diversification. Lederman and Maloney (2012) explore the relationship between economic size (also measured by the size of a country's labor force) and diversification. The authors simultaneously estimate the determinants of merchandise export-revenue concentration, terms-of-trade volatility, and GDP-per-capita growth volatility by means of a three-stage least squares estimation. They find that size has a negative impact on concentration.

The literature presented above leads to two important questions: Do the factors that affect standard measures of diversification (namely GDP per capita and size of the labor force) also affect latent export

[17] In addition, the higher the number of risky projects a country invests in, the larger the share of savings it is willing to allocate to the portfolio of risky projects.

diversification? And, if so, do they also affect latent export diversification's relationship with cross-sectional diversification? This section sheds some light upon these questions.

We begin by exploring the link between latent export diversification, on the one hand, and GDP per capita and the size of the labor force on the other. In particular, we estimate the following specification:

$$\ln\left(N_{i,2012}^{L}\right) = \beta X_i + u_i, \qquad (5)$$

where, depending on the specification, X_i includes variables such GDP per capita, size of the labor force, the log of the initial size of the cross-sectional export basket, (geographic) trade openness, and net exports of energy and agricultural products per worker.[18] The initial size of the cross-sectional export basket is included in order to capture the effects of size and income per capita on latent diversification beyond those operating through the cross-sectional export basket.

We estimate equation (5) both through OLS and the pseudo-maximum likelihood estimator proposed by Santos Silva, Tenreyro and Wei (2014), hereafter SSTW, which takes into account the potential OLS bias arising from the bounded nature export-product counts. The rest of the exercises described in this and the next section uses a sample of 95 countries for which we have data on trade-related variables, ToT volatility, GDP per capita volatility and size of the labor force. We use the values of variables in X_i in 1980 to reduce the risk of reverse causality. This precaution notwithstanding, the results we present below should not be interpreted as causal relations; they are partial correlations.

Table 1, Panel A, presents the results for the estimation exercise described above. Columns (1)-(6) show that, for the case of the OLS estimation, there is a robust correlation between the log of the latent export basket, on the one hand, and the log of the average initial size of the cross-sectional export basket, GDP per capita, the size of the labor force, and the average net exports of energy and mining per worker, on the other. Similar to what has been found for other measures of diversification, countries with a larger number of people in the labor force and with higher values of GDP per capita tend to have more diversified latent export baskets. This result holds even after controlling for the battery of trade-related variables described above. Similarly, and consistent with Figure 3, Panel A of Table 1 documents a positive correlation between latent export diversification and cross-sectional diversification. Interestingly, the results suggest that countries with a higher incidence of energy and mineral exports, which typically have

[18] Geographic trade openness is calculated as in Frankel and Romer (1999), who calculate the geographic component of trade derived from a gravity estimation. Net exports of energy and agricultural products per worker are calculated as in Lederman and Maloney (2012).

more concentrated contemporaneous export baskets than other countries, have larger latent export baskets once we control for cross-sectional diversification. This already points to the potential role of latent export diversification as way for countries with certain economic structures to cope with external volatility.

Columns (7)-(12) of Table 1, Panel A, presents the results for the estimation of equation (5) using the method proposed by SSTW. The results found using this alternative estimation broadly resemble those found using OLS. However, the correlations appear less robust once we take into account the truncated nature of the data, especially in the case of the initial size of the cross-sectional export basket. In particular, the size of the labor force and GDP per capita remain significant in four of the specifications in which they are included (6 and 5, respectively) while the size of the average cross-sectional basket is significant in 3 of the six specifications in which it is included. In contrast to the OLS case, net exports of minerals and energy per worker stops being significant in all of the specifications in which it is included. Finally, the difference in the magnitude between the SSTW estimates and the OLS estimates suggest that truncation does affect the estimated impact of our variables of interest and latent diversification. However, as was discussed above, the statistical significance of size and income per capita does not appear to be driven entirely by the truncated nature of latent diversification.

The results in Columns (1)-(12) of Table 1, Panel A, point to a tight link between latent diversification, size and GDP per capita, even after controlling for measures of cross-sectional diversification. This, together with the results in the literature, suggests that larger and richer countries are better equipped to diversify their export baskets. However, as shown in Figure 3, countries with small cross-sectional export baskets tend to have large latent export baskets relative to their cross-sectional baskets. The question we turn now to is: Do variables like size and income play a role in explaining the differences between latent diversification and cross-sectional diversification documented in Figure 3?

To address this question we estimate an equation similar to (5), where we replace the log of the latent export basket with the difference between the log of the latent export basket and the log of the cross-sectional export basket.[19] With the exception of the log of the initial size of the cross-sectional export basket, we use the same controls and sample as in equation (5). Once again, we apply both OLS and SSTW methodologies to estimate this equation for the reasons discussed above.

[19] In this analysis we use the size of the average cross-sectional export basket from 1962 to 2012.

The results, presented in Panel B of Table 1, show that the gap between cross-sectional and latent export baskets is correlated both to GDP per capita and the size of the labor force, and this is robust to controlling for other trade-related variables. In particular, the gap between the two measures of diversification is increasing in both the size of the labor force and GDP per capita (see columns (1)-(6) of Table 1, Panel B); that is, poorer and smaller countries have latent export baskets that are relatively bigger than their cross-sectional basket. Importantly, this relationship does not seem to be driven by the truncated nature of the data; columns (7)-(12) show that the correlation between the gap between latent and cross-sectional diversification, on the one hand, and GDP per capita and the size of the labor force, remains significant when we estimate the relationship using the SSTW methodology.

In sum, the results in Table 1 suggest that, while poorer, smaller countries have less diversified latent and cross-sectional export baskets, they have on average larger latent export baskets relative to their cross-sectional baskets than richer and bigger countries. This suggests that small and poor countries partially overcome their inability to build diversified cross-sectional export baskets by holding diversified latent export baskets.

5.1. Open and Closed Export Lines, Labor Force and GDP per capita

What explains the finding that smaller and poorer countries have a higher number of latent export lines (relative to their average number of export lines per year) than bigger, richer countries? One possible explanation is that such countries have a larger turnover in their cross-sectional export baskets.

We run four sets of regressions to assess the likelihood of this explanation: one for the number of export lines opened relative to *t-1* as a share of the export basket in year *t*; one for the number of export lines closed relative to *t-1* as a share of the size of the export basket in year *t-1*; one for the number of export lines opened in *t* relative to the initial year as a share of the size of the export basket in *t*, and one for the number of export lines closed relative to the initial year as a share of the size of the export basket in initial year.[20] We assess the role of economic size and income levels on the rate at which countries open and close export lines through three specifications—(i) one which controls for the log of the size of the labor

[20] In the first set of regressions we define an opened export line relative to t-1 as an export line that is present in the contemporaneous export basket in period t and was not present in t-1. In the second set of regressions we define a closed export line relative to t-1 as an export line that is present in the contemporaneous export basket in period t-1 and is not present in t. In the third set of regressions we define an opened export line relative to the initial year as an export line that is present in period t and was not present in the initial year. In the fourth set of regressions we define a closed export line relative to the initial year as an export line that is present in the initial year and is not present in t.

force in 1980, (ii) one which controls for the log of GDP per capita in 1980, and (iii) one which controls for both variables and year fixed effects simultaneously. In each of these regressions we control for year fixed effects. Similar to when estimating equation (5), we estimate these specifications both through OLS and through the pseudo maximum likelihood method proposed by SSTW, which takes into account the truncated nature of the left hand side variable. The results of the estimation of the four sets of regressions are presented in Table 2.

The results presented in Table 2 confirm the hypothesis set forth at the beginning of this subsection—small and poor countries have higher turnover in their export baskets, as they open and close export lines at a higher rate. In particular, columns (1)-(3) and (7)-(9) of Table 2, Panel A show the OLS estimates of the creation and destruction rates (opened export lines relative to the size of the export basket in *t* and closed export lines relative to the size of the export basket in *t-1*), respectively. The coefficients for size and income per capita are negative in each of these columns, which implies that small and poor countries present high year-to-year turnover in their export baskets. Clearly, the results are partly driven by the bounded nature of creation and destruction rates; columns (4)-(6) and (10)-(12) present the SSTW estimates, which have a smaller magnitude compared to the OLS estimates. However, the effect of size and GDP per capita remains significant under the SSTW estimation. Indeed, except for the coefficient on size in the regression for the creation rate (column (4)), all of the coefficients are statistically significant.

The results for creation and destruction relative to the initial export basket, presented in Panel B of Table 2, reinforce the conclusion that small and poor countries have higher rates of creation and destruction relative to the initial export basket than bigger and richer countries. Overall, the results in this subsection suggest that the disproportionately large latent basket relative to cross-sectional baskets observed in poor and small countries is indeed a byproduct of the large turnover observed in their cross-sectional baskets.

6. Latent Trade Diversification and Macroeconomic Volatility

Concentration has been deemed by academics and policy makers as a possible cause of economic volatility. For example, Acemoglu and Zilibotti (1997) and Imbs and Wacziarg (2003) highlight, theoretically and empirically, that poorer countries, which on average are more volatile, have more concentrated economic structures. More recently, di Giovanni and Levchenko (2012) show that poorer countries have more concentrated firm size distributions. Under certain assumptions, this implies that they are more vulnerable to the amplification of shocks hitting big firms.

The literature has paid special attention to the relationship between export concentration and income volatility. Jansen (2004) and Bachetta et al. (2007) show that countries with more diversified export baskets (measured by the HHI) have lower output volatility. They also show that export diversification at the product level (as opposed to the destination level) is a better conduit to mitigate output volatility for poorer economies. As mentioned in section 5, the literature has also stressed the fact that smaller and poorer countries, which are on average more volatile, tend to be more concentrated than bigger and richer countries (see Klinger and Lederman (2004, 2006), Hummels and Klenow (2005), and Cadot, Carrère and Strauss-Khan (2009, 2011)).

Lederman and Maloney (2012) study a specific channel through which export concentration affects income volatility, namely through terms of trade (ToT) volatility. They estimate a structural model where export concentration determines ToT volatility, which, in turn, determines GDP per capita growth volatility. The authors confirm that higher levels of export concentration have a positive and statistically significant effect on ToT volatility, and that ToT volatility has a positive and significant effect on income growth volatility.

We expand this literature by studying the connection between our measure of latent export diversification and volatility. Figure 7 provides a preliminary look at this relationship. It plots the ratio of the number of latent export lines in 2012 to the total number of export lines defined in the SITC Rev. 1, 5-digit classification, against the standard deviations of ToT growth (Panel A) and GDP per capita growth (Panel B) for the period 1980-2012. We observe a negative correlation between the number of latent export lines and the two measures of volatility.

To be sure, these negative correlations can be attributed to other factors that are positively correlated with volatility and our measure of latent diversification. For instance, section 4 documented a positive correlation between our measure of latent export diversification and standard measures of the extensive margin which, as Panels C and D in Figure 7 show, are also correlated with volatility.[21]

To address this concern, Table 3 presents estimated coefficients of a multivariate econometric model linking volatility of ToT and GDP per capita growth, captured by the standard deviation between 1980 and 2012, with export diversification (both latent and cross sectional) and other country characteristics that may affect volatility. Among the country characteristics included are: the log of the initial level of

[21] This is in line with Lederman and Maloney (2012), who document a positive correlation between the HHI of export concentration and ToT volatility.

development (GDP per capita in 1980); the log of the size of the labor force in 1980; and net exports of minerals and agricultural products per worker.[22] We use the levels of GDP per capita and the size of the labor force in 1980 to minimize the risks of reverse causality between the right hand side variables and volatility. We use the log of the number of latent export lines in 2012 to proxy latent export diversification. Similarly, we use the log of the average number of products exported per year in the 1962-1980 period to proxy for the size of the initial cross-sectional export basket. We choose the initial value of our measure of cross-sectional diversification (as opposed to the average in the period that goes until 2012) to minimize the risk of colinearity between our two measures of diversification. Finally, as discussed above, the sample we use includes 95 countries for which we have trade, economic, and demographic data.

