Final report

International Growth Centre

Linkages between factor endowments and export diversification

Empirical insights from the firm-level data of Pakistan

Salamat Ali
Shafqat Ali Khan

March 2017

When citing this paper, please use the title and the following reference number:
F-37317-PAK-1

DIRECTED BY LSE UNIVERSITY OF OXFORD

FUNDED BY UKaid from the British people

Linkages between Factor Endowments and Export Diversification: Empirical Insights from the Firm-Level Data of Pakistan

Salamat Ali and Shafqat Ali Khan

Abstract

This paper uses transaction-level data for Pakistan over the period 2000-2013 to examine the relationship between firm-level measures of export diversification and the degree of dissimilarity in factor endowments in the destinations. The measure of export diversification exploits a unique feature of the data: detailed product descriptions within narrow HS8 categories. The main findings are that firms export a larger number of varieties to countries with more dissimilar endowments of physical and human capital, and to destinations with lower trade costs. The effect is larger for manufacturing exports compared with those from the agriculture sector and is particularly large for exports to developed economies. The physical capital remains the largest factor positively affecting diversification, although its contribution has diminished over time and that of human capital has increased as the economy has moved towards re-specialisation in recent years.

Keywords**:** Export diversification; endowment inequality; trading partner's choice; firm-level data, Pakistan

JEL codes: F1, F13, F14

Disclaimer

The study uses administrative datasets of the Government of Pakistan, some of which are completely confidential in nature. We accessed these datasets as internal researchers and have taken extreme care to ensure confidentiality. The research datasets may not exactly reproduce aggregates reported in the government publications. The use of statistical data in this work does not imply the endorsement of the organisations in relation to the interpretation or analysis of the information.

1 Introduction

International trade theory is based mainly on the principle that specialisation according to comparative advantage is optimal. Developing countries, however, are concerned about the challenges caused by export concentration. These include price volatility, sudden closure of markets owing to regulatory changes, entry of new competitors or changes in commodity demand; all these factors have implications for the balance of payments and other fiscal developments. Although theory has not made much progress on this front, empirical interest in export diversification has risen substantially. There are signs that international bodies are moving away from a narrow range of structural reforms, such as liberalization and deregulation, to a wide range of actions needed to diversify the structure and level of sophistication of economic activity (UNCTAD, 2016; IMF, 2016). Investigating this aspect of development is of considerable importance to policy-makers as a country's diversification pattern has implications for subsequent economic performance (Hausmann and Rodrik, 2003; Hummels and Klenow, 2005; Klinger and Lederman, 2004; Hausmann et al., 2007).

In the recent literature on determinants of export diversification, Chowdhury et al. (2014) and Sekkat (2015) examine the relationship between export diversification and exchange rate, and Regolo (2013) investigates the role of factor endowments, Parteka and Tamberi (2013) assess the role of country-specific factors and Persson and Wilhelmsson (2016) examine the role of trade preferences. Cadot et al. (2013) provide a nice survey of the earlier literature and highlight the linkages between firms' productivity, economic growth and level of export diversification. In extension of this stream of research, we document the connections between factor endowments inequality of trading partners and diversification at a firm level, which remains less examined formally, both theoretically and empirically, especially in developing country contexts.

We empirically investigate the linkages between inequality in factor endowments (land, labour and capital) and the level of diversification of Pakistan's exports, and decompose the effect across products, markets and over time. With a highly-disaggregated data, we directly measure export

diversification along EM, instead of computing concentration indices[1]. This is because Cadot et al. (2011) find that diversification and subsequent re-concentration take place mainly along EM. Moreover, this approach offers straightforward interpretation of results.

Three findings emerge from the analysis. First, Pakistan's exports to differently endowed markets are more diversified. Inequality of physical and human capital across trading partners affects diversification positively but inequality in land endowment affects it negatively. The differences in physical capital endowment have a much larger effect than is the case for human capital or land endowments. One standard deviation increase in the inequality of physical capital is associated with increased diversification of 4.77% compared with the corresponding effect of human capital by 1.3% and land endowments by -2.1%.

Second, it shows that the dissimilarity in factor endowments has a differential effect across sectors. Overall, the effect is larger for manufacturing exports than for agricultural products, and is particularly large for exports to developed economies. The differences in human capital matter the most for the diversification of manufactured goods while the differences in land endowments matter for the diversification of agricultural products. The sectoral decomposition further indicates that the diversification occurs for the most part in manufacturing sectors and the contribution of the agriculture sector is quite limited. Third, it finds that the effect of various factor endowments is heterogeneous over time. The role of human capital has increased gradually, whereas that of physical capital has diminished, especially in recent years, as the country has closed some product lines and moved towards re-specialisation[2]. This kind of heterogeneity across sectors, although of a direct interest to policy-makers, remains masked in the analysis at the aggregate level.

These results hold to alterative measures of diversification and estimations at various levels of aggregation of data as well as to a demanding set of fixed effects. In addition to internal validity of the results to a battery of robustness checks, WE also show their external validity by using information

[1] Frequently used export diversification indices are the Herfindhal, Gini and Thiel indices, all of which measure inequality in export shares. The Thiel index using active product lines, as in Regolo (2013), measures export diversification along intensive margins (Bacchetta et al., 2012).
[2] This re-specialisation point appears to arrive at a much lower level of development than that found in other economies (Cadot et al., 2013).

from the Exporter Dynamic Database for 52 developing countries. WE further demonstrate that low trade costs amplify diversification, and that the effect of inequality of factor endowments is still significant even after controlling for trade costs and other policy-related variables. These stylised facts are not explained by existing trade models[3] and thus warrant theoretical extension to catch up with the pattern revealed in the transaction-level data.

The main contribution of this article lies in its examining the diversification of firm-level exports from an emerging economy by using a relatively precise indicator. Thanks to the availability of highly disaggregated data containing actual descriptions of products, WE can identify and account for the differentiation of varieties within the HS8 classification (see Section 2.2). The primary source, in addition to containing information on the products at the highest level of disaggregation, provides for relatively longer coverage of the recent period, 2000–2013. These unique features allow greater flexibility in exploring diversification patterns across firms, products, markets and over time. This administrative dataset of Pakistan is being explored for the first time for such an empirical analysis.

We are not aware of any firm-level study that explores a similar research question. However, in a closely related work Regolo (2013) develops and tests a model predicting a linkage between factor endowments and export diversification using product-level data for a set of developing countries. Since firms are assumed to produce a single product variety, the framework developed by Regolo (2013) does not provide any guidance on the diversification at a firm-level. This paper, therefore, fundamentally differs from Regolo (2013) as the unit of analysis in this work is firm within the country and it follows a common trading partners approach, as in Levchenko (2007), Defever et al. (2015) and many others.

Besides extending the literature on trade diversification, the study also contributes to the stream of literature exploring the effect of trade costs on trade composition. Recently, Milner and McGowan (2013) show that trade costs affect the export mix between OECD countries. Ali and Milner (2015)

3 Neither classical trade theory (Hecksher Ohlin or Ricardian models) nor more recent firm-based models (Melitz, 2003) provide any guidance on the relationship between firm-level export diversification and the destinations' factor endowments or trade cost. Similarly, the framework developed by Regolo (2013) does not provide any guidance on the relationship at a firm level.

document similar effects of trade costs for developing countries. This paper supports the findings of these studies in that the exports to low-trade cost countries are more diversified. Finally, in terms of methodological improvement, we estimate an empirical model in a panel structure, without aggregation of trade flow data, so as to avoid aggregation bias in the results (Anderson and Van Wincoop, 2004). Besides overcoming the problem of measurement error, the use of panel setting makes it possible to control for time-varying omitted variables for firms and products, which was not possible in earlier studies because estimations were made at the higher levels of aggregation[4].

The paper is structured as follows. Section 2 introduces the data and Section 3 discusses estimation strategies and presents the main estimation results. Section 4 conducts a heterogeneity analysis and performs further sensitivity checks. Section 5 concludes with a summary of the findings and their policy implications.

2 Data and Descriptive Analysis

This research uses data on Pakistan's exporting firms from government's primary sources. Pakistan is the sixth-most populous country in the world, with more than 200 million people. It is the 26th-largest economy and is characterised as being among the top 10 emerging and growth-leading countries of the developing world[5]. This section introduces the dataset, discusses the importance of measuring diversification at a micro level and investigates the linkages between diversification and factor endowments.

2.1 Firm-Level Export Data

The micro-level information on EM of firms and products is collected from national data sources[6] of the Federal Board of Revenue (FBR) of the Government of Pakistan. This dataset contains information on product code, prices, quantities and description of each item at the transaction level. The cleaned

[4] Regolo (2013) averages the observations over the sample period, 1995-2007.
[5] http://thelondonpost.net/2015/06/pakistan-an-emerging-economy/
[6] These datasets are subjected to confidentiality. Most of the information is available from the Export Dynamics Database (EDD) of the World Bank, however.

dataset contains 15 million transactions comprising all product categories in manufacturing as well as in agriculture. It covers the universe of firms that shipped to any of 215 trading partners of Pakistan between 2000 and 2013. The econometric estimations, however, use this information for 140 export markets[7]. The large set of export markets provides for wide geographical and temporal coverage of developed and developing countries. We group the data in 98 sectors as per two-digit level of Harmonised System and compute export diversification indicator for each sector. This transformation yields 1,020,257 observations. The estimation sample contains 43,348 unique firms and represents 97% of Pakistan's exports. We test the integrity and accuracy of the data by performing aggregation tests and comparing the results with the same information retrieved from the UN Comtrade dataset.

This firm-level dataset is superior for the analysis of export diversification for the following reasons. First, it identifies the products at the highest possible level of disaggregation (HS8) and thus allows for examining precise growth along EM. Second, the availability of actual descriptions of products at the item level makes it possible to account for the differentiation of varieties within the HS8 categories. Third, unlike Regolo (2013), which uses the UN Comtrade dataset for 1995–2007, this dataset covers the recent period also. Profound changes have occurred in trade patterns since the financial crisis of 2008–2009, and Pakistan, like many other counties, has entered many preferential trade agreements in this period. As these policy measures also affect diversification by reducing trade costs, this long-time span (2000–2013) thus permits examining the behaviour of exports over time.

While the information on actual description of products is an interesting and potentially useful feature of the data, it is not devoid of limitations. We take extreme precaution to check the data set for typos in descriptions because, in some cases, descriptions of products are not consistently defined across firms, and due to its large size of the data it is difficult to manually standardize the detailed descriptions. To circumvent this shortcoming, we test the robustness of results by using HS8 product count as in in Persson and Wilhelmsson (2016) and Dennis and Shepherd (2011).