The estimates in columns 1 and 2 of Table 3, Panel A, confirm the negative correlation between latent diversification and ToT volatility found in Figure 7. Latent and cross-sectional diversification are negatively correlated with ToT growth volatility, and the estimated coefficients are statistically significant. Importantly, column 2 of Panel A shows that the negative correlation between ToT volatility and latent export diversification survives the inclusion of cross-sectional diversification, although the magnitudes of the coefficients are smaller.

It is important to note that the negative correlation between ToT volatility and cross-sectional and latent diversification can be the result of omitting other country characteristics linked to volatility and correlated with diversification. For instance, as highlighted above, volatility has been found to be negatively correlated to size and income, which, as shown in Table 1, are positively correlated with latent diversification. Yet, columns (3) and (4) of Table 3, Panel A show that the effect of latent diversification on ToT volatility is unaffected by the inclusion of size and income as control variables. Also, the results in these columns show that size and income have small point estimates, which are not statistically significant, and that the effect of initial cross-sectional diversification loses its significance.

We explore the sensitivity of the results to the inclusion of other trade characteristics in columns (5)-(7). In particular, we control for the role of natural resources, measured by net exports of energy and mining and agricultural exports per worker. First, some argue that commodity prices are more volatile than those of other products, which would imply that exporters of primary goods might face high ToT volatility.[23] In

[22] Alternative specifications which included additional controls (trade openness and financial depth) did not yield much change in the coefficients of trade diversification.
[23] To be sure, there is some evidence that commodity prices are not necessarily more volatile than the prices of manufactured-products (Arezki, Hongyan and Lederman (2014)).

addition, there is evidence that exports of primary products have an effect on export diversification (Lederman and Maloney (2012)). This implies that the results in columns (1)-(4) could be driven by the omission of measures of export concentration in the primary sector.

The results in columns (5)-(7) show that the effect of diversification on ToT volatility remains significant even after controlling for the cross-sectional measure of diversification and other trade characteristics. This indicates that the effect of latent diversification on ToT volatility goes beyond the omission of a country's exports of primary goods. Interestingly, latent diversification remains the only variable with a significant coefficient, despite a decline in the magnitude of its coefficients.

In contrast, Panel B of Table 3 suggests that the link between latent diversification and the volatility of GDP per capita growth is less robust than its link with ToT volatility. Indeed, the coefficient of (the log of) latent diversification is negative and remains significant even after controlling for ToT volatility, the (log of the) initial size of the labor force, and the log of initial GDP per capita (columns (2)-(4)). However, the effect of latent diversification loses its statistical significance once we include other trade and economic characteristics (columns (5)-(8)).

Two additional observations from Table 3, Panel B, are worth mentioning. First, there is a tight and robust link between GDP per capita growth volatility and ToT growth volatility (as indicated by a statistically significant coefficient on ToT volatility in all of the estimated specifications). Second, the coefficient on the size of the labor force is negative and significant in some specifications estimating the determinants of GDP per capita growth volatility. This result is consistent with the findings in Lederman and Maloney (2012).

In sum, Table 3 suggests that latent diversification is a robust determinant of ToT growth volatility and, through it, of GDP per capita growth volatility. In section 6.2 we investigate the plausibility of this channel in more detail through a three-stage least squares estimation (3SLS) that uses the results of Table 3. But first, we address the potential effect of aggregation on the link between latent diversification and volatility.

6.1. Latent Diversification, Aggregation and Volatility

As discussed in Section 4.1, the level of aggregation of trade data affects our measures of cross-sectional and latent diversification, which in turn affects the econometric exercise presented above. In particular, higher levels of aggregation of trade data impact the variables entering the regression in Table 3 in two

ways: i) the variance of both measures of diversification across countries declines with the level of aggregation; and ii) the colinearity between the measures of diversification rises with the level of aggregation (see section 4.1 for a discussion). Given the above, do changes in aggregation affect the coefficient of latent diversification in the regressions presented in Table 3?[24]

To address this question we re-estimate the specifications in columns (1)-(7) of Table 3, Panel A using measures of diversification calculated with the SITC Rev. 1 2-digit, 3-digit and 5-digit classifications. The results of this exercise are presented in Figure 8.

The coefficients on our measure of latent diversification tend to be more precisely estimated when using data at higher levels of disaggregation. This is partly a result of the attenuation of variance across countries. The reduction in statistical significance of latent diversification estimates in specifications where we also control for cross-sectional diversification is larger when we use more aggregated data. This is partly a result of the large degree of colinearity between cross-sectional and latent diversification when using aggregated data. Likewise, Figure 8 shows that in most cases more aggregated data is associated with larger point estimates. This is also partly due to lower cross-country variance, because the coefficient is inversely proportional to the variance of the explanatory variables.

6.2. Latent Diversification, Terms-of-Trade Volatility and GDP Volatility

The estimation results in Table 3 highlight the link between latent diversification and terms-of-trade volatility. Latent diversification is negatively correlated with ToT volatility (but not necessarily with GDP per capita volatility), whereas economic size and income per capita do not necessarily have a direct effect on either volatility measure. To explore these links further, we present the results of an econometric exercise that simultaneously explores the determinants of latent diversification, terms-of-trade volatility, and GDP per capita growth volatility. In the model, latent diversification affects GDP per capita growth volatility through its effect on terms-of-trade volatility. We estimate these relationships through a three-stage least squares estimation (3SLS). The results are presented in Table 4.

The model of latent diversification is consistent with the results obtained throughout this paper—bigger and richer countries typically have larger latent export baskets. In addition, a higher prevalence of commodity exports does not have a statistically significant impact on latent diversification. Regarding

[24] We focus our analysis exclusively on the link between latent diversification and ToT growth volatility. This choice was made given the fact that the coefficient for latent diversification in the GDP per capita growth volatility equation is significant only in a limited number of cases and, as is discussed in section 4.1, this expected to hold for higher levels of aggregation.

volatility, latent diversification has a negative impact on ToT volatility, as expected from Table 3, and ToT volatility increases GDP per capita growth volatility.

The estimates presented in Table 4 suggest a large impact of latent diversification on the volatility of GDP per capita growth. An increase from the 25th to the 50th percentile of the latent export basket distribution in our restricted sample of 95 countries reduces ToT growth volatility by 1.37 percentage points, or one fourth of the average standard deviation in the sample. In turn, this implies a 0.8 percentage point reduction in the volatility of GDP per capita growth.

7. Sectoral Distribution of Latent Export Baskets: The Role of Factor Endowments

Thus far the analysis has highlighted the importance of latent export baskets in mitigating external volatility. This section provides preliminary evidence on another important policy question—which types of goods does a country add to its latent export basket over time?

To answer this question we explore the importance of the cross-sectoral composition of cross-sectional export baskets in a given point in time on the cross-sectoral composition of the goods added to the latent export basket from that point onwards. In particular, we focus on the role of initial factor endowments as a potential determinant of cross-sectional and latent export baskets. We do so by splitting a country's export basket into the ten broad industry clusters proposed by Leamer (1995), which group products according to factor intensities,[25] and analyze the relationship between the share of goods in the cross-sectional export basket belonging to an industry cluster and the share of goods added to the latent export basket belonging to that industry cluster. To avoid the problem of truncation present in the trade data, the analysis focuses on the four industry clusters with the highest number of export lines in them, namely labor intensive goods, capital intensive goods, chemicals and machinery.

We begin our analysis by estimating for each year from 1980 to 2005 the following equation:

$$S^{L}_{c,j,t-2012} = \alpha_j + \theta S^{c}_{c,j,t} + \varepsilon_{c,t} \qquad (6)$$

[25] Leamer (1995) proposes a set of 10 industrial clusters. The distinguishing feature of each of these clusters is that the products included in each tend to be exported by similar countries in terms of their endowments of labor, land and natural resources. The ten broad clusters are: petroleum; raw materials; forest products; tropical agriculture; animal products; cereals; labor intensive goods; capital intensive goods; machinery; and chemicals.

where $S^{C}_{c,j,t}$ is the share of goods in country *c*'s cross-sectional export basket in year *t* that belongs to industry cluster *j*, $S^{L}_{c,j,t-2012}$ is the share of goods added to country c's latent export basket from year *t* until year 2012 that belongs to industry cluster *j*, and α_j is an industry fixed effect.

Notice that the analysis described above can be interpreted as a test of the importance of factor endowment-induced comparative advantages on latent export baskets, in the spirit of the Heckscher-Ohlin model. In such a world, $S^{C}_{c,j,t}$ is a proxy for country *c*'s comparative advantage in industry *j*. This pattern of comparative advantage holds over time to the extent that country *c*'s ranking in terms of factor endowments remains constant over time and to the extent that there are no factor intensity reversals.

Figure 9 illustrates the results of the estimation of equation 6. There are two salient points stemming from Figure 9. First, the persistence of initial cross-sectional industry shares is statistically different from zero regardless of the timeframe analyzed. However, the mapping from cross-sectional shares to latent export shares is not one-to-one—the coefficient for the cross-sectional share (θ in equation (6)) hovers around 0.4 and 0.6, which implies that 1 percentage points increase in the share of sector *j* in the cross-sectional basket translates approximately into a 0.5 percentage points increase in the share of goods added to the latent basket belonging to sector *j*. The second point worth highlighting is that the coefficient of persistence declines as the time period analyzed increases. We will return to this result later on this section

To be sure, Figure 9 masks potential differences in the persistence coefficient across industry clusters. We explore this potential heterogeneity by estimating versions of equation (6) on an industry by industry basis. The results of this exercise, which are depicted in Figure 10, are twofold. First, there is indeed heterogeneity in the persistence coefficient across industry clusters. Chemicals displays the highest persistence levels both in short and in long time windows, while capital intensive goods display a relatively low persistence. Machinery displays an interesting pattern—it's the industry cluster with the lowest persistence in the long-run but is one of the industry clusters with the highest persistence terms in the short-run. A second point emerging from Figure 10 is that, in line with the results presented in Figure 9, all industry clusters experience increases in persistence as the time horizon analyzed shrinks.

Coming back to the relationship between persistence and the time window, we explore two potential hypothesis for such pattern. One possibility is that this pattern reflects the relative importance of initial factor endowments over time—one could expect the degree of factor-endowment dependence of the goods added to the latent export basket to be higher in the short-run and to dim in the long-run. Yet, the specific timing of the increase in persistence in all sectors, which occurs somewhere around the mid-1990s, opens the door to another potential explanation. As was highlighted in Section 3, trade liberalization was an important force behind the strong process of export diversification observed around the world. Nevertheless, trade liberalization could either ignite higher or lower persistence. On the one hand, trade openness could lead to a stronger prevalence of factor endowments on export baskets (i.e. the specialization patterns predicted by the HO model). On the other hand, trade liberalization removed some of the distortions to export decisions introduced by import substitution strategies present in the 1980s, hence, rebalancing export baskets and breaking the link between initial shares and shares in the latent export basket.

To disentangle between these two hypothesis (i.e. a fall in the relative importance of initial factor endowments over time and a change in persistence due to trade liberalization) we perform an additional exercise. In particular, we estimate an equation similar to (6) where instead of having a shrinking time window, we keep the time window fixed at seven years and analyze how the persistence of initial cross-sectional industry shares on the share of goods added to the latent export basket in those seven years in an industry varies over time. Similar to the exercise summarized in Figure 9, we estimate (6) as a panel with industry fixed effects.

The results of this exercise, presented in Figure 11, show that the persistence coefficient in the 7 year fixed window exercise is relatively stable in the 1980s, falls in the 1990s, and then increases in the 2000s. Moreover, the increase in the persistence in this latter exercise coincides with the increase in the persistence observed in the shrinking window exercise presented in Figure 9. For this reason, and to get a clearer picture of which of these two forces dominates the results in Figure 9, we calculate the difference in each year between the persistence coefficients obtained in the two exercises, which captures the effect fall in the relative importance of initial factor endowments over time. This difference, which is plotted in Figure 11 alongside the persistence coefficients of the two exercises, supports the initial conjecture that the relative importance of initial factor endowments falls as the window of time analyzed increases—the degree of factor-endowment dependence of the goods added to the latent export basket is higher in the short-run and dims in the long-run.

The results in this section reinforces the potential of latent export baskets in providing countries with a tool to cope with external volatility. Indeed, while countries tend to open new export lines in industries in which they had a factor-endowment driven comparative advantage in the short-run, this link weakens in the long-run. Hence, over long-periods of times the industry composition of initial export baskets may differ substantially from the industry composition of the goods added to the latent export basket, which in turn results in latent basket that are more evenly distributed across industry clusters.

8. Conclusion

This paper introduces the concept of latent trade diversification, which measures trade diversification based on the number of export lines that have been active in a given time period (instead of currently active only). Differences in diversification across countries are much smaller when looking at latent diversification than when looking at cross-sectional diversification. Moreover, latent diversification has a significant and economically meaningful impact on volatility in terms-of-trade and indirectly on GDP per capita growth. These results are robust to the estimation of numerous variations of our baseline specifications.

Our results open the door to further research on the concept of latent diversification and its implications for volatility. This paper presents a highly stylized and simple measure of latent diversification, which is easily implementable. This simplicity notwithstanding, the measure has some limitations. It only takes into account products that have been exported within a finite time frame, which may be arbitrary depending on data availability to researchers. This implies that the proposed measure of latent diversification might miss products that are in the latent basket but that have not been exported in a country's history within the data sample.

More importantly, the findings call for an explanation of changes in an economy's basket of exported products as the paper took a shortcut in this regards. We utilized Leamer's (1995) industry clusters as a proxy for products with similar production functions in terms of factor proportions. This highlights the connection between this paper and the literature on dynamic comparative advantage. Future research could aim to parameterize Ricardian and factor-proportions models of comparative advantage enhance our understanding of the forces that could explain the relationship between the products in an economy's latent export basket and the basket observed in a given period of time.

References

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

Arezki, R., D. Lederman, H. Zhao. 2014. "The Relative Volatility of Commodity Prices: A Reappraisal," American Journal of Agricultural Economics vol. 96: 939-951.

Artuc, E., D. Lederman, and G. Porto. 2013. "A Mapping of Labor Mobility Costs in the Developing World," Policy Research Working Paper 6556, The World Bank.

Bacchetta, M., M. Jansen, R. Piermartini, and R. Amurgo-Pacheco. 2007. "Export Diversification as an Absorber of External Shocks," WTO mimeo.

Brenton, P. and R. Newfarmer. 2007. "Watching More than the Discovery Channel: Export Cycles and Diversification in Development," Policy Research Working Paper 4302, The World Bank .

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

Cadot, O., C. Carrère, and V. Strauss-Kahn. 2009. "Trade Diversification, Income, and Growth: What Do We Know?", CERDI Working Paper 2009.31 (Clermont-Ferrand: Centre d'Etudes et de Recherches sur le Developpement International).

de Ferranti, D., G. Perry, D. Lederman, and W. Maloney. 2002. *From Natural Resources to the Knowledge Economy: Trade and Job Quality*. Washington, DC: World Bank.

di Giovanni, J. and A. Levchenko. 2009. "Trade Openness and Volatility," Review of Economics and Statistics, vol. 91 (3), pp. 558-585.

di Giovanni, J. and A. Levchenko. 2012. "Country Size, International Trade, and Aggregate Fluctuations in Granular Economies," Journal of Political Economy, vol. 120 (6), pp. 1083-1132.

Easterly, W and A. Kraay, 2000, "Small States, Small Problems? Income, Growth, and Volatility in Small States," World Development, 28(11), pp. 2013-27.

Gervais, A. and J. B. Jansen. 2013. "The Tradability of Services: Geographic Concentration and Trade Costs," NBER working paper 19759.

Hanson, G. H. 2012. "The Rise of Middle Kingdoms: Emerging Economies in Global Trade." Journal of Economic Perspectives, 26(2), pp. 41-64.

Hanson, G. H., N. Lind, and M. Muendler. 2014. "The Dynamics of Comparative Advantage," Unpublished manuscript.

Hummels, D. and P. Klenow. 2005. "The Variety and Quality of a Nation's Exports" American Economic Review vol. 95, pp. 704-723.

Imbs, J. and R. Wacziarg. 2003. "Stages of Diversification." American Economic Review, March, 93(1), pp. 63-86.

Jansen, M. 2004. "Income Volatility in Small and Developing Economies: Export Concentration Matters." World Trade Organization WP.

Klinger, B., and D. Lederman. 2004. ''Discovery and Development: An Empirical Exploration of 'New' Products,'' Policy Research Working Paper Series 3450, The World Bank.

Klinger, B. and D. Lederman. 2006. "Diversification, innovation, and imitation inside the Global Technological Frontier," Policy Research Working Paper Series 3872, The World Bank.

Leamer, E. 1995. "The Heckscher-Ohlin Model in Theory and Practice," Graham Lecture, Princeton Studies in International Finance, No. 77, February 1995.

Lederman, D. and W F. Maloney. 2012. *Does What You Export Matter? In Search of Empirical Guidance for Industrial Policies*. Latin American Devolopment Series. Washington DC: World Bank.

Levchenko, A. and J. Zhang. 2011. "The Evolution of Comparative Advantage: Measurement and Welfare Implications," NBER Working Paper 16806.

Longmore, R., Jaupart, P., and M. R. Cazorla. 2014. "Toward Economic Diversification in Trinidad and Tobago," Policy Research Working Paper 6840, The World Bank.

Proudman, J. and S. Redding. 2000. "Evolving Patterns of International Trade," Review of International Economics, vol. 8(3), pp. 373-396.

Redding, S. 1999. "Dynamic Comparative Advantage and the Welfare Effects of Trade," Oxford Economic Papers, vol. 51, pp. 15-39.

Sach, J. and A. Warner. 1995. "Economic Reform and the Process of Global Integration," Brookings Papers on Economic Activity, vol. (26), pp. 1-118.

Santos Silva, J.M.C., S. Tenreyro, K. Wei. 2014. "Estimating the Extensive Margin of Trade," Journal of International Economics, forthcoming.

Schott, P. 2004. "Across-Product versus within Product Specialization in International Trade", Quarterly Journal of Economics, vol. 119(2), pp. 647-678.

Wacziarg, R. and K. Welch. 2008. "Trade Liberalization and Growth: New Evidence," World Bank Economic Review, vol. 22 (2), pp. 187-231.

Appendix 1. Composition of the Latent Export Basket by Status of Activity of its Products

One of the assumptions made in the paper is that products inside a country's latent export basket are more likely to be exported in a given year than those outside of their export basket, even if they were not exported in the previous years. However, there may be reason to question whether the likelihood of a country exporting a latent export good in a given year is inversely related to the time that the good has being inactive. This could happen, for instance, if a country's know-how of an inactive export product depreciates over time or if technological change makes some export products obsolete.

To address this concern, we divide products within a country's export basket by status of "activity." In particular, we split the latent export basket in each year into three groups. The first comprises export products that are actively being exported in that year. The second group consists of export goods that have been exported in at least one year out of the last x years ($x \in \{5,15\}$). The third group is comprised of goods that have not been exported in the last x years ($x \in \{5,15\}$). The underlying assumption in this exercise is that x captures the lifespan of a product once it's inactive. We focus the discussion on the results for the world average.[26] The results of the exercises are depicted in Figure A2.1.

The decomposition of the latent export basket by activity status suggests that, in the most recent years, cross-sectional export baskets account for between 50 and 60 percent of the average country's latent export basket. The composition of the 40-50 percent residual share depends on our assumption regarding the rate of decay of inactive export lines. The most stringent case of our two exercises, the one with $x = 5$ (Panel A in Figure 8) shows that the residual share is split evenly between products inside and outside the interval. This suggests that, under this assumption, approximately 20 to 25 percent of the products initially included in a country's export basket do not buffer against external shocks. Obviously, when $x = 15$ the share of "obsolete" export lines, that is products that have been inactive for more than 15 years, falls to approximately 15 percent of the average latent export basket.

[26] We also performed a similar exercise splitting countries into four groups according to two variables: a) their level of income per capita in 1980 and b) the size of their labor force in 1980. The four groups were determined using the quartiles of the distribution of each of the two variables in 1980. While the levels varied by group, the qualitative results across groups were similar.

The above suggests that our measure of latent diversification may overestimate a country's ability to use export goods within their latent export basket to overcome negative shocks, especially for small and poor countries. Nevertheless, the severity of this problem will depend on how quickly goods in the latent basket lose value for these countries.

Appendix 2. Description of the Data

Table A presents summary statistics of the main variables used in the empirical analysis in the paper along with the sources of the data. The number of observations used throughout the paper reflects the inclusion of countries for which we have data for all the variables used in the analysis.