[7] The number of markets is dictated by the availability of data on factor endowments.

2.2 Measuring Export Diversification

The existing literature mostly measures trade diversification indirectly through various concentration indices, such as the Thiel, Gini and Herfindhal indices. Some studies use direct measures such as product count at HS4 or HS6 levels[8]. WE follow the second approach but dig deeper into extensive margins and find empirical support for the concept of 'product proximity' (Hausmann and Klinger, 2006), and exploit this in estimations. The reason is that 'it is very easy to get puzzled or misled by indicators, and the more complicated the trickier' (Bacchetta et al., 2012).[9] We therefore count the number of products exported by firm across markets at HS8 level, which is the highest level of disaggregation in the Customs dataset. In addition, we consider the differentiation of products within the HS8 categories, which is made possible thanks to the availability of actual descriptions of products at the item level. The micro focus on diversification is needed to account for product varieties within the narrowly defined HS8 categories, as discussed in the following paragraphs.

Table 1 presents the description of four HS8 products exported by the same firm. It indicates a great deal of heterogeneity of products within the same HS8 level of classification. As the data suggests these four HS8 products actually cover 32 different varieties. A firm may change the style of the garments, produce different types of products for men, women and children or change the input mix to generate different variety. For some products there are more than 100 varieties within the same HS8 code. Both supply- and demand-side factors can explain this differentiation of products. Costs of adaptation are lower for similar products within firms. Similarly, once a firm has entered a specific market, the additional cost of introducing new varieties of similar products may be lower. If a firm sells its products under a specific brand, it may benefit from brand loyalty of consumers. Similar observations about Chinese firms have been made in Defever et al. (2015) but this differentiation of products within the HS8 categories has not been considered in any empirical study on export diversification. This approach of measuring diversification at the detailed level appears to be more in tune with the actual volume of exports, whereas measuring it as per HS8 classification or at the higher level (HS6 or HS4) may lead to misleading inference (Table 2).

[8] For survey, see Cadot et al. (2011).
[9] UNCTAD's Handbook of Trade Policy Analysis advises looking at the numbers, not just the indicators (Bacchetta et al., 2012).

Table 1: Product Varieties within HS8 Categories, 2013

Firm ID	HS8 Code	Description of Products
1216908	6105.1000	100% cotton printed tops mix print/style # 3420c &3454c top-ref.100% cotton
1216908	6105.1000	Unisex dyed top style # 8862
1216908	6105.1000	Unisex dyed bottoms # 8861
1216908	6105.1000	Unisex scrub white top style # 8862
1216908	6105.1000	Unisex scrub dyed top style # 8864
1216908	6105.1000	Unisex scrub dyed top style # 8865
1216908	6105.1000	Unisex basic coats (pkwc & wccw) in woven white fabric style cc-121 & cc-122
1216908	6105.1000	Unisex executive coats (ewcc) style cc-101 in woven white fabric
1216908	6105.1000	Unisex Syed jewel neck raglan sleeve jacket style # 102
1216908	6105.1000	Unisex white neck raglan sleeves jacket style # 102
1216908	6105.1000	Women top style # 102 in woven printed fabric
1216908	6105.1000	Women warp top style # 109 in woven printed fabric
1216908	6105.1000	Men lab coats 40" 3 pocket with side slits
1216908	6105.1092	Men's white consultation jackets 33" style # 5730
1216908	6105.1092	Women lab coats 40" 3 pockets with faux boa belt
1216908	6105.1092	Men lab coats 40" 3 pockets with side slits
1216908	6105.1093	Unisex scrub v neck top
1216908	6105.1093	Unisex dyed and white scrubs bottom
1216908	6105.1093	Unisex dyed and white scrubs v neck top
1216908	6105.1093	Women's white consultation jackets 33" style # 5740
1216908	6105.1093	1(a) women dyed top with one chest pocket & two pockets waist style # p.m.
1216908	6105.9000	1(b) women dyed bottom w/one cargo pocket one back pocket &w/self fab.drawcord stylpm
1216908	6105.9000	60% cotton 40% polyester knitted dyed long sleeve burnout crew neck tee
1216908	6105.9000	60% cotton 40% polyester dyed knitted men's fleece long sleeve hooded sweatshirts size,
1216908	6105.9000	70% cotton 30% polyester dyed & white knitted men's fleece sleeve hooded sweatshirts dyed
1216908	6105.9000	70% cotton 30% polyester knitted dyed fleece men's long hooded sweat shirts
1216908	6105.9000	Blended knitted fleece men's long sleeve hooded sweatshirt 0% cotton 30% poly
1216908	6105.9000	Men's bleached white consultation jacket 33" style # 5740
1216908	6105.9000	Men's bleached white lab coats 40" 3 pocket with side slit & faux back belt style # 5710
1216908	6105.9000	Mix cloth assorted size & colour
1216908	6105.9000	Women bleached white lab coats 40" 3 pockets with side slit & faux back belt style # 5720
1216908	6105.9000	Women shirt in woven fabric counts 26x26/99x50 dyed & women pant in woven fabric

Note: The table presents heterogeneity in product variety within the four different HS8 codes exported by the same firm
Source: Pakistan Customs dataset. These four HS8 products represent 32 different varieties.

Table 2: Decomposition of Exports and Product Varieties across Markets, 2013

Region	Exports		HS8 Products		Products+ Description	
	Value	%	#	%	#	%
	1	2	3	4	5	6
East Asia and Pacific	487.06	19.9	2,004	47.7	53,391	19.7
Europe and Central Asia	721.54	29.4	2,355	56.1	124,245	46.0
Latin America and Caribbean	55.96	2.3	630	15	11,356	4.2
Middle East & North Africa	301	12.3	2,705	64.4	47,255	17.5
North America	402.5	16.4	1,532	36.5	70,725	26.2
South Asia	321.52	13.1	1,953	46.5	32,742	12.1
Sub-Saharan Africa	161.49	6.6	1,675	39.9	21,726	8.0
All	2,572		4,200		270,335	

Notes: The table compares the number of HS8 products and export varieties with the fraction of exports absorbed in each region. The fraction of varieties in column (6) are quite comparable with the fraction of exports in column (2) but the fraction of HS8 products in column (4) are not.
Source: Constructed using administrative datasets for the financial year 2013.

The difference between product count and actual number of varieties becomes more evident in the decomposition of data across sectors and over time (Figures 1 and 2). As Figure 1 indicates, the HS8-level product count has been fairly consistent over time but the set of products within the HS8 level experienced a significant drop during the period of great trade collapse (2008-2009). This variation in diversification level over time is masked at the HS8 level. Not only the effect of global recession

(2008–2009), which adversely affected diversification, is masked at HS8 level, it also conceals the rebound in product variety in the period following the financial crisis.

Similarly, a great deal of variation between HS8 products and actual varieties is evident even at the sector level (Figure 2). This dissection of the data along two broad categories, agriculture and manufacturing, shows that, although manufacturing constitutes the largest volume of exports, agriculture also has a substantial share in the distribution. Overall, agriculture constitutes around 20% of national exports but manufacturing exports are much more diversified.

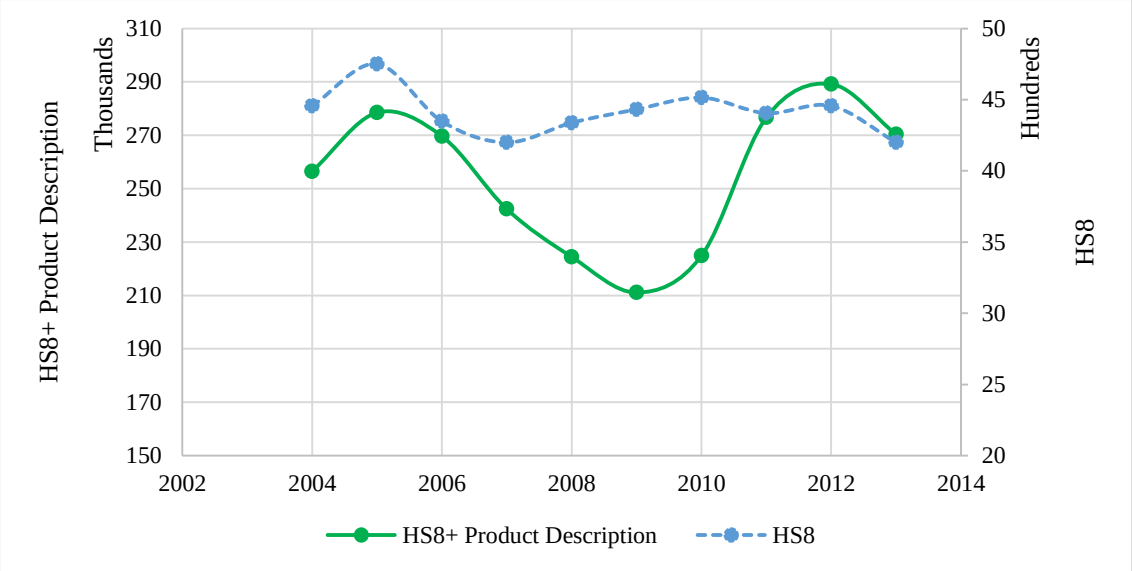
The drop in the number of products from 2012 onwards, indicates that the country has closed some product lines (Panel B of Figure 2). The existing literature suggests countries close some product lines as they travel along their development path (Imbs and Waczairg, 2003; Cadot et al., 2011); however, in the case of Pakistan, this turning point seems to be reached at a very early stage of development[10]. This shift seems to coincide with a reduction in the number of exporting firms in commodity and low-tech sectors and the corresponding increase in medium- and high-tech sectors as shown in Figure A2 in the appendix.

The analysis of diversification across markets according to level of development indicates that these two different approaches of measuring export diversification yield contrasting results (Figure 3). The HS8-level examination indicates that exports to developing countries are more diversified, whereas the micro-level measurements show that developed economies absorb a relatively large set of products and over time these differences have narrowed down. It follows that using product count as a diversification indicator might affect inferences about heterogeneity in the diversification level across markets. For instance, an exploration of diversification patterns at HS8 level across regions shows that prior to 2009 exports to Europe and Central Asia were most diversified. And in the later years the diversification level in this market dropped, whereas those to the countries of the Middle East and North America have risen (Figure A3). By contrast, examining the same pattern at micro level

[10] A similar early turning point has been observed in the Caribbean economies (Mohan, 2016).

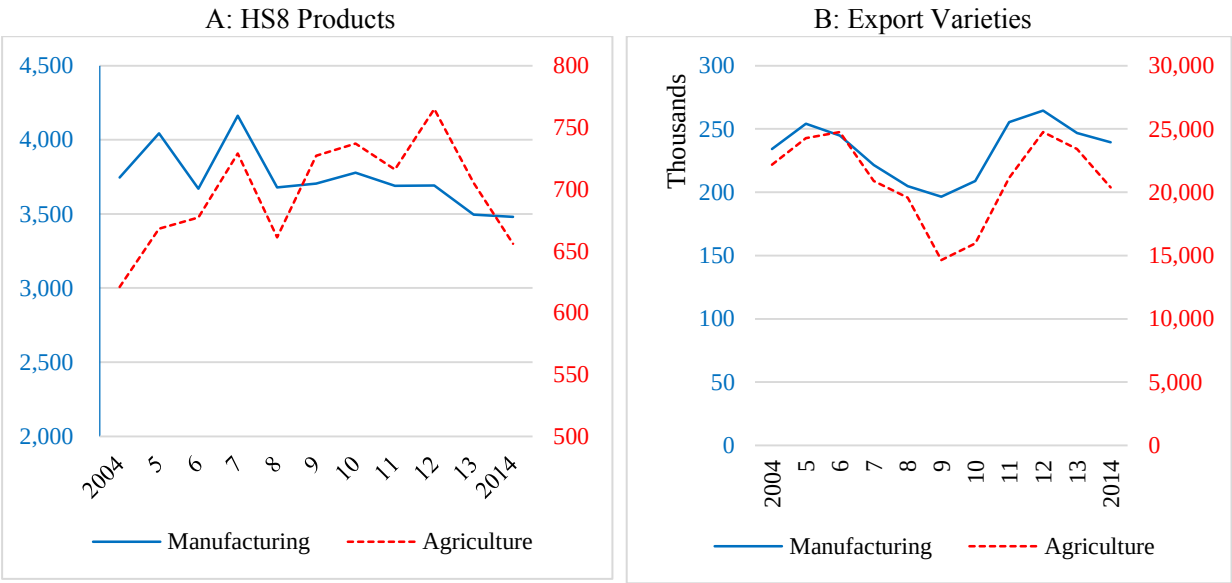
indicates that the level of diversification is higher for the markets of Europe and Central Asia and this pattern has been consistent over time (Figure A4). Therefore, to ensure a precise comparison across markets we account for the variation in the number of products within the HS8 level.

Figure 1: Divergence between Export Varieties and HS8 Products, 2003–2014



Source: Pakistan Customs dataset.

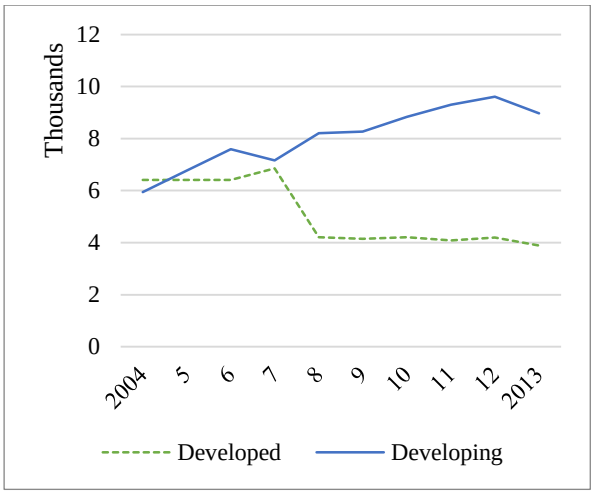
Figure 2: Diversification across Sectors and over Time, 2004–2014



Source: Constructed using administrative datasets

Figure 3: Diversification across Developed and Developing Countries

A: HS8 Products B: Export Varieties



Source: Pakistan Customs dataset.

2.3 Data on Factor Endowments

The data for the main explanatory variables – that is, factor endowment (physical capital, human capital and labour) – comes from Penn World Tables (PWT 08) (Feenstra et al., 2013). This new generation of PWT computes the physical capital from the data on initial assets, investment and depreciation for 167 countries. Similarly, it generates the Human Resources Index, based on average years of schooling of the population aged 15 and above and the assumed returns to education, as discussed in Barro and Lee (2013) and Psacharopulos (1994). The data on land endowments and population comes from the World Development Indicators (WDI). Gravity model variables are retrieved from the CEPII[11] and gross domestic product (GDP) data is downloaded from the open data sources of the World Bank. The remaining information on other economic variables is retrieved from the open data sources of the World Trade Organization (WTO) and the World Bank.

Figure 4 plots the graphs of average differences in factor endowments between Pakistan and its export markets over time. The export markets are grouped in two categories, developed and developing, following UNCTAD's classification of economies. As the charts indicate, inequality in human and physical capital endowments is relatively higher for developed markets; however, the difference in land endowment is larger for developing economies. Moreover, factor endowments inequality appears to have increased over time but the differences across developed and developing markets have narrowed. This trend warrants the deconstruction of diversification patterns along temporal dimensions

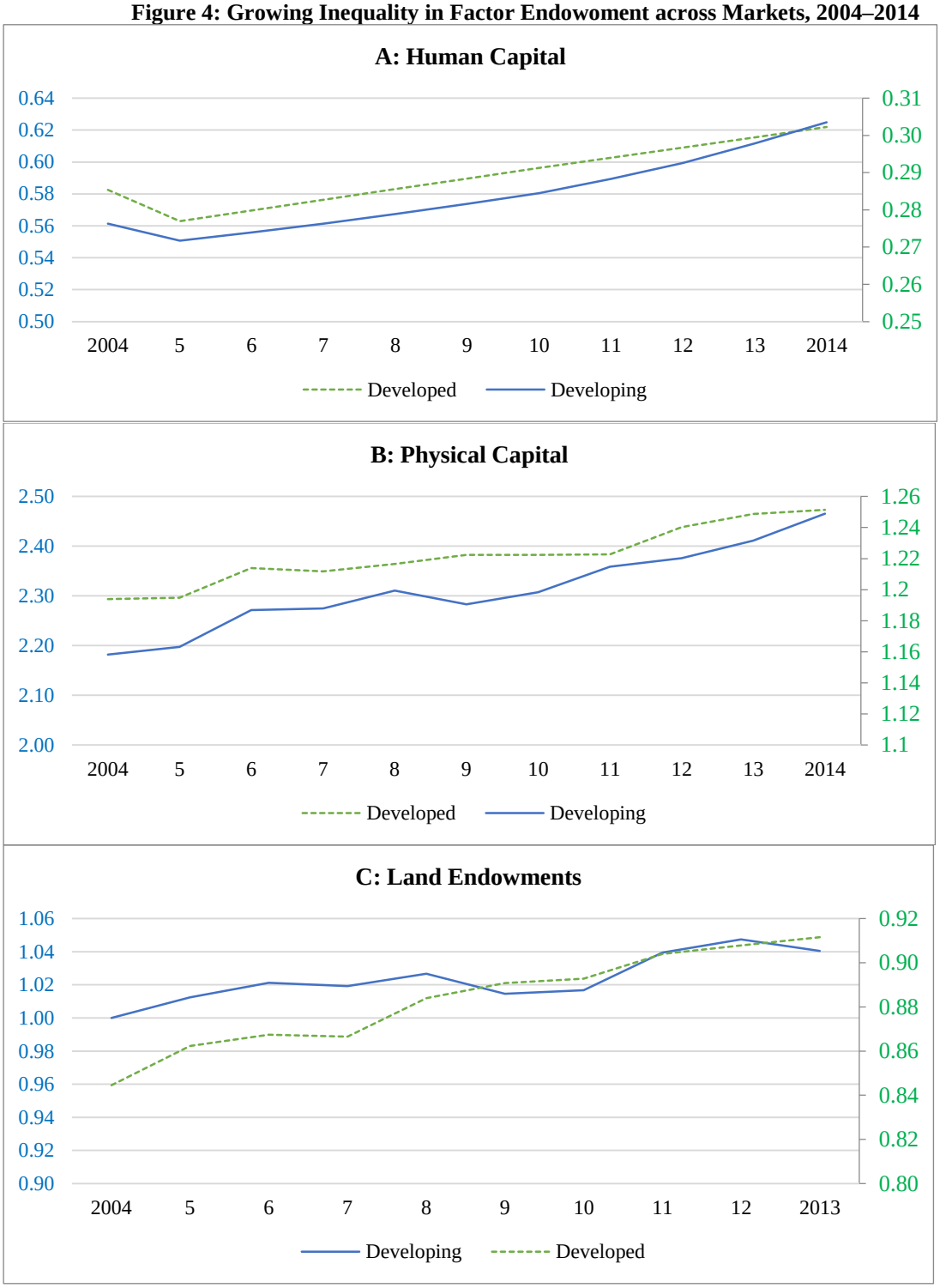
[11] http://www.cepii.fr/

to understand the underlying heterogeneity. It further suggests that the effect of human capital and that of physical capital on export diversification could be similar whereas that of land endowment may differ across these market groups.

2.4 Linkages between Factor Endowments and Export Diversification

Figure 5 presents the relationship between factor endowments and EM of products across markets. As the charts show, inequality in physical and human capital with trading partners is positively associated with the set of exported products and inequality in land endowments is negatively associated. Overall, this pattern appears to contradict the predictions of Regolo (2013), which finds that exports are more diversified for similarly endowed countries (South–South and North–North) than they are for differently endowed ones (North–South or South–North). The following sections examine this relationship between inequality in factor endowments and diversification at a firm level in a regression framework.