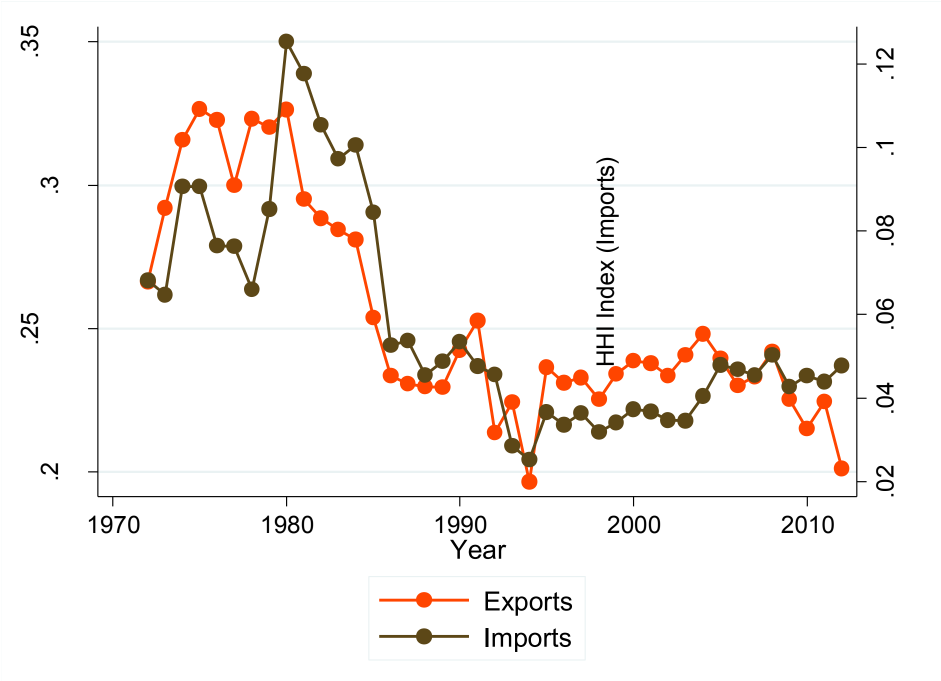
Table A. Description of Main Variables

	N	mean	p50	st. dev.	min	max	
Log of GDP per capita in 1980	95	8.460	8.425	1.187	6.187	11.474	PWT
Log of people engaged in 1980	95	0.808	0.888	1.878	-3.400	5.517	PWT
Latent Export Basket, 1980-2012	95	645.126	746.000	241.826	12.000	901.000	COMTRADE
Size of the Average Cross-Sectional Export Basket, 1980-2012	95	394.055	365.276	272.108	4.500	800.303	COMTRADE
Size of the Average Cross-Sectional Export Basket, 1962-1980	95	270.215	169.133	272.094	1.000	913.389	COMTRADE
S.D of GDP per worker growth rate, 1980-2012	95	0.061	0.057	0.032	0.021	0.155	PWT
S.D of terms of trade, 1980-2012	95	0.060	0.055	0.029	0.020	0.153	PWT
Net exports of Mining and Energy per capita before 1980	95	-14.388	-0.549	259.734	-2,357.332	655.509	Lederman and Maloney (2012)
Net exports of Agricultural goods per capita before 1980	95	-0.632	-0.156	9.954	-29.233	53.924	Lederman and Maloney (2012)
Geographical trade openness over GDP	95	24.871	18.470	20.057	2.560	98.140	Lederman and Maloney (2012)

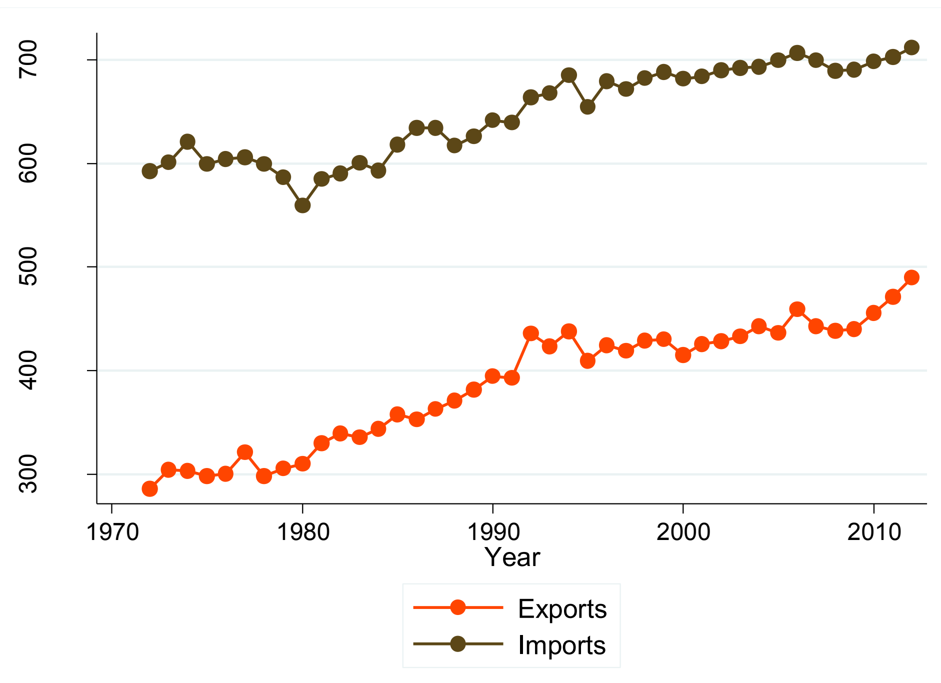
Appendix 3. Figures and Tables

Figure 1. Evolution of Trade Concentration

Panel A. Average Herfindahl-Hirschman Index of Export and Import Values, by Year



Panel B. Average Number of Goods Exported and Imported, by Year

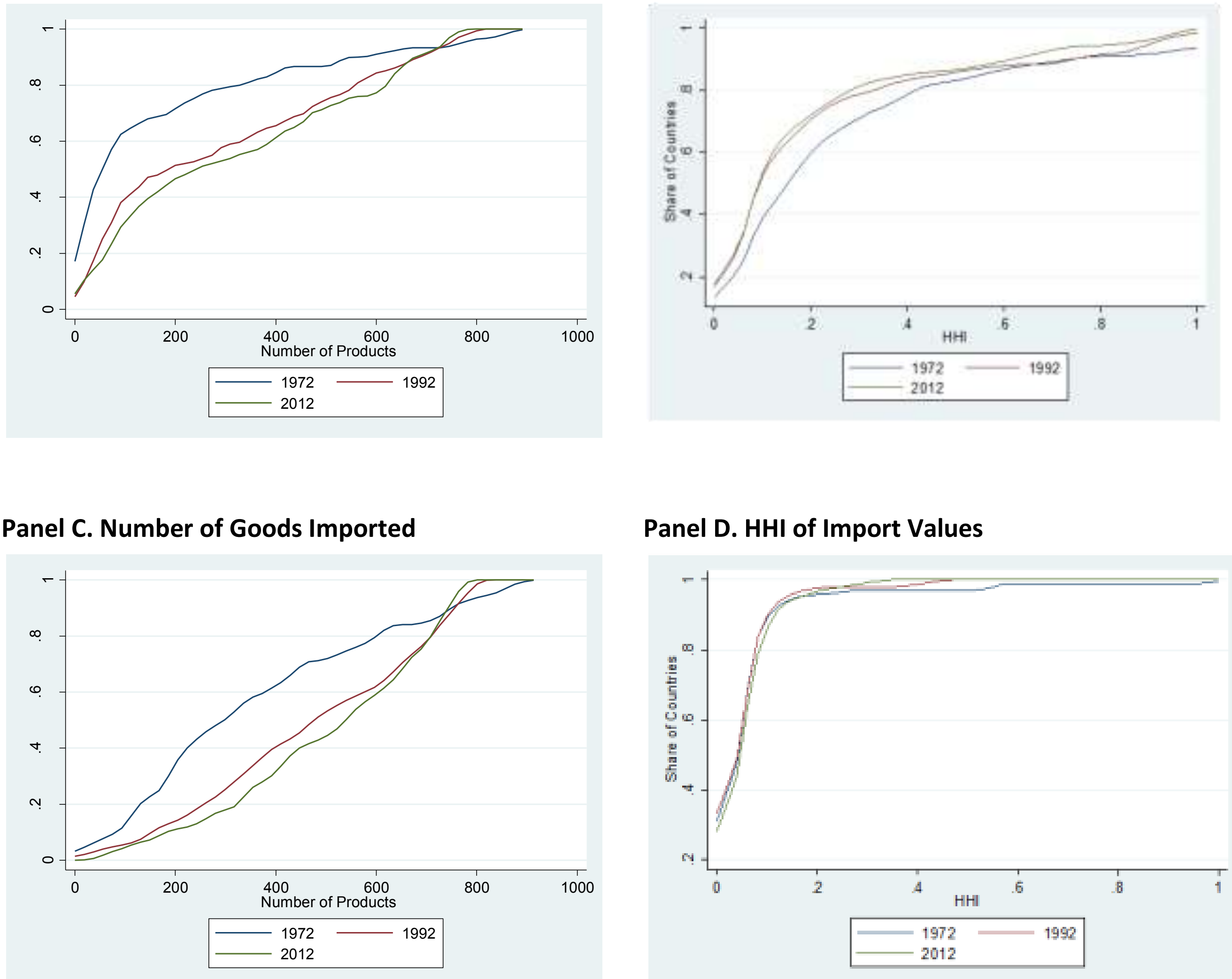


Source: Authors' calculation based on data from the UN's Comtrade.

Figure 2. Cumulative Distribution of Measures of Trade Concentration

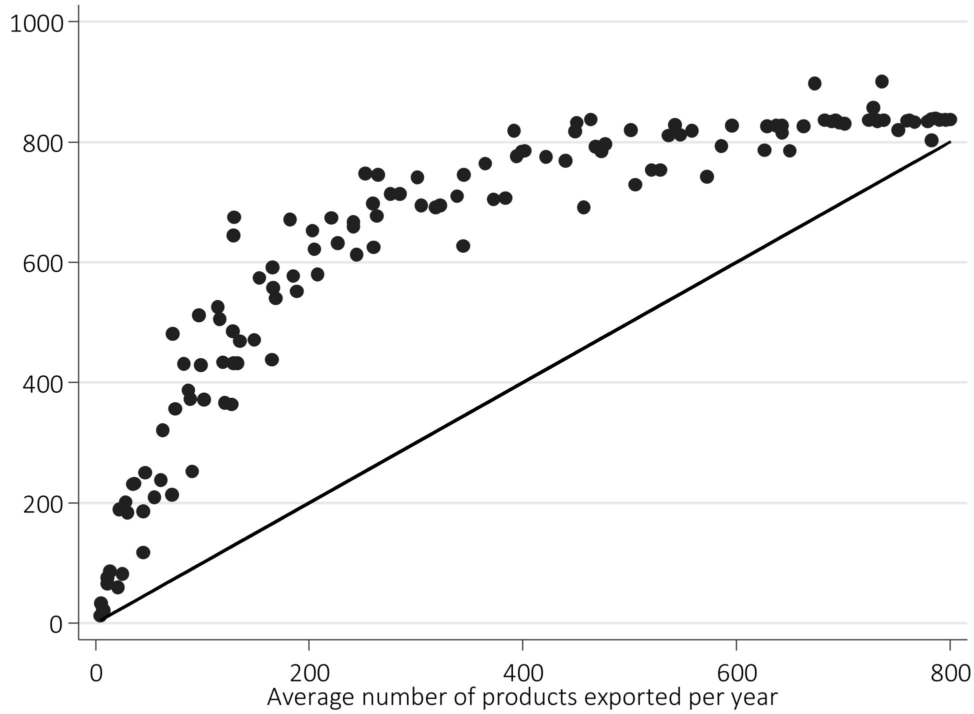
Panel A. Number of Goods Exported **Panel B. HHI of Export Values**