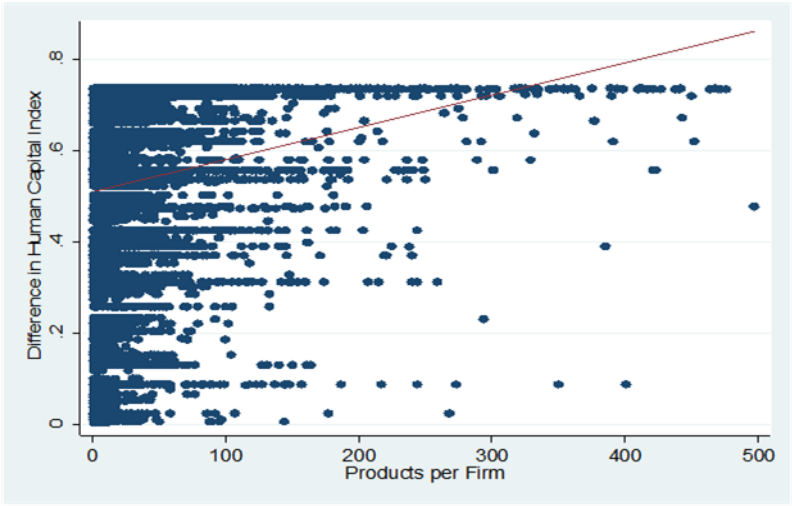
Figure 4: Growing Inequality in Factor Endowoment across Markets, 2004–2014



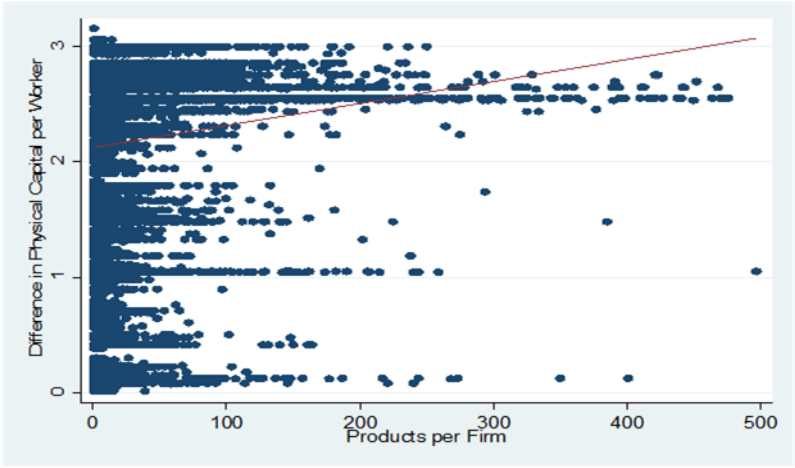
Note: The factor endowments on the y-axis are measured in absolute difference of log.
Source: Authors' construction using data from PWT08

Figure 5: Relationship between Factor Endowments and Export Diversification

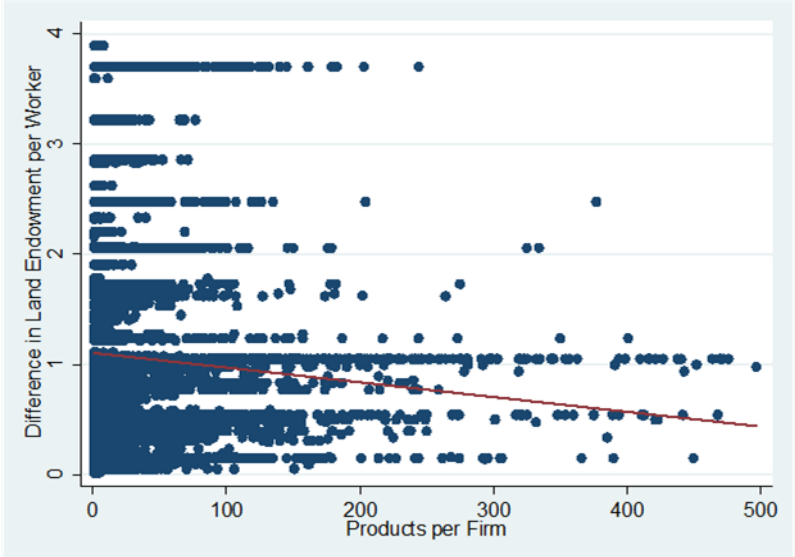
A: Inequality in Human Capital



B: Inequality in Physical Capital



C: Inequality in Land Endowments



Note: The factor endowments on the y-axis are measured in absolute difference of log.
Source: Authors' calculations.

3 Estimation Strategy and Main Results

We empirically investigate the relationship between diversification and endowments inequality with three different specifications. They comprise (1) estimations at the firm level by using trade diversification indicator, (2) adding controls for trade costs and other gravity variables, and (3) replicating the estimations at market level by using the Herfindhal-Hirschman Index (HHI) of export concentration. These specifications provide robust evidence of linkages between endowment inequality and export diversification.

3.1 Specification I: Firm-Level Estimations

We regress the level of diversification on bilateral differnces in factor endowments using the following estimation equation.

$$ln(XD_{ijkt}) = \beta_0 + \beta_1(DIFF^{K/L})_{jt} + \beta_2(DIFF^{H/L})_{jt} + \beta_3(DIFF^{T/L})_{jt} + \beta_4(\text{ln GDP/capita})_j + \alpha_i + \gamma_k + \lambda_t + \varepsilon_{ijkt} \quad \ldots(1)$$

where *i* denotes firm, *j* market (trading partners of Pakistan), *k* product and *t* time (year).

The dependent variable, XD_{ijkt} is total number of varieties exported by a firm in the given year by sector (N_{ijkt}) relative to the average of the same for all firm across all markets ($\bar{N}_{ijkt}$). This diversification indicator measures export intensity at a firm level in terms of extensive margins. A higher value of 'XD' means a wider set of exported products, indicating more diversified exports, whereas lower values indicate relatively concentrated structure. For constuction of this indicator we identify products at the HS8 level of disaggregation and take into account the differenciation of varieties within the HS8 categories as discussed in Section 2.2. We compute this indicator for 98 product groups (across markets and over time) following the two-digit HS classification. These broad categories, besides allowing incorporation of controls for time invariant and time-varing omitted variables for products, permit decomposing the estimation results across sectors., which is important to observe the heterogeneity of effect across agriculture and manufacturing.

This approach of measuring export diversification in relative terms[12] has several advantages over using other indicators. First, by using average number of products exported by all firms as a denominator, the product diversity is measured relative to the average of all firms exporting in that particular sector (Michler and Josephson, 2016). Second, we update the denominator each year to allow for changes in production and the trading environment. Third, by computing the indicator at the sector level, we capture heterogeneity across firms within the same sector.

The explanatory variables $DIFF^{E/L}_{jt} = Abs[\ln(E/L)_t - \ln(E/L)_{jt}]$ for $E \in \{K; H; T\}$, and K, H, T, and L denote phyical, human, land and labour endowments, respectively. The coefficients of interest on these variables, β_1 and β_2, are expected to be positive, whereas β_3, is expected to be negative as suggested in the graphical analysis (Figure 5).

This identification strategy exploits the variation in factor endowments across countries and over time and the heterogeneity of the set of products exported by firms across markets as well as over time. A wide variation in the EM of products and factor endowments across market and over time (as demonstrated in Section 2) provides a nice setting to investigate the question using econometric techniques. A similar functional form of the regression equation was developed in Regolo (2013) for exploring the linkages between diversification and factor endowments at the product level. I apply the same specification at a higher level of disaggregation in a different empirical setting. Due to very large sample, the firm-level estimations generate relatively precise estimates and allow controls for potentially time-invariant or time variant omitted factors.

α_i, γ_k, λ_t are a set of fixed effects for firms, products and time. The firm and product fixed effect soak up potentially omitted factors that are time-invariant. The year fixed effects, λ_t, accounts for year-specific effects that are common for all firms. Estimating this fixed effect model assumes that there are major differences across firms and products but they vary little over time. To account for these omitted factors, the fixed effect estimator transforms all the variables into deviation from their mean.

[12] Most of the the commonly used export concentrataion indices (Thiel, Gini, Herfindhanl) are in the ratio form but their computation and interpretation is complicated.

In robustness checks, however, we use time-varying fixed effects for firms and products. Firm-year fixed effects soak up any changes in the firms' productivity or technological improvements over time that can influence their exports to various markets. Similarly, product-year fixed effects account for fluctuation in product-specific factors. The control for GDP account for the variation in the level of economic development across Pakistan's trading partners.

All estimations are in a logarithmic form and the estimation method is ordinary least squares (OLS); however, we apply non-linear a Pseudo Poisson Maximum Likelihood (PPML) estimator for a robustness test (Silva and Santos, 2006). To account for autocorrelation, standard errors are clustered at market-year level, the level of variation of independent variables. We use the Stata command 'reghdfe', as suggested in Guimaraes and Portugal (2010), to estimate equation (1) with high dimensional fixed effects.

Table 3: Main Estimation Results: Effect of Factor Endowments on Diversification Level
The dependent variable is log of export diversification indicator

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Human Capital	0.033***	0.051***	0.059***	0.063***	0.058***	0.057***	0.065***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Physical Capital	0.031***	0.041***	0.053***	0.055***	0.059***	0.058***	0.063***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Land Endowments	0.003**	-0.010***	-0.016***	-0.018***	-0.019***	-0.022***	-0.021***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
GDP/capita	0.061***	0.058***	0.070***	0.068***	0.071***	0.078***	0.079***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Prod. FE		Y		Y	Y	Y	
Firm FE			Y	Y	Y		
Time FE					Y		
Firm-year FE						Y	Y
Prod-year FE							Y
R^2	0.003	0.048	0.158	0.205	0.206	0.272	0.284
Observations	1,020,257	1,020,257	1,020,257	1,020,257	1,020,257	965,535	965,535

Note: The estimations show standardised beta coefficients with robust standard errors in parentheses. These coefficients were obtained using Stata 13 SE. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Y indicates the inclusion of fixed effects. Columns (6) to (8) have fewer observations as singleton observations are dropped in the estimations. The dependent variable, XDijkt, is total number of varieties exported by a firm (Nijkt) in the given year by sector relative to the average of the same for all firm across all markets ($\bar{N}$ijkt).

Table 3 reports the main estimation results. These estimates provide a conditional correlation of export diversification indicator with the bilateral differences in factor endowments. Column (1) presents pure variation in the data and columns (2) to (4) add fixed effects for products, firms and time to account for potentially omitted variables. Column (5) examines the combined effect of all time-invariant fixed effects, which also generate similar results. As the estimates show, bilateral

differences in human and physical capital are positively associated with the diversification variable, whereas the reverse is true for land endowments, and this effect is statistically significant at the 1% significance level. The positive coefficients in rows (1) and (2) suggests exports to differently endowed markets in terms of physical and human capital are more diversified. By contrast, the negative coefficients in row (3) suggests that trade with similarly endowed markets in terms of land endowments is more diversified.

The estimates in column (5) are used as baseline. Columns (6) and (7) present initial robustness checks by introducing controls for time-varying omitted variables for firms and products.

To make the interpretation of the results easier, regression coefficients in Table 3 are standardised. The coefficients in column (7) are interpreted as follows. If a firm re-orients its exports to a trading partner that is more different in terms of human capital by '1' standard deviation, its exports to that market will be more diversified by 1.30% (1.30%=0.065x0.20), where 0.20 is the standard deviation of human capital), keeping all other things constant. Similarly, if differences in physical capital or natural endowments increase by '1' standard deviation from their mean, the diversification changes by 4.77 % (0.063x0.77) and -2.1% (-0.021x1.018), respectively.