Panel C. Number of Goods Imported **Panel D. HHI of Import Values**



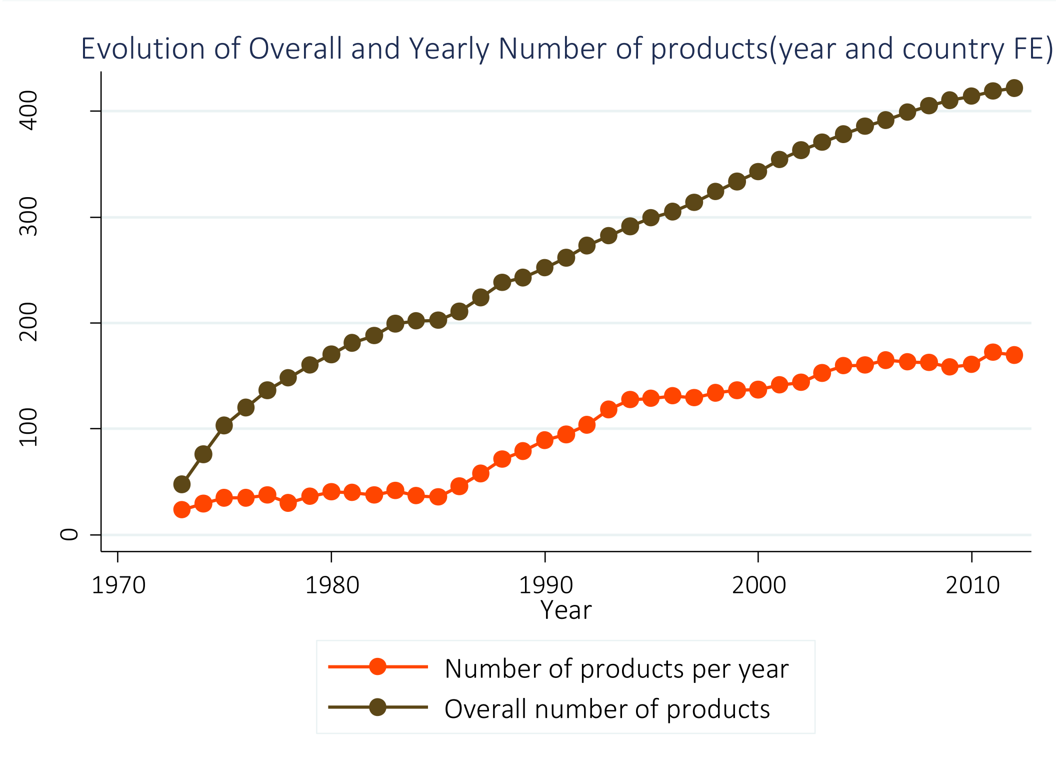
Source: Authors' calculation based on data from the UN's Comtrade.

Figure 3. Average Number of Goods Exported per year vs. Number of Latent Export Lines:, 1972-2012



Source: Authors' calculation based on data from the UN's Comtrade, SITC rev. 1, at the 5-digit level.

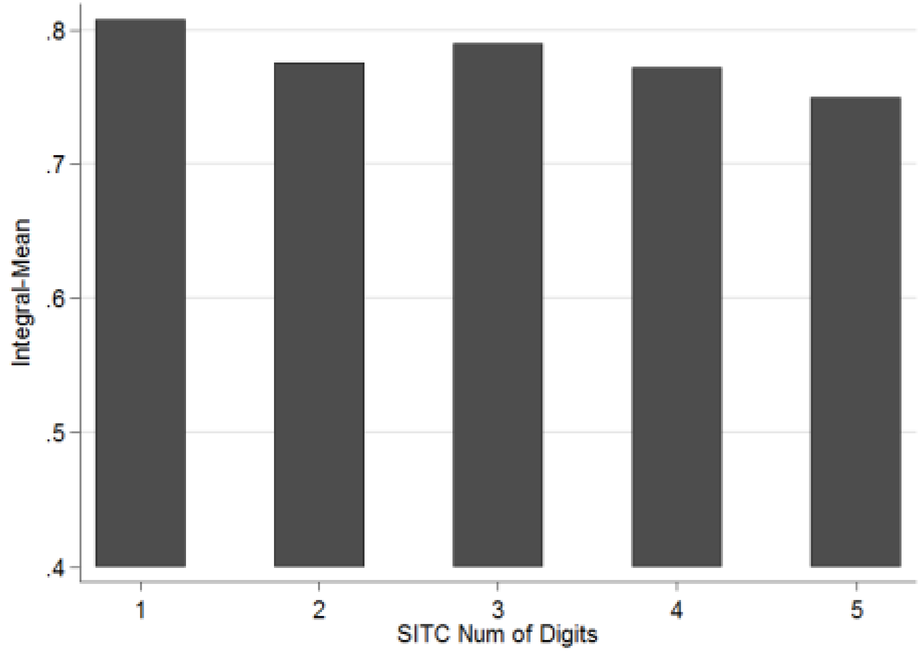
Figure 4. Evolution of the Number of Goods Exported per year and the Number of Latent Export lines: 1972-2012



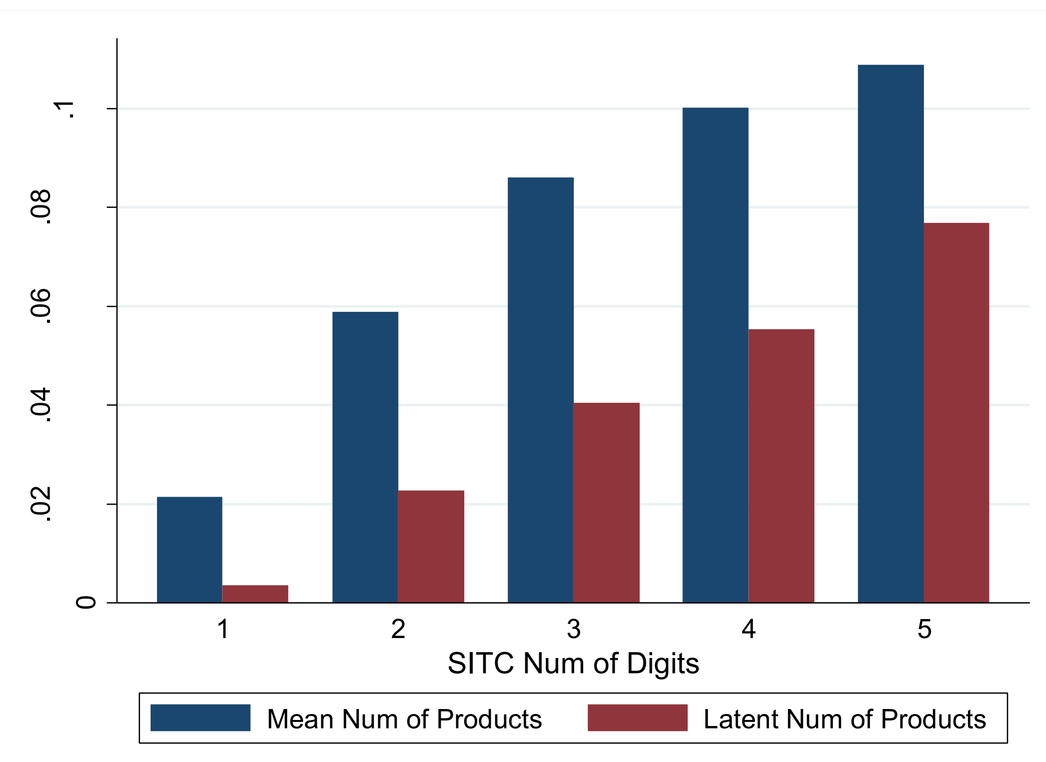
Source: Authors' calculation based on data from the UN's Comtrade, STIC rev. 1, at the 5-digit level.

Figure 5. Export Diversification, Latent Export Diversification, and Aggregation

Panel A. Estimated Average Distance between Export Diversification and Latent Export Diversification (D_{SITC-k}), by Level of Aggregation

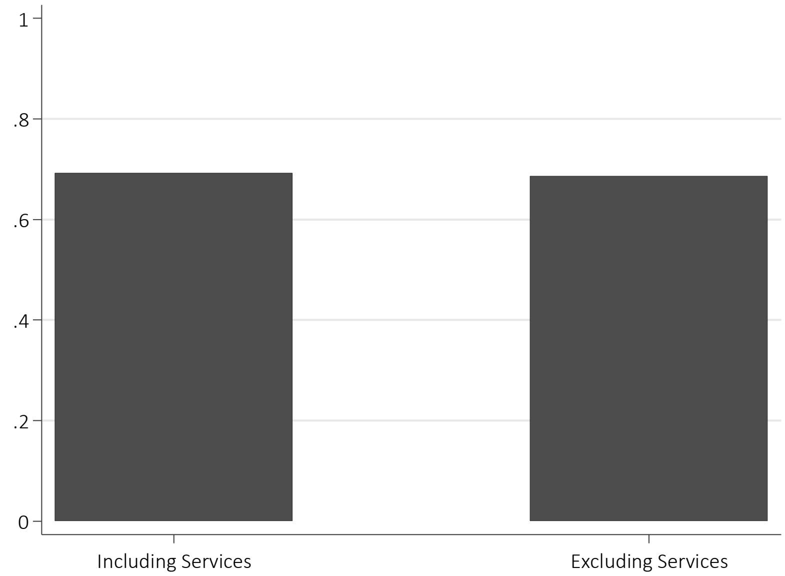


Panel B. Variance by Measure of Diversification and Level of Aggregation



Source: Authors' calculation based on data from the UN's Comtrade, SITC rev 1, 1-5 digits. Note: See text for a definition of the variables.

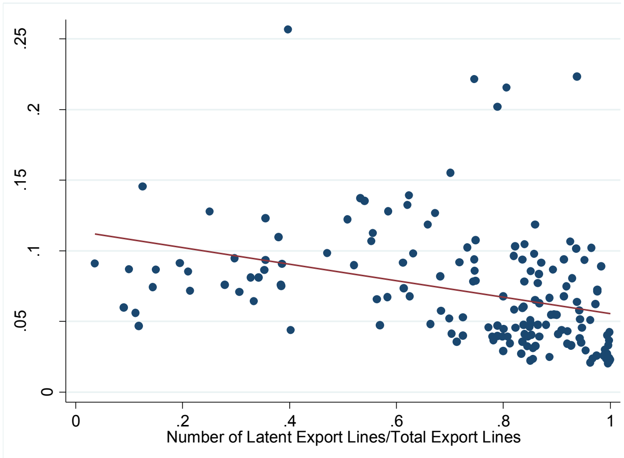
Figure 6. Export Diversification, Latent Export Diversification, and Trade in Services



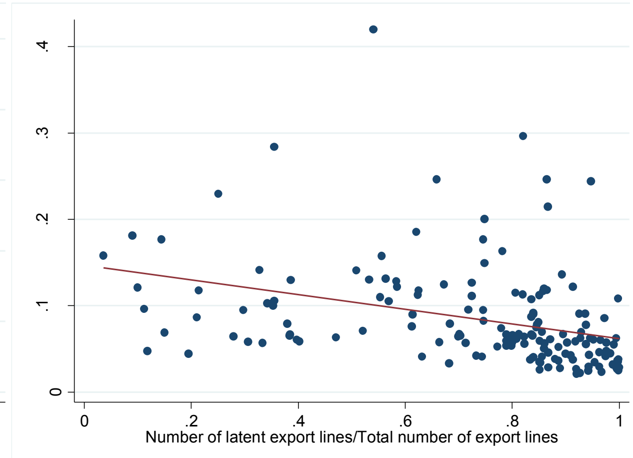
Source: Authors' calculation based on data from the UN's Comtrade. Note: See text for a definition of the variables.

Figure 7. Latent Export Lines, Average Export Lines and Volatility

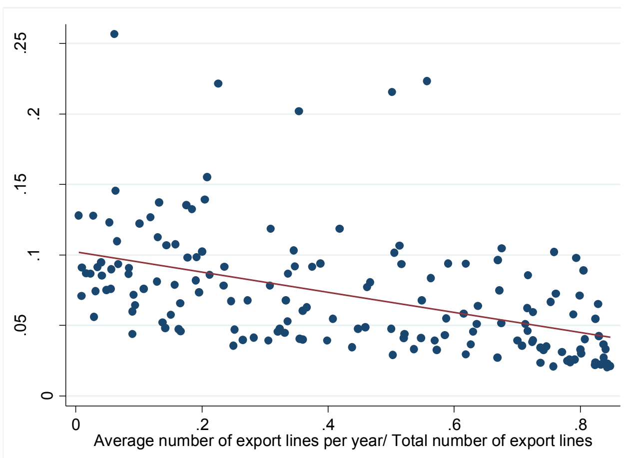
Panel A. Latent Export Lines and ToT Growth Volatility



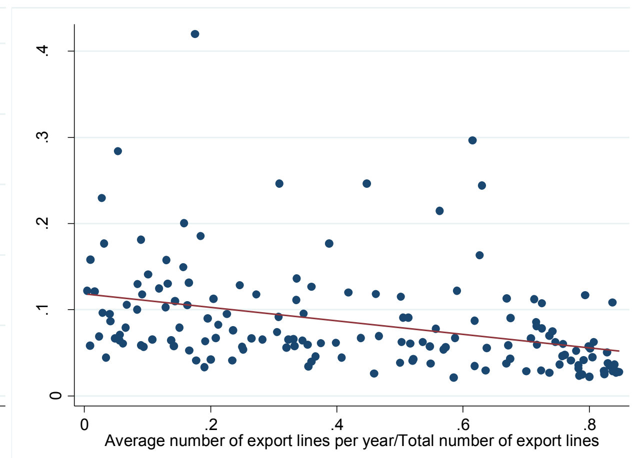
Panel B. Latent Export Lines and Real GDP per capita Growth Volatility



Panel C. Average Export Lines and ToT Growth Volatility

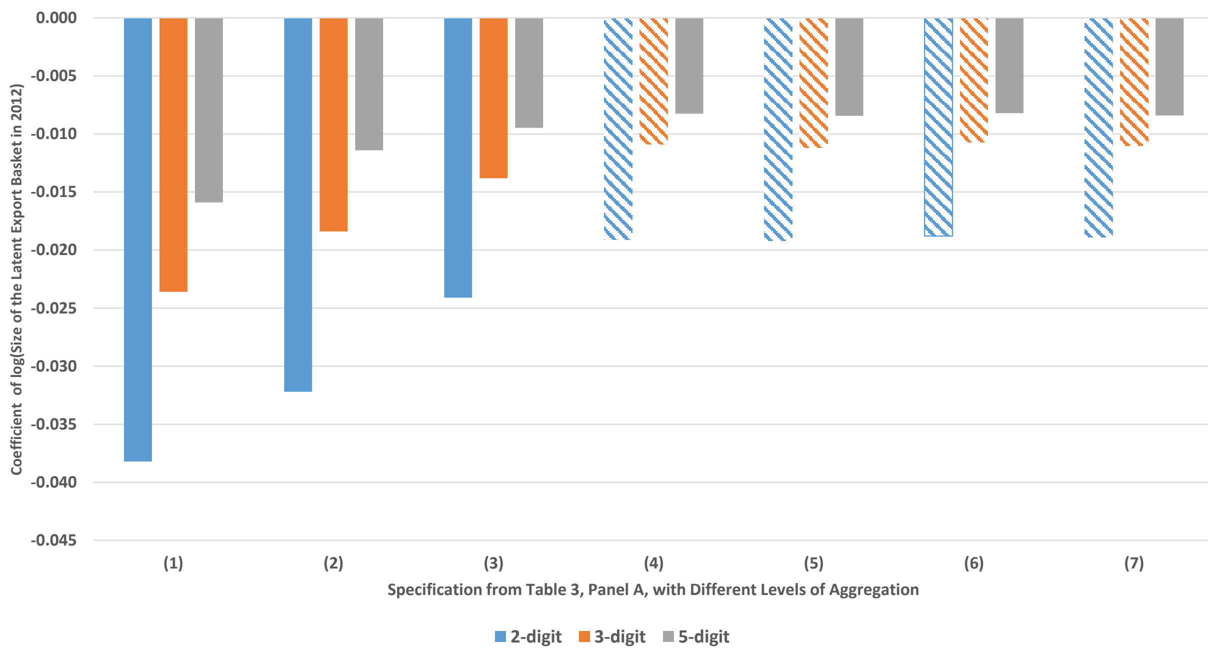


Panel D. Average Export Lines and Real GDP per capita Growth Volatility



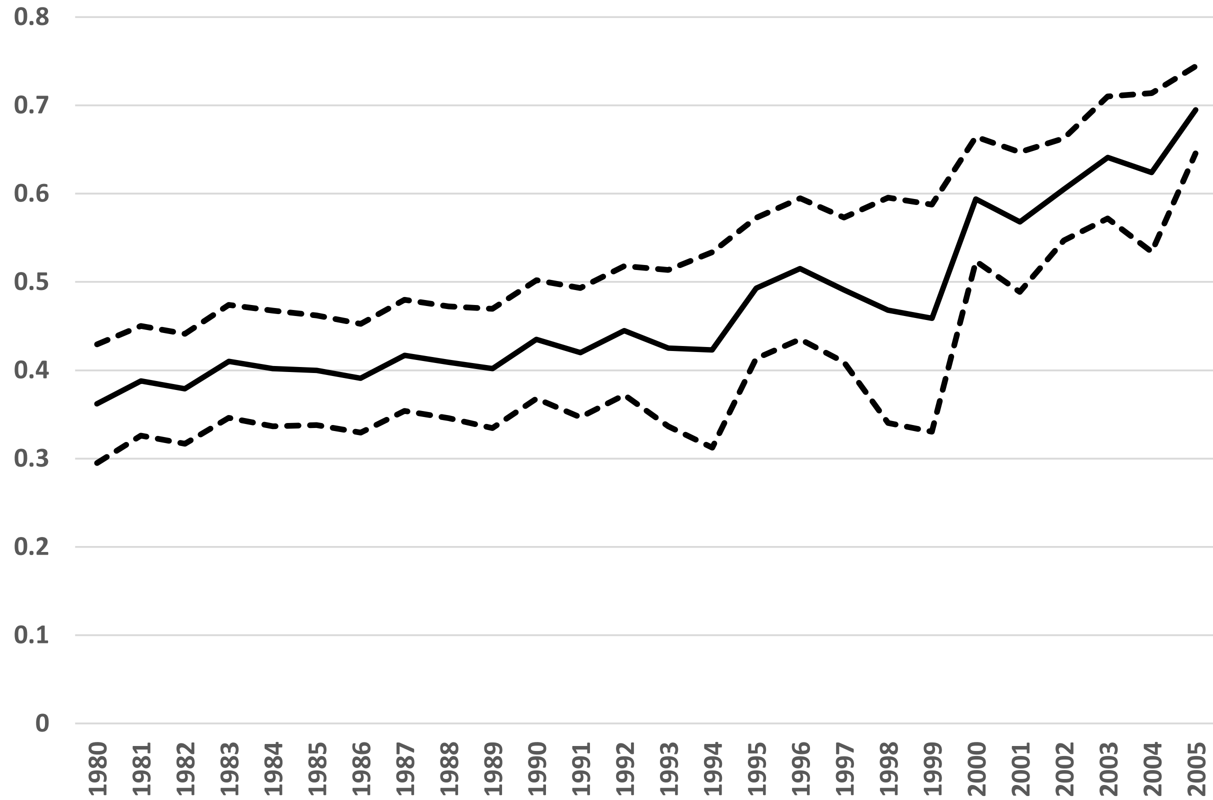
Sources: UN's COMTRADE and PWT rev 8.1. Notes: Solid lines are the fitted values of an OLS regression of volatility on the number of export lines. The number of latent export lines are calculated for the period 1972-2012. Average number of export lines per year are calculated for the period 1980-2012. Volatilities are calculated for the period 1980-2012.

Figure 8. ToT Volatility and Latent Diversification—the Effect of Aggregation



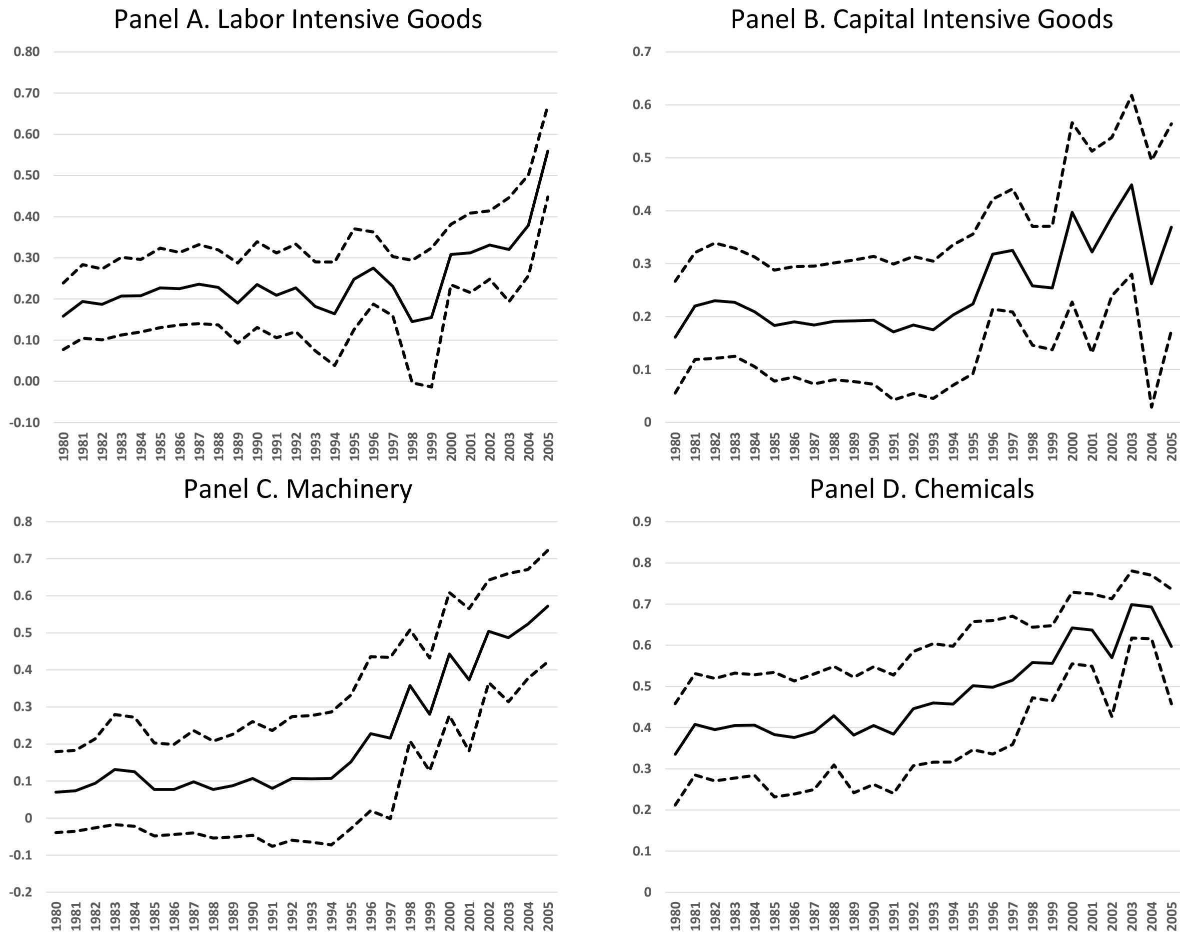
Sources: Authors' estimates based on data from UN's COMTRADE and PWT rev. 8. Note: Each bar represents the coefficient of the log of the Size of the Latent Export Basket in 2012 using different levels of aggregation of exports in each of the specifications in columns (1)-(7) of Table 3, Panel A. Dashed bars represent coefficients that are **not** statistically significant at the 10% level of significance. 2-digit stands for data coming from the SITC rev.1 dataset at the 2 digit level of aggregation, 3-digit stands for data coming from the SITC rev.1 dataset at the 3 digit level of aggregation, 5-digit stands for data coming from the SITC rev.1 dataset at the 5 digit level of aggregation.