3.2 Specification II: Gravity Controls

The notion that factor endowment fully explains the export diversification pattern is naïve, of course. Intra-industry trade models and a burgeoning literature on trade costs indicate that many other factors and policies affect trade patterns. To examine the effect of trade cost on the level of diversification, we augment the above equation with standard gravity model variables, such as distance between trading partners, GDP, common official language and presence or absence of a trade agreement. These variables are derived from the CEPII dataset and entail the definitions therein. The modified regression equation takes the following form.

$$ln(XD_{ijkt}) = \beta_0 + \beta_1(DIFF^{K/L})_{jt} + \beta_2(DIFF^{H/L})_{jt} + \beta_3(DIFF^{T/L})_{jt} + \beta_4 \, ln(dist.)_j + \beta_5 \, ln(GDP/capita)_{jt} + \beta 6 \, (lang.)_j + \beta_7 \, (PTA)j + \alpha_i + \gamma_k + \lambda_t + \epsilon_{ijkt} \quad \ldots(2)$$

The effect of remoteness (ln dist.) is expected to be negative, whereas that of other regressors, such as GDP, common language (lang.) and trade agreements (PTA), is expected to be positive.

Table 4 presents the estimation results of equation (2). Column (1) shows that distance to market of trading partner negatively affects diversification as transportation costs are higher for distant markets. Column (2) controls for the GDP of trading partners. The addition of GDP per capita in regression alleviates the concern that the three variables capturing the differences in factor endowments do not capture the level of development of trading partners. The results indicate that trade is more diversified with rich trading partners as rich consumers want more varieties of products. Similarly, common official language and trade agreement are associated with a higher level of diversification as they reduce the cost of information and tariffs, respectively (column 3).

Table 4: Effect of Trade Costs and Factor Endowments on Diversification Level

The dependent variable is log of export diversification indicator

	(1)	(2)	(3)	(4)
Human Capital	0.069***	0.026***	0.019***	0.048***
	(0.009)	(0.010)	(0.011)	(0.011)
Physical Capital	0.051***	-0.002	0.016***	0.035***
	(0.002)	(0.003)	(0.003)	(0.003)
Land Endowments	-0.017***	0.004***	-0.015***	-0.021***
	(0.001)	(0.001)	(0.001)	(0.001)
ln (Distance)	-0.010***	-0.012***	-0.030***	
	(0.003)	(0.003)	(0.003)	
ln (GDP/capita)		0.134***	0.119***	
		(0.001)	(0.001)	
C. Language (1, 0)			0.051***	
			(0.003)	
FTA (1, 0)			0.009***	
			(0.008)	
ln (BTC)				-0.074***
				(0.006)
R^2	0.205	0.211	0.212	0.214
Observations	1,020,257	1,020,257	1,020,257	583,307

Note: The estimations show standardised beta coefficients with robust standard errors in parentheses. These coefficients were obtained using Stata 13 SE. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The estimations include fixed effects for firms, products and time. BTC stands for bilateral trade costs. The dependent variable, XDijkt, is total number of varieties exported by a firm (Nijkt) in the given year by sector relative to the average of the same for all firm across all markets ($\bar{N}$ijkt).

The estimated coefficients on human and physical capital in columns (2) and (3) drop in magnitude on incorporating various gravity controls in the regression framework. This indicates that trade costs are an important predictor of level of diversification. The effect of GDP of trading partner is much higher than that of geographical distance. Interestingly still, the coefficient of interest on human and physical capital is positive, whereas that on land endowments is negative, and all these coefficients are

statistically significant at the 1% significance level. These results support the findings of Amurgo-Pacheco and Pierola (2008) in that the diversity of trade is related to distance and size of export markets.

Columns (4) uses an alternative indicator of trade costs. The reason is that Limao and Venables (2001) argue that variation of costs of international business across countries does not necessarily depend on distance. Moreover, using geographical distance between trading partners as a proxy for trade costs ignores the frictions induced by other sources, such as non-tariff measures and procedural obstacles, which too impose additional burdens on exporting firms. Similarly, many other studies suggest that improved shipping line connectivity overcomes the resistance posed by geographical remoteness (Arvis et al., 2016). To account for these issues, we use the bilateral trade cost (BTC) indicator from the World Bank (column 4), which specifically factors in the costs of numerous other barriers. As the results in column (4) indicate, the change of trade cost indicator does not affect the baseline results for factor endowments.

3.3 Specification III: Herfindhal-Hirschman Index (HHI) of Export Concentration

We compute this HHI across markets and replicate the same estimations by using the HHI as an alternative dependent variable. The HHI as a measure of the degree of market concentration by countries is expressed as follows (UNCTAD, 2012).

$$H_j = \frac{\sqrt{\sum_{i=1}^{n} \left(\frac{x_{ij}}{X_j}\right)^2} - \sqrt{1/n}}{1 - \sqrt{1/n}} \qquad \text{……..…(3)}$$

where Hj is concentration index of Pakistan's export to market *j*, x_{ij} is value of exports from Pakistan to trading partner *j* in product *i,* n is number of products and

$$X_j = \sum_{i=1}^{n} x_{ij}$$

The export concentration index for Pakistan has a mean of 0.46 and a standard deviation of 0.24, suggesting highly concentrated export structure. The modified regression equation is as follows.

$$ln(HHI_{jt}) = \beta_0 + \beta_1(DIFF^{K/L})_{jt} + \beta_2(DIFF^{H/L})_{jt} + \beta_3(DIFF^{T/L})_{jt} + \lambda_t + \varepsilon_{jt} \quad \ldots\ldots(4)$$

where $_j$ denotes market (trading partners of Pakistan) and t time (year).

As this indicator measures level of concentration, not diversification, we expect opposite signs (to that of baseline estimation) on coefficients of various regressors of interest (Table 5). The negative coefficients on human and physical capital indicate that the greater differences in factor endowments are associated with less concentrated (that is more diversified) exports. The effect of various trade cost indicators (such as common official language, FTA) is negative as these measures reduce concentration of exports. GDP of trading partner has similar effects but remoteness of markets has an opposite but relatively weak effect on export concentration. Overall, these estimates support the baseline results.

Table 5: Estimations with Export Concentration Index

The dependent variable is log of HHI

	(1)	(2)	(3)
Human Capital	-0.208***	-0.208***	-0.064***
	(0.022)	(0.012)	(0.011)
Physical Capital	-0.333***	-0.326***	-0.139***
	(0.024)	(0.013)	(0.013)
Land Endowments	0.088***	0.090***	0.061***
	(0.022)	(0.011)	(0.013)
ln (Distance)		0.006	0.030*
		(0.016)	(0.015)
C. Language (1, 0)			-0.064***
			(0.020)
FTA (1, 0)			0.001
			(0.086)
ln (GDP)			-0.551***
			(0.005)
R2	0.207	0.203	0.432
N	2,644	2,644	2,644

Note: The estimations show standardised beta coefficients with robust standard errors in parentheses. These coefficients were obtained using Stata 13 SE. * p < 0.10, ** p < 0.05, *** p < 0.01. The estimation includes fixed effects for time.

In the absence of any firm-level study on the effect of factor endowment inequality on the diversification of exports along EM, these results are not directly comparable with earlier work. However, unlike Regolo's (2013) prediction (that measures diversification along IM using the Thiel Concentration Index), our diversification indicator is positively correlated with bilateral differences in factor endowments in human and physical capital. However, the correlation is negative with the

inequality in land endowments. This means greater differences in human and physical capital per worker are associated with increased diversification, whereas increased similarity in natural endowments is has similar effect. Second, these results show that inequality in physical capital has a much larger effect than does inequality in human capital or land endowments

4 Heterogeneity Analysis and Robustness Checks

This section decomposes the estimated conefficient on factor endowments across products and markets. It also conducts a series of other robustness tests by using alternative measures of diversification, estimation tecniques and dataset.

4.1 Heterogeneity across Sectors and Markets

Given the significance of agriculture in overall exports, Table 6 deconstruct the baseline estimates[13] into two broad sectors, agriculture and manufacturing. It shows that the effect of differences in endowments is much higher for diversification of manufactured goods compared with that for agricultural products. Moreover, the effect of various endowments is heterogeneous across sectors: differences in human capital have large effect for diversification of manufactured goods compared with that of agricultural exports. Similarly, dissimilarity in land endowments positively affects the diversification of agricultural exports but negatively influences the same for manufactured goods.

Table 6: Heterogeneity of the Effect of Factor Endowments across Sectors

The dependent variable is log of export diversification indicator

	(1)
Manufacturing #	
Human Capital	0.064***
	(0.001)
Physical Capital	0.059***
	(0.001)
Land Endowment	-0.023***
	(0.001)
Agriculture #	
Human Capital	0.018***
	(0.004)
Physical Capital	0.054***
	(0.004)
Land Endowment	0.011***
	(0.003)
R^2	0.206
Observations	1,020,257

Note The estimations show standardised beta coefficients with robust standard errors in parentheses. These coefficients were obtained using Stata 13 SE. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The estimation includes fixed effects for firms, products and time. The dependent variable, XDijkt, is total number of varieties exported by a firm (Nijkt) in the given year by sector relative to the average of the same for all firm across all markets ($\bar{N}$ijkt).

[13] estimated coefficients in column (5) of Table 3

Table 7: Heterogeneity of the Effect of Factor Endowments across Markets

The dependent variable is log of export diversification indicator

	(1)
Developed Countries #	
Human Capital	0.121***
	(0.002)
Physical Capital	0.149***
	(0.004)
Land Endowment	-0.037***
	(0.003)
Developing Countries #	
Human Capital	-0.006**
	(0.002)
Physical Capital	0.055***
	(0.002)
Land Endowment	0.002
	(0.001)
R^2	0.206
Observations	1,020,257

Note: The estimations show standardised beta coefficients with robust standard errors in parentheses. These coefficients were obtained using Stata 13 SE. * p < 0.10, ** p < 0.05, *** p < 0.01. The estimation includes fixed effects for firms, products and time. The dependent variable, XDijkt, is total number of varieties exported by a firm (Nijkt) in the given year by sector relative to the average of the same for all firm across all markets ($\tilde{N}$ijkt).