Figure 9. Relationship between Sectoral Shares in the Cross-Section and Sectoral Shares in Latent Baskets



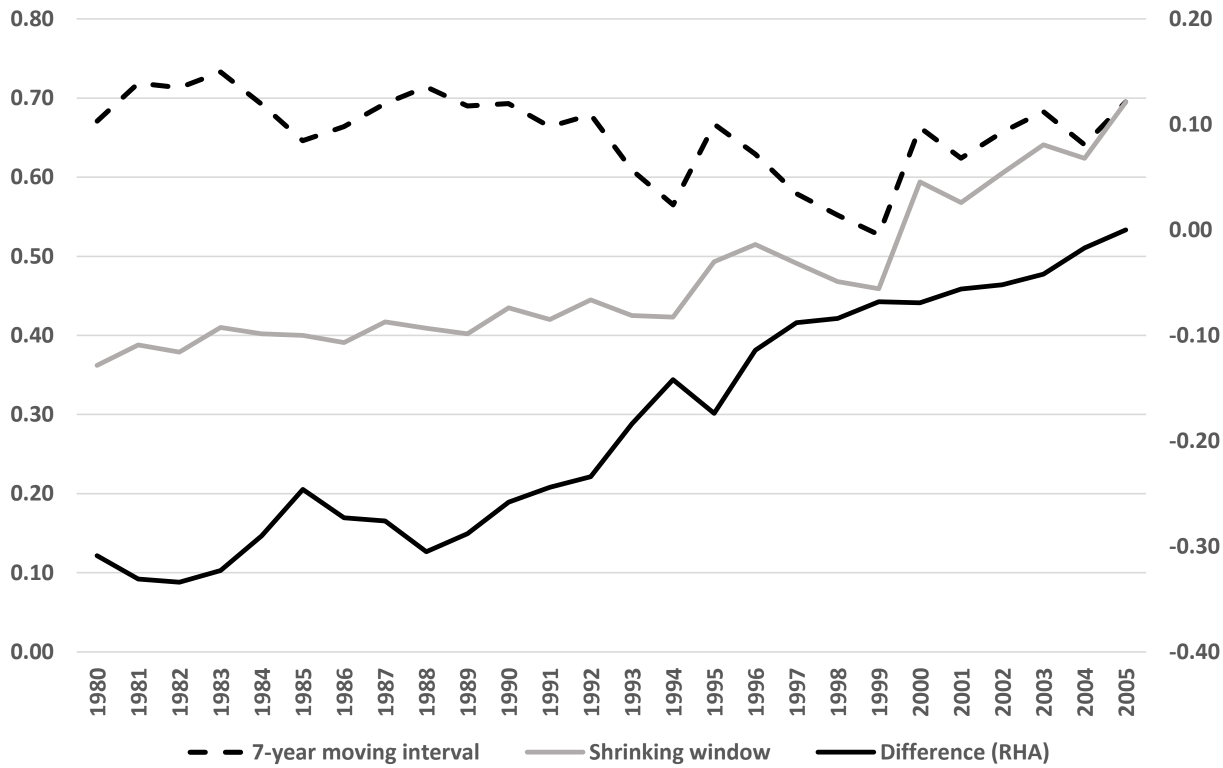
Source: Authors' calculation based on data from the UN's Comtrade. Note: The graph plots the coefficient of a regression between the share of sector j in the goods added to the latent export basket from t to 2012 and the share of sector j in the cross-sectional export basket in year t, controlling for sector fixed effects. Sectors are defined as in Leamer (1995) based on factor intensities. In the estimation we use the four sectors with the largest number of export lines in them, namely labor intensive goods, capital intensive goods, chemicals, and machinery. Dotted lines represent the 95 percent confidence interval. For more details see text.

Figure 10. Relationship between Sectoral Shares in the Cross-Section and Sectoral Shares in Latent Baskets, by Industry Cluster



Source: Authors' calculation based on data from the UN's Comtrade. Note: Each graph plots the coefficient of a regression between the share of sector j in the goods added to the latent export basket from t to 2012 and the share of sector j in the cross-sectional export basket in year t. Sectors are defined as in Leamer (1995) based on factor intensities. Dashed lines represent the 95 percent confidence interval. For more details see text.

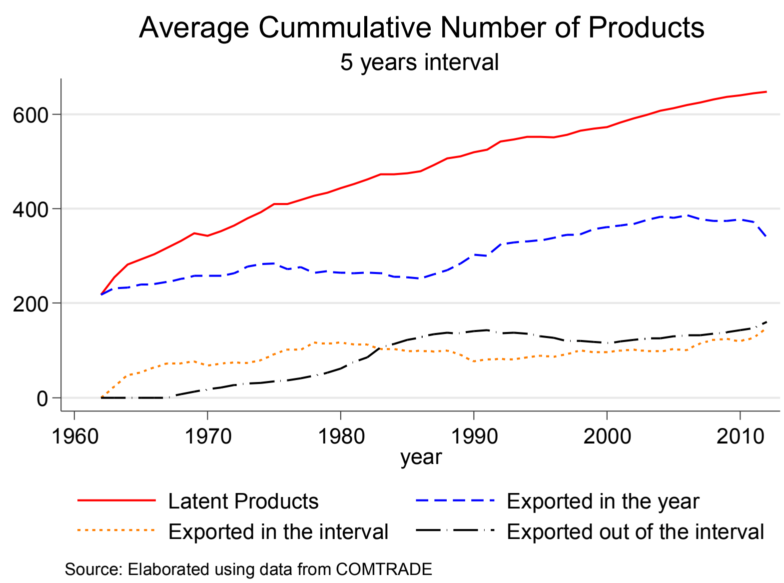
Figure 11. Relationship between Sectoral Shares in the Cross-Section and Latent Baskets: Stickiness of Initial Distributions vs. Time Effects



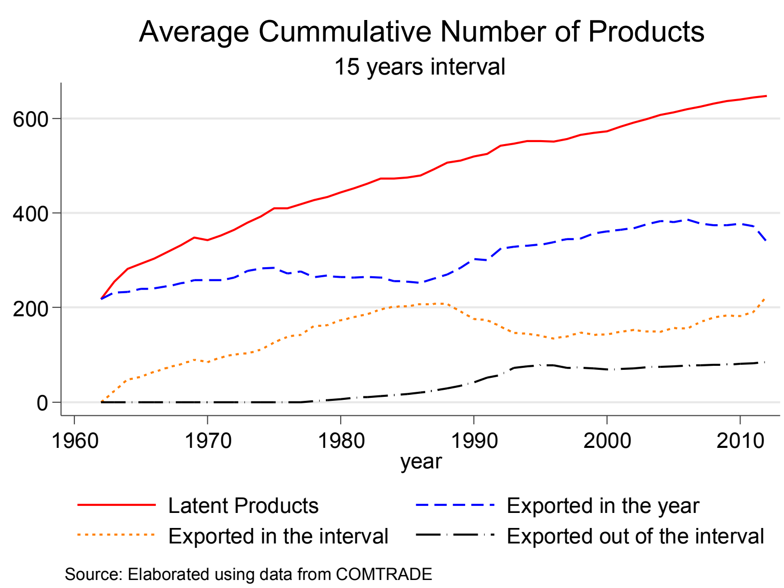
Source: Authors' calculation based on data from the UN's Comtrade. Note: The solid gray line plots the coefficient of a regression between the share of sector j in the goods added to the latent export basket from t to 2012 and the share of sector j in the cross-sectional export basket in year t, controlling for sector fixed effects. The dashed line plots the coefficient of a regression between the share of sector j in the goods added to the latent export basket from t to t+7 and the share of sector j in the cross-sectional export basket in year t, controlling for sector fixed effects. The solid black line plots the difference between the gray and the dashed lines. Sectors are defined as in Leamer (1995) based on factor intensities. For more details see text.

Figure A2.1. Distribution of latent export products by status of activity and population

Panel A. 5-year Interval



Panel B. 15-year Interval



Sources: UN's COMTRADE. Notes: Lines represent the world average for each of the latent export basket and each of the three subsets comprising it—exported in year, exported in the interval, and exported outside the interval.

Table 1. Determinants of Latent Export Baskets and its Gap with Cross-Sectional Export Baskets

Panel A. ln (Size of Latent Export Basket in 2012)

	OLS						SS-T-W					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Log of Labor Force in 1980, in millions	0.122***	0.155***	0.155***	0.152***	0.152***	0.156***	0.0738	1.236	1.137**	1.421*	1.421*	1.849***
	(0.0329)	(0.0329)	(0.0331)	(0.0334)	(0.0334)	(0.0546)	(0.0784)	(1.818)	(0.559)	(0.742)	(0.742)	(0.706)
Log of GDP per capita in 1980, in thousand		0.206***	0.206***	0.206***	0.206***	0.205***		1.835	1.678*	2.171*	2.171*	2.362**
		(0.0559)	(0.0557)	(0.0555)	(0.0555)	(0.0533)		(2.689)	(0.932)	(1.138)	(1.138)	(1.185)
Average Net exports of Agriculture per worker prior to 1980			-0.00949	-0.0485	-0.0485	-0.0448			-0.552	-1.761	-1.761	-1.600
			(0.192)	(0.180)	(0.180)	(0.181)			(1.578)	(1.713)	(1.713)	(1.398)
Average Net exports of Mining and Energy per worker prior to 1980				0.0167**	0.0167**	0.0164**				0.114	0.114	0.101
				(0.00637)	(0.00637)	(0.00752)				(0.0791)	(0.0791)	(0.0905)
Geographic trade openness over GDP in 1980						0.000398						0.0276
						(0.00455)						(0.0243)
log of the Size of the Average Cross-Sectional Export Basket, 1962-1980	0.178***	0.0979**	0.0980**	0.0969**	0.0969**	0.0968**	0.105	0.488	0.448**	0.531**	0.531**	0.579
	(0.0527)	(0.0426)	(0.0432)	(0.0429)	(0.0429)	(0.0432)	(0.0688)	(0.670)	(0.188)	(0.221)	(0.221)	(0.417)
Constant	5.347***	3.962***	3.962***	3.970***	3.970***	3.966***	0.305	-7.524	-6.918*	-8.877*	-8.877*	-10.18**
	(0.284)	(0.510)	(0.510)	(0.515)	(0.515)	(0.531)	(0.518)	(10.09)	(4.138)	(4.693)	(4.693)	(4.895)
Observations	95	95	95	95	95	95	95	95	95	95	95	95
R-squared	0.317	0.385	0.385	0.388	0.388	0.388						

Panel B. ln (Size of Latent Export Basket in 2012)-ln (Size of the Average Yearly Export Basket, 1962-1980)

	OLS						SS-T-W					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Log of people engaged in 1980	-0.216***	-0.218***	-0.216***	-0.213***	-0.213***	-0.220***	-0.632**	-0.836***	-0.841***	-0.845***	-0.845***	-0.887***
	(0.0186)	(0.0191)	(0.0189)	(0.0193)	(0.0193)	(0.0235)	(0.247)	(0.136)	(0.152)	(0.189)	(0.189)	(0.187)
Log of GDP per capita in 1980		-0.266***	-0.269***	-0.269***	-0.269***	-0.267***		-1.015***	-1.037***	-1.043***	-1.043***	-1.015***
		(0.0342)	(0.0342)	(0.0361)	(0.0361)	(0.0360)		(0.167)	(0.163)	(0.191)	(0.191)	(0.198)
Average Net exports of Agriculture per worker prior to 1980			-0.508**	-0.465**	-0.465**	-0.473***			-1.648**	-1.669**	-1.669**	-1.690**
			(0.202)	(0.177)	(0.177)	(0.178)			(0.793)	(0.806)	(0.806)	(0.785)
Average Net exports of Mining and Energy per worker prior to 1980				-0.0180	-0.0180	-0.0174				0.00400	0.00400	0.00791
				(0.0214)	(0.0214)	(0.0213)				(0.0750)	(0.0750)	(0.0736)
Geographic trade openness over GDP						-0.000891						-0.00650
						(0.00209)						(0.00752)
Constant	0.961***	3.216***	3.237***	3.228***	3.228***	3.236***	-0.835	8.051***	8.284***	8.343***	8.343***	8.271***
	(0.0469)	(0.280)	(0.281)	(0.293)	(0.293)	(0.298)	(0.856)	(1.711)	(1.725)	(2.123)	(2.123)	(2.154)
Observations	95	95	95	95	95	95	95	95	95	95	95	95
R-squared	0.473	0.760	0.767	0.773	0.773	0.774						

Source: COMTRADE and PWT rev.8.

Note: Robust standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Latent basket stands for the empirical approximation of the number of latent export lines in 2012 described in the text. Average yearly basket stands for the average number of export lines per year between 1980 and 2012. GDP per capita is PPP 2005 US$. SS-T-W stands for the estimation methodology proposed by Santos Silva, Tenreyro and Wei (2014).

Table 2. Opened Export, Closed Export Lines, Size of the Labor Force and GDP per capita

Panel A. Relative to t-1

	Opened Export Lines						Closed Export Lines					
	OLS	OLS	OLS	SS-T-W	SS-T-W	SS-T-W	OLS	OLS	OLS	SS-T-W	SS-T-W	SS-T-W
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
log of the Labor Force in 1980, in millions	-3.445***		-4.133***	-0.844		-1.120***	-3.657***		-4.330***	-0.901**		-1.066***
	(0.375)		(0.354)	(0.554)		(0.315)	(0.354)		(0.345)	(0.352)		(0.260)
log of GDP per capita in 1980, in thousands		-5.148***	-6.409***		-0.198**	-1.580***		-4.886***	-6.215***		-0.197**	-1.396***
		(0.992)	(0.687)		(0.0974)	(0.455)		(0.992)	(0.674)		(0.0880)	(0.359)
Constant	16.55***	26.59***	33.60***	0.276	-1.734***	4.181**	14.80***	23.68***	30.75***	-0.0434	-1.903***	3.160**
	(1.332)	(2.933)	(2.162)	(2.130)	(0.377)	(1.989)	(1.319)	(2.707)	(2.059)	(1.198)	(0.325)	(1.598)
Observations	3,363	3,363	3,363	3,272	3,272	3,272	3,267	3,267	3,267	3,267	3,267	3,267
R-squared	0.189	0.157	0.368				0.207	0.136	0.398			

Panel B. Relative to Initial Year

	Opened Export Lines						Closed Export Lines					
	OLS	OLS	OLS	SS-T-W	SS-T-W	SS-T-W	OLS	OLS	OLS	SS-T-W	SS-T-W	SS-T-W
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
log of the Labor Force in 1980, in millions	-3.378**		-4.580***	-0.0621		-0.558**	-4.533***		-4.865***	-0.604		-0.907***
	(1.367)		(1.149)	(0.0651)		(0.262)	(0.482)		(0.501)	(0.558)		(0.269)
log of GDP per capita in 1980, in thousands		-9.795***	-11.19***		-0.276**	-1.261***		-2.351**	-3.539***		-0.0256***	-0.712***
		(1.909)	(1.640)		(0.127)	(0.458)		(1.144)	(0.926)		(0.00731)	(0.255)
Constant	51.19***	73.21***	80.98***	-0.907*	0.121	5.525**	20.78***	23.62***	29.65***	-0.409	-2.361***	2.118
	(2.903)	(4.545)	(4.061)	(0.528)	(0.550)	(2.551)	(1.875)	(3.455)	(3.311)	(1.849)	(0.0372)	(1.568)
Observations	3,363	3,363	3,363	3,272	3,272	3,272	3,869	3,869	3,869	3,778	3,778	3,778
R-squared	0.200	0.292	0.376				0.189	0.056	0.223			

Source: COMTRADE and PWT rev.8.

Note: Standard errors are clustered at the country level and are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1. GDP per capita is PPP 2005 US$. All regression include year fixed effect. SS-T-W refers to the estimation methodology proposed by Santos Silva, Tenreyro, and Wei (2014). See text for a detailed explanation of the variables.

Table 3. Volatility, Cross-Sectional Diversification, and Latent Diversification

Panel A. Determinants of Terms-of-Trade Growth Volatility

Dependent variable: St. Dev. of ToT growth	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log of the Size of the Latent Export Basket in 2012	-0.0159***	-0.0114***	-0.00948**	-0.00825*	-0.00843*	-0.00820*	-0.00840*
	(0.00298)	(0.00398)	(0.00407)	(0.00455)	(0.00461)	(0.00450)	(0.00454)
log of the Size of the Average Cross-Sectional Export Basket, 1962-1980		-0.00421*	-0.00283	-0.00224	-0.00217	-0.00222	-0.00215
		(0.00215)	(0.00276)	(0.00291)	(0.00292)	(0.00295)	(0.00294)
log of GDP per capita in 1980, in thousand			-0.00587	-0.00661*	-0.00670*	-0.00623	-0.00625
			(0.00372)	(0.00395)	(0.00395)	(0.00396)	(0.00399)
log of Labor Force in 1980, in millions				-0.00137	-0.00141	-0.00266	-0.00294
				(0.00157)	(0.00157)	(0.00239)	(0.00247)
Average Net exports of Mining and Energy per worker prior to 1980					0.000651		0.000769
					(0.00167)		(0.00166)
Average Net exports of Agriculture per worker prior to 1980					-0.0145		-0.0162
					(0.0172)		(0.0180)
Geographic trade openness over GDP						-0.000156	-0.000183
						(0.000188)	(0.000197)
Constant	0.160***	0.152***	0.183***	0.180***	0.182***	0.181***	0.183***
	(0.0190)	(0.0203)	(0.0304)	(0.0312)	(0.0306)	(0.0314)	(0.0306)
Observations	95	95	95	95	95	95	95
R-squared	0.192	0.238	0.281	0.286	0.292	0.291	0.298

Panel B. Determinants of GDP per capita Growth Volatility

Dependent variable: St. Dev. of GDP per capita growth	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
S.D of terms of trade after 1980	0.619***	0.532***	0.534***	0.561***	0.531***	0.528***	0.550***	0.550***
	(0.106)	(0.120)	(0.123)	(0.121)	(0.117)	(0.116)	(0.113)	(0.113)
log of the Size of the Latent Export Basket in 2012		-0.00720**	-0.00730**	-0.00791**	-0.00307	-0.00322	-0.00307	-0.00311
		(0.00355)	(0.00352)	(0.00358)	(0.00429)	(0.00434)	(0.00416)	(0.00423)
log of the Size of the Average Cross-Sectional Export Basket, 1962-1980			0.000122	-0.000417	0.00198	0.00196	0.00198	0.00195
			(0.00166)	(0.00176)	(0.00191)	(0.00192)	(0.00183)	(0.00185)
log of GDP per capita in 1980, in thousand				0.00279	-0.000465	-0.000469	-0.00154	-0.00149
				(0.00278)	(0.00288)	(0.00287)	(0.00267)	(0.00271)
log of Labor Force in 1980, in millions					-0.00572***	-0.00577***	-0.00172	-0.00176
					(0.00140)	(0.00141)	(0.00179)	(0.00183)
Average Net exports of Mining and Energy per worker prior to 1980						0.000456		0.000135
						(0.000551)		(0.000566)
Average Net exports of Agriculture per worker prior to 1980						-0.00235		0.00243
						(0.0124)		(0.0134)
Geographic trade openness over GDP							0.000480**	0.000478**
							(0.000217)	(0.000223)
Constant	0.0233***	0.0738***	0.0738***	0.0550*	0.0470	0.0484	0.0398	0.0399
	(0.00563)	(0.0266)	(0.0268)	(0.0312)	(0.0341)	(0.0338)	(0.0318)	(0.0321)
Observations	95	95	95	95	95	95	95	95
R-squared	0.328	0.355	0.355	0.363	0.438	0.439	0.473	0.474

Source: COMTRADE, PWT rev.8.

Note: *** p<0.01, ** p<0.05, * p<0.1. Latent basket stands for the empirical approximation of the number of latent export lines in 2012 described in the text. Average cross-sectional basket stands for the average number of export lines per year between 1962 and 1980. GDP per capita is PPP 2005 US$. See main text for additional variables definitions and discussion.

Table 4. Determinants of Export Concentration, ToT growth Volatility, and GDP per capita Growth Volatility, 1980-2012 (3SLS)

	(1)	(2)	(3)
VARIABLES	log of the Size of the Latent Export Basket in 2012	S.D. of ToT Growth	S.D. of GDP per capita Growth
log of Labor Force in 1980, in millions	0.154***	0.00339	-0.00147
	(0.0395)	(0.00260)	(0.00239)
Log of GDP per capita in 1980, in thousand	0.208***		
	(0.0625)		
Average Net exports of Mining and Energy per worker prior to 1980	0.00312		
	(0.0203)		
log of the Size of the Average Cross-Sectional Export Basket, 1962-1980	0.0991**	0.000766	
	(0.0484)	(0.00333)	
log of the Size of the Latent Export Basket in 2012		-0.0390***	
		(0.0148)	
S.D. of ToT Growth, 1980-2012			0.589***
			(0.195)
Geographic trade openness over GDP			0.000467**
			(0.000199)
Constant	3.939***	0.299***	0.0147
	(0.465)	(0.0795)	(0.0161)
Observations	95	95	95

Source: Authors'calculation from Comtrade and PWT rev. 8.

Note: *** p<0.01, ** p<0.05, * p<0.1. Cross-sectional error correlations are assumed to be unstructured. Volatility is measured by the standard deviation of the annual growth rate of each variable in the period between 1980 and 2012. The first stage and the intercepts are not reported. The results correspond to corss-sectional estimates for the 1980-2012 period.