Table 8: Heterogeneity of the Effect across Markets and Sectors Combined

The dependent variable is log of export diversification indicator

	(1)
Developed x Manufacturing x	
Human Capital	0.132***
	(-0.003)
Physical Capital	0.168***
	(-0.005)
Land Endowments	-0.045***
	(-0.003)
Developed x Agriculture x	
Human Capital	0.092***
	(-0.011)
Physical Capital	-0.161***
	(-0.021)
Land Endowments	0.044***
	(0.013)
Developing x Manufacturing x	
Human Capital	-0.005*
	(-0.003)
Physical Capital	0.052***
	(-0.002)
Land Endowments	0.002*
	(-0.002)
Developing x Agriculture x	
Human Capital	-0.013*
	-0.007
Physical Capital	0.087***
	(-0.005)
Land Endowments	-0.001
	(-0.004)
R^2	0.206
Observations	1,020,257

Note: The estimations show standardised beta coefficients with robust standard errors in parentheses. These coefficients were obtained using Stata 13 SE. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The estimations include fixed effects for firms, products and time. The dependent variable, XDijkt, is total number of varieties exported by a firm (Nijkt) in the given year by sector relative to the average of the same for all firm across all markets ($\tilde{N}$ijkt).

Table 7 decomposes the baseline results[14] for two broad market groups: developed and developing. As the estimates show, the effect of inequality in endowments is much higher for diversification of exports to developed countries. Moreover, the differences in human capital are positively associated with diversification level in developed economies and negatively associated with the same for developing markets. It indicates that export to developing countries that are similarly endowed in terms of human capital are more diversified, while the reverse is true for developed markets.

Table 8 examines the effect of differences in endowments on the diversification level across markets and sectors in tandem. It shows that the differential effect of human capital across markets holds at sector level also. For instance, differences in human capital are positively associated with diversification to developed countries, both in agriculture and in manufacturing, but similarity in human capital has the same effect for exports to developing countries. Similarly, the effect of physical capital and land endowments also varies across markets and sectors.

4.2 Further Robustness Tests

This sub-section tests the robustness of baseline results (model 5 in Table 3) by using alternative measures of diversification, replicating the same estimations at higher levels of aggregation of data, deconstructing the effect of endowment inequality over time, and by using alternative estimation techniques and data sources.

Alternative Dependent Variables

Table 9 replicates the baseline estimations (equation 1) with alternative measures of diversification. Column (1) uses the number of HS8 products as an indicator of diversification, as in Persson and Wilhelmsson (2016) and Dennis and Shepherd (2011)[15]. Column (2) uses number of firms per market as an alternative dependent variable. In both regressions, the estimated coefficients on factor

[14] estimated coefficients in column (5) of Table 3

[15] Although this indicator ignores the differentiation of products within HS8 level of disaggregation, it has been used in the earlier studies due to data limitations.

endowments bear signs similar to those in benchmark results and the results are statistically significant at the 1% significance level.

Table 9: Alternative Measures of Export Diversification

Dep. Variables	(1) HS8 Products	(2) Firms
Human Capital	0.051***	0.040***
	(0.001)	(0.002)
Physical Capital	0.045***	0.082***
	(0.001)	(0.002)
Land Endowments	-0.017***	-0.009***
	(0.001)	(0.002)
R^2	0.153	0.471
Observations	1,020,257	286,095

Note: The dependent variables are number of products per firm (column 1) and firms per product (column 2) by destination in log. The estimations show standardised beta coefficients with robust standard errors in parentheses. These coefficients were obtained using Stata 13 SE. * p < 0.10, ** p < 0.05, *** p < 0.01. Column (1) includes fixed effects for firms and time and column (2) includes those for products and time.

Computing Export Diversification Indicator at Higher Levels of Product Aggregation

As some firms export multiple products from various sub-sectors, limiting the product count within the 98 base categories might influence the results because exports of the same firm from the other sub-sections would be treated as a separate observation, which might underestimate the response of dependent variable. Therefore, Table 10 replicates the same estimations at higher levels of aggregation. Column (1) groups the data in 16 broad sectors and column (2) treats all products as one group and column (3) removes temporal dimension of the data. The disadvantage of this approach is that fixed effects for products and time cannot be incorporated in columns (2) and (3). These results indicate that the level of aggregation matters for the magnitude of the coefficients but its sign and significance level remains quite stable.

Table 10: Estimations at Higher Levels of Aggregation

The dependent variable is log of export diversification indicator

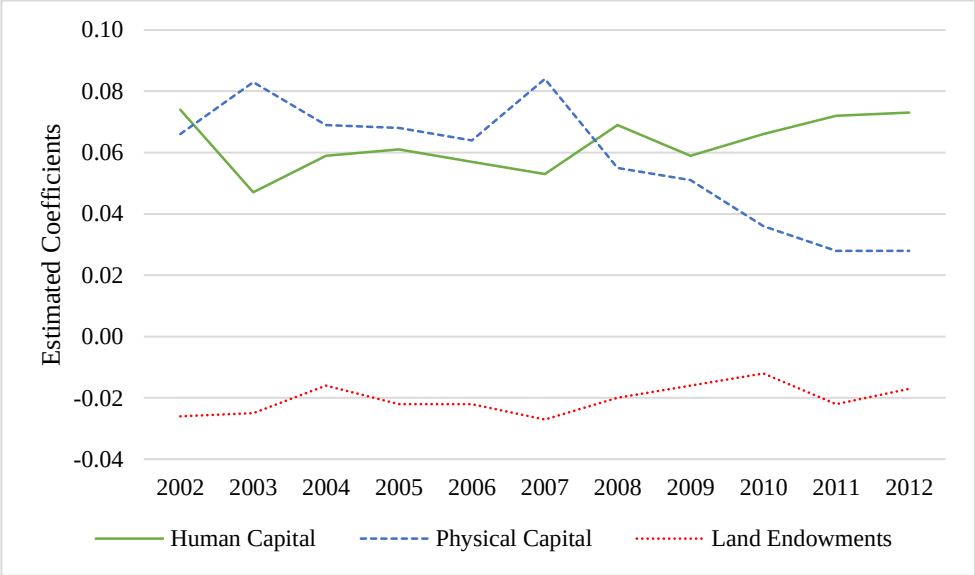
	Sixteen Sectors	Single Sector	Single Period
	(1)	(2)	(3)
Human Capital	0.064***	0.084***	0.074***
	(0.002)	(0.002)	(0.003)
Physical Capital	0.072***	0.094***	0.142***
	(0.002)	(0.002)	(0.003)
Land Endowments	-0.024***	-0.031***	-0.036***
	(0.001)	(0.001)	(0.002)
R^2	0.232	0.280	0.289
Observations	822,295	581,708	210,213

Note: The estimations show standardised beta coefficients with robust standard errors in parentheses. These coefficients were obtained using Stata 13 SE. * p < 0.10, ** p < 0.05, *** p < 0.01. Column (1) contains the results at sector level, column (2) for all products at market level and column (3) for the period average, 2000–2013. The reduced number of observations in column (2) and (3) owes to the higher level of aggregation.

Decomposition of Effect along Time

Figure 6 plots the decomposed results along temporal dimensions to identify the changes over the study period and to make sure that a particular year might not be driving the results. The detailed estimates are reported in the appendix (Table A2). As Figure 6 indicates, the effect of physical and human capital is positive in all the years but that of land endowments is negative, which further supports the baseline estimates.

Figure 6: Deconstructing the Effect of Factor Endowments over Time



Notes: The figure plots the estimates coefficients on endowment inequality over time. It highlights the increasing contribution of human capital in diversification compared with that of physical capital or natural endowments, especially in the post-2008 period.

These estimates highlight the increasing contribution of human capital in diversification compared with that of physical capital or natural endowments, especially in the post-2008 period. This period is associated with the drop in some product lines (Figures 1 and 2). Moreover, in this period the orientation of exports to developing countries has increased (Figure 3) and the country seems to have comparative advantage in human capital relative to these markets.

Alternative Estimation Approach

Using a product count as a dependent variable, we replicate the base line regressions with Poisson Pseudo Maximum Likelihood (PPML) estimator. The estimated coefficients on all regressors bear the expected signs and are statically significant, which further supports the baseline results (Table 11).

Table 11: Estimations Using the PPML Estimator

The dependent variable is a number of products per firm by destination

	(1)
Human Capital	0.295***
	(0.037)
Physical Capital	0.052***
	(0.005)
Natural Endowments	-0.014***
	(0.005)
FE (Firm-Prod-Market)	Y
FE (Time)	Y
N	680,173

Note: Standard errors in parentheses, * p < 0.10, ** p < 0.05, *** p < 0.01. The number of observations is smaller as many dropped because they pertained only to one group. The individual fixed effects for firms and products could not be included owing to computational problems but the combined fixed effects for firms-product-markets were added in regression using the Stata xtpoisson command.

Effect of Policy Variables

Table 12 examines the effect of policy-related variables on diversification in addition to that of factor endowments. These variables are taken from the WDI and take the definitions contained therein. In all these regressions, the sign and significance level on the factor endowments remains stable. Moreover, all the explanatory variables bear the expected signs. For instance, trade openness increases diversification, as does the improvement in connectivity (Logistics Performance Index and the Liner Shipping Connectivity Index). Higher values of these indices indicate a reduction in trade costs. Similarly, the effect of the high cost of imports in destination markets, the requirement of many documents and the high cost of starting up business reduce diversification.

Table 12: Effect of Policy-Related Variables on Diversification Level

The dependent variable is log of export diversification indicator

	(1)	(2)	(3)	(4)	(5)	(6)
Human Capital	0.018***	0.031***	0.042***	0.056***	0.061***	0.042***
	(0.004)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
Physical Capital	0.007*	0.043***	0.069***	0.053***	0.057***	0.067***
	(0.004)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
Natural Endowments	-0.023***	-0.021***	-0.021***	-0.018***	-0.019***	-0.014***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Trade openness	0.485***					
	(0.019)					
Logistics Performance Index		0.084***				
		(0.002)				
Liner Shipping Connectivity Index			0.009***			
			(0.001)			
Cost to Import				-0.003		
				(0.003)		
Documents to Import					-0.036***	
					(0.004)	
Cost of business start-up						-0.110***
						(0.002)
R2	0.256	0.220	0.225	0.223	0.223	0.209
N	770,109	770,109	717,506	702,487	702,487	1,018,688

Note: The estimations show standardised beta coefficients with robust standard errors in parentheses. These coefficients were obtained using Stata 13 SE. * p < 0.10, ** p < 0.05, *** p < 0.01. The number of observations varies because of data limitations. The estimations include fixed effect for firms and products as well as a control for time trend.

External Validity

Finally, we check the validity of these results for other developing countries. We replicate the same estimations on the Exporter Dynamic Database for 52 developing economies. This cross-country analysis allows for controlling multilateral resistance but does not permit including fixed effects for firms and products. Despite this, the estimations support the baseline results, suggesting that this pattern of association between inequality in factor endowments and diversification is not peculiar to Pakistan.

Table 13: External Validity – Estimations using the Exporter Dynamic Database for 52 Countries

	(1)	(2)
Dependent variables	Mean HS6 products per firm	Median HS6 products per firm
Human Capital	0.101***	0.075*
	(0.028)	(0.039)
Physical Capital	0.036**	0.062**
	(0.017)	(0.025)
Natural Endowments	-0.058**	-0.061*
	(0.028)	(0.039)
FE (Origin-Destination)	Y	Y
FE (Time)	Y	Y
R2	0.740	0.506
N	43,692	43,692

Note: The estimations show standardised beta coefficients with robust standard errors in parentheses. These coefficients were obtained using Stata 13 SE. * p < 0.10, ** p < 0.05, *** p < 0.01.

5 Conclusion

This study examines whether inequality in factor endowments between trading partners affects the diversification of exports along extensive margins at a firm level. The availability of highly disaggregated data, which contains the description of goods at the transaction level, permits accounting for the differentiation of varieties within the 8-digit level of Harmonised System (HS) and thus enables us to measure export diversification at a micro level.

The study finds that endowment inequality, specifically the differences in physical and human capital, increases diversification, whereas inequality in land endowments reduces it. Moreover, the comparative advantage appears to be quite dynamic. The role of physical capital endowment is declining over time and that of human capital is increasing. The study further shows that low trade costs positively affect diversification. The level of diversification rises with the GDP of the trading partner, with sharing a common official language and with having a trade agreement but it drops with

remoteness. The results are robust to a battery of sensitivity tests. Moreover, a replication of estimations on the Exporter Dynamic Database corroborates their validity for other developing countries. The analysis emphasises the need to account for the differentiation of varieties within narrowly defined product categories. It shows that analysis based on product count could mask stark differences in the diversification level across sectors and markets.

To the best of my knowledge, the existing theoretical literature does not provide clear guidance about the relationship between the number of export varieties and the dissimilarity of factor endowments in the destination country. The stylised facts documented in this work provide ground for the extension of trade models to capture this pattern in the data.

Since export diversification has assumed considerable importance from development policy perspectives, these findings imply that choice of trading partners matters for increasing diversification. It shows that, at a micro-level, exports to developed markets are more diversified than those to developing. Besides other factors, these markets are more heterogeneous in terms of factor endowments and thus attract a wider variety of goods. Second, policy intervention to reduce trade costs could increase diversification level. Finally, developing economies might need to shift their development focus from investing in physical capital to developing human capital because the contribution of human capital in diversification of exports is increasing over time.

6 References

Ali, S. and Milner, C. 2015. Trade costs and the composition of developing countries' exports. Globalization and Economic Policy Working Paper 13/2015. Nottingham: University of Nottingham.

Amurgo-Pacheco, A. and Pierola, M. D. 2008. Patterns of export diversification in developing countries: intensive and extensive margins. Policy Research Working Paper 4473. Washington, DC: World Bank.

Anderson, J. E. and Van Wincoop, E. 2004. Trade costs. *Journal of Economic Literature*, 42, 691–751.

Arvis, J.-F., Duval, Y., Shepherd, B., Utoktham, C. and Raj, A. 2016. Trade costs in the developing world: 1996–2010. *World Trade Review*, 15, 451–74.

Bacchetta, M., Beverelli, C., Cadot, O., Fugazza, M., Grether, J.-M., Helble, M., Nicita, A. and Piermartini, R. 2012. A practical guide to trade policy analysis. Geneva: WTO.

Barro, R. J. and Lee, J. W. 2013. A new data set of educational attainment in the world, 1950–2010. *Journal of Development Economics*, 104, 184–98.

Cadot, O., Carrère, C. and Strauss-Kahn, V. 2011. Export diversification: what's behind the hump? *Review of Economics and Statistics*, 93, 590–605.

Cadot, O., Carrère, C. and Strauss-Kahn, V. 2013. Trade diversification, income, and growth: what do I know? *Journal of Economic Surveys*, 27, 790–812.

Chowdhury, M. T. H., Bhattacharya, P. S., Mallick, D. and Ulubaşoğlu, M. A. 2014. An empirical inquiry into the role of sectoral diversification in exchange rate regime choice. *European Economic Review,* 67, 210-227.

Defever, F., Heid, B. and Larch, M. 2015. Spatial exporters. *Journal of International Economics*, 95, 145–56.

Dennis, A. and Shepherd, B. 2011. Trade facilitation and export diversification. *The World Economy*, 34, 101–22.

Feenstra, R. C., Inklaar, R. and Timmer, M. P. 2015. The next generation of the Penn World Table. *The American Economic Review*, 105, 3150–82.

Fontagné, L., Gaulier, G. and Zignago, S. 2008. Specialization across varieties and North–South competition. *Economic Policy*, 23, 52–91.

Guimaraes, P. and Portugal, P. 2010. A simple feasible procedure to fit models with high-dimensional fixed effects. *Stata Journal*, 10, 628.

Hausmann, R. and Klinger, B. 2006. Structural transformation and patterns of comparative advantage in the product space. Cambridge, MA: John F. Kennedy School of Government, Harvard University.

Hausmann, R. and Rodrik, D. 2003. Economic development as self-discovery. *Journal of Development Economics*, 72, 603–33.

Hausmann, R., Hwang, J. and Rodrik, D. 2007. What you export matters. *Journal of Economic Growth*, 12, 1–25.

Hummels, D. and Klenow, P. J. 2005. The variety and quality of a nation's exports. *The American Economic Review*, 95, 704–23.

Imbs, J. and Wacziarg, R. 2003. Stages of diversification. *The American Economic Review,* 93**,** 63-86.

IMF. 2016. Fiscal Policy, How to adjust to a large fall in commodity prices, Fiscal Affaris Department, International Monetary Fund, Washington

Klinger, B. and Lederman, D. 2004. Discovery and development: an empirical exploration of new products. Washington, DC: World Bank.

Levchenko, A. A. 2007. Institutional quality and international trade. *The Review of Economic Studies*, 74, 791–819.

Limao, N. and Venables, A. J. 2001. Infrastructure, geographical disadvantage, transport costs, and trade. The World Bank Economic Review, 15, 451-479.

Melitz, M. J. 2003. The impact of trade on intra- industry reallocations and aggregate industry productivity. *Econometrica,* 71**,** 1695-1725.

Michler, J. D. and Josephson, A. L. 2016. To specialize or diversify: agricultural diversity and poverty dynamics in Ethiopia. *World Development*, http://dx.doi.org/10.1016/j.worlddev.2016.08.011

Milner, C. and McGowan, D. 2013. Trade costs and trade composition. *Economic Inquiry*, 51, 1886–902.

Mohan, P. 2016. Caribbean diversification and development. *The World Economy*, 39, 1434–53.

Parteka, A. and Tamberi, M. 2013. What determines export diversification in the development process? Empirical assessment. *The World Economy,* 36**,** 807-826.

Paudel, R. C. and Burke, P. J. 2015. Exchange rate policy and export performance in a landlocked developing country: the case of Nepal. *Journal of Asian Economics*, 38, 55–63.

Persson, M. and Wilhelmsson, F. 2016. EU trade preferences and export diversification. *The World Economy*, 35, 295–315.

Psacharopoulos, G. 1994. Returns to investment in education: a global update. *World Development*, 22, 1325–43.

Regolo, J. 2013. Export diversification: how much does the choice of the trading partner matter? *Journal of International Economics*, 91, 329–42.

Romalis, J. 2004. Factor proportions and the structure of commodity trade. *The American Economic Review*, 67–97.

Sekkat, K. 2015. Exchange rate misalignment and export diversification in developing countries. *The Quarterly Review of Economics and Finance*, 1–14.

Shepherd, B. 2010. Geographical diversification of developing country exports. *World Development*, 38, 1217–28.

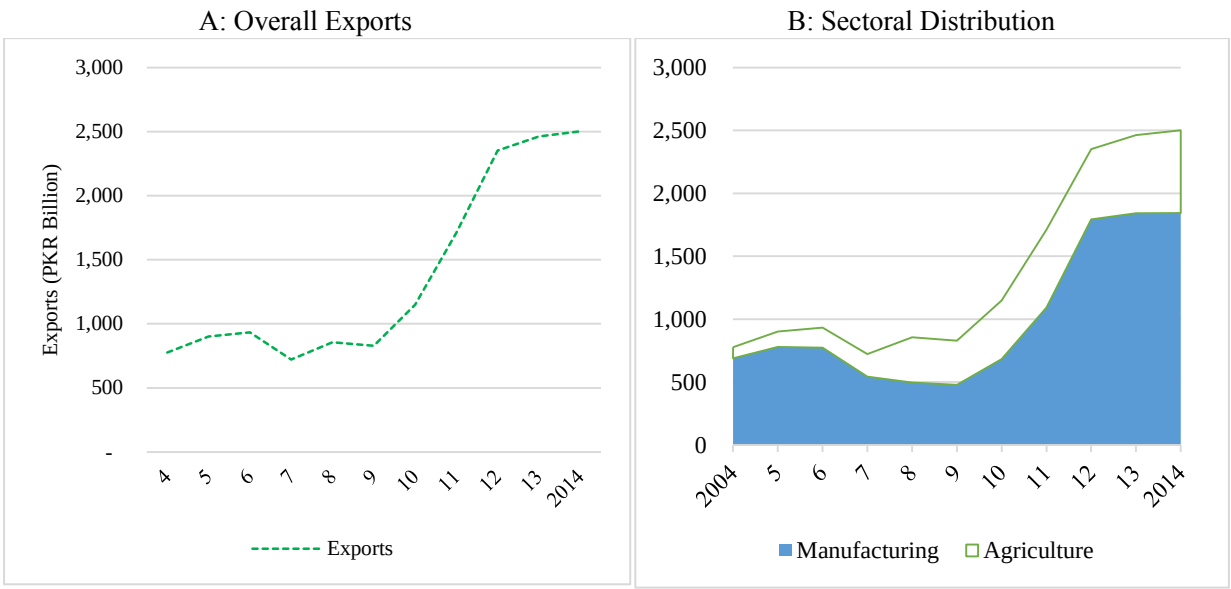
Silva, J. S. and Tenreyro, S. 2006. The log of gravity. *The Review of Economics and Statistics*, 88, 641–58.

UNCTAD. 2012. Handbook of statistics. Geneva: UNCTAD.

UNCTAD. 2016. Trade and Development Report, Structural transformation for inclusive and sustained growth, Geneva: UNCTAD

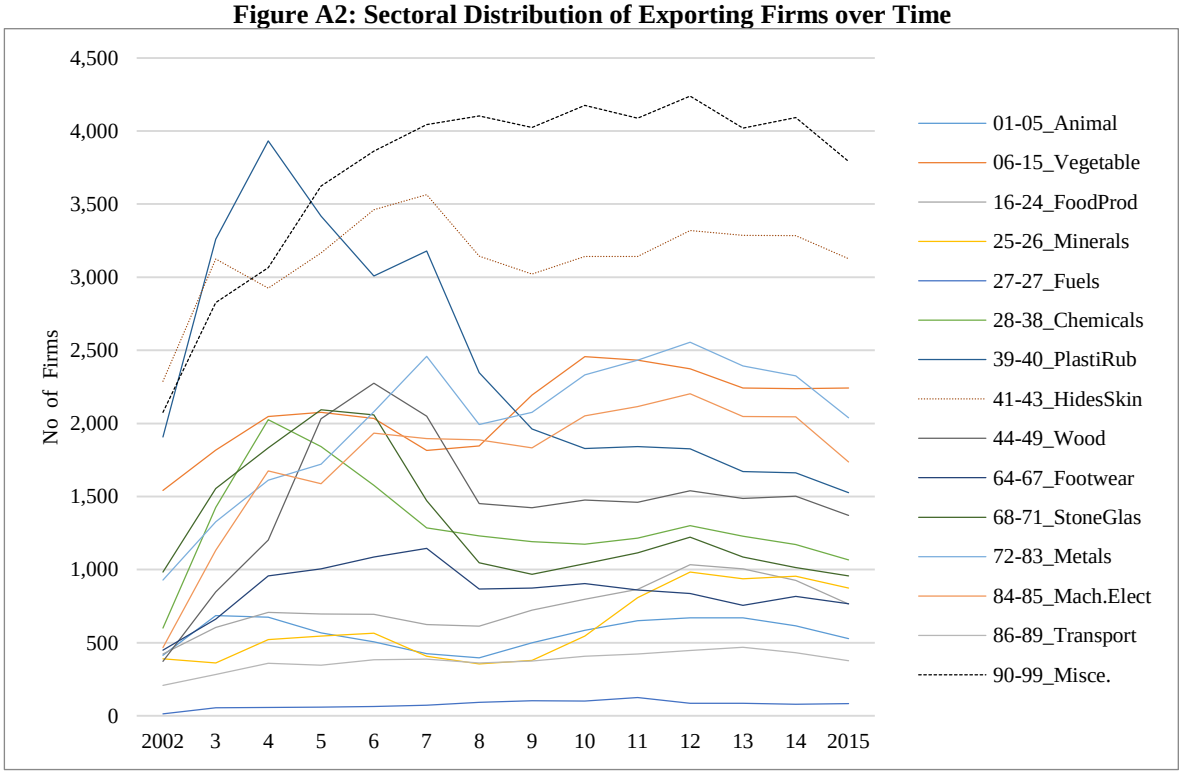
7 Appendix

Figure A1: Distribution of Exports over Time, 2004–2014



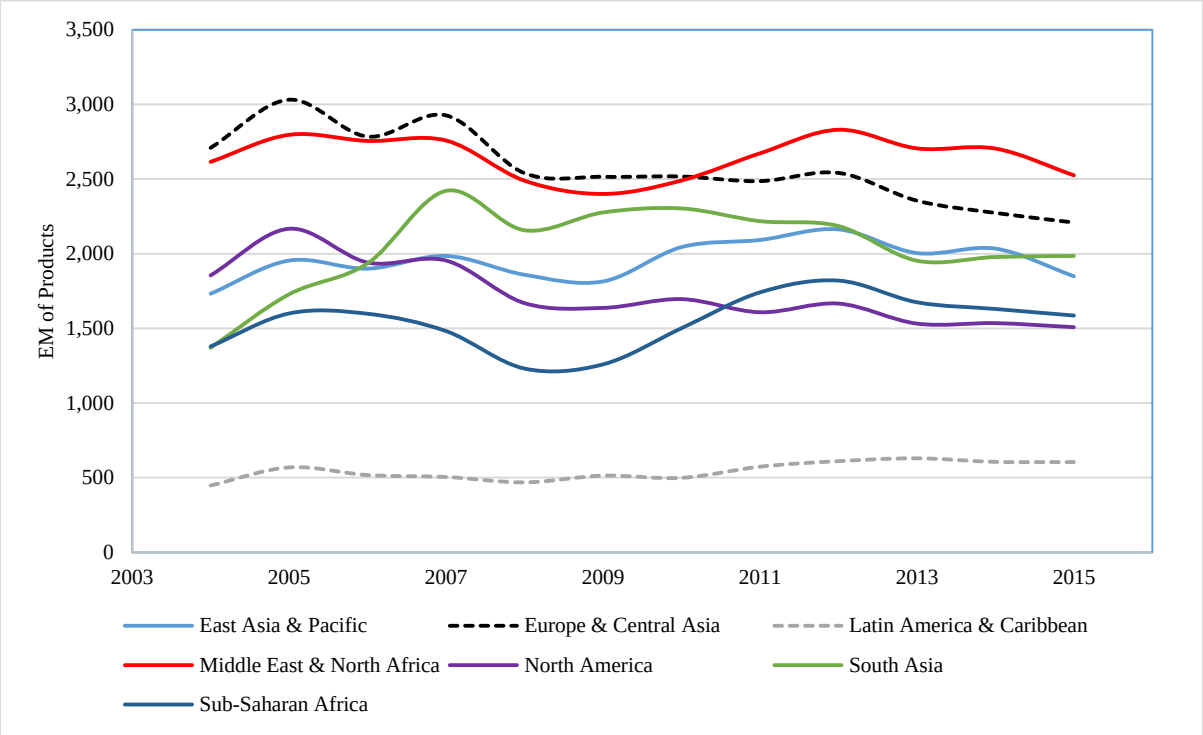
Note: Export values are in billion PKR.
Source: Constructed using administrative datasets

Figure A2: Sectoral Distribution of Exporting Firms over Time



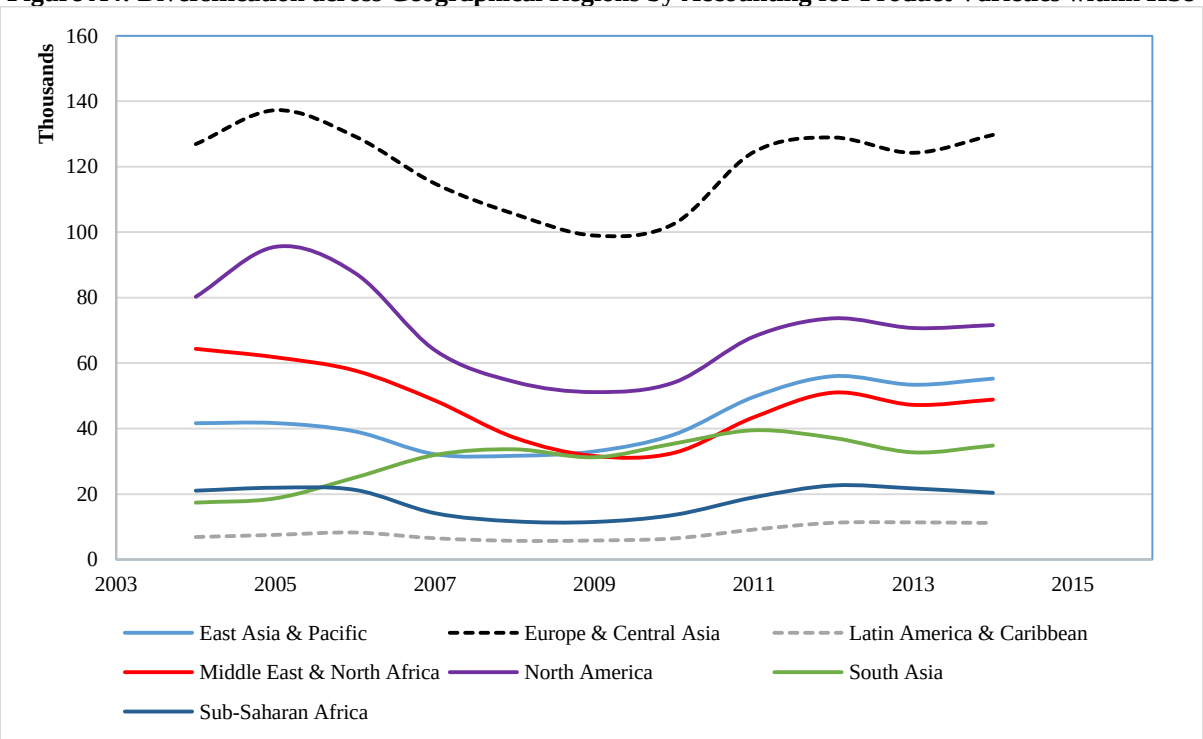
Source: Constructed using administrative datasets.

Figure A3: Diversification across Geographical Regions at HS8 Level, 2004–2015



Source: Pakistan Customs dataset.

Figure A4: Diversification across Geographical Regions by Accounting for Product Varieties within HS8



Source: Pakistan Customs dataset.

Table A1: Summary Statistics

Variables	Obs	Mean	Std. Dev.	Min	Max
DIFF H/L	1,020,257	0.5197	0.2003	0.0005	0.837
DIFF K/L)	1,020,257	2.2639	0.7765	0.0045	3.71
DIFF T/L	1,020,257	1.2301	1.29	0.0051	7.368
HHI	2,644	0.4657	0.2383	0.1012	1
C. Language (1, 0)	1,020,257	0.3655	0.4816	0	1
FTA (1, 0)	1,020,257	0.0258	0.1585	0	1
ln (Distance)	1,020,257	8.5756	0.6103	7.041	9.701
ln (GDP)	1,020,257	13.4573	1.9466	6.236	16.396

Note: HHI is Herfindhal-Hirschman Index of export concentration and $DIFF_{ij}^{K/L} = \left| \ln\left(K/L\right)_i - \ln\left(K/L\right)_j \right|; DIFF_{ij}^{H/L} = \left| \ln\left(H/L\right)_i - \ln\left(H/L\right)_j \right|; DIFF_{ij}^{T/L} = \left| \ln\left(T/L\right)_i - \ln\left(T/L\right)_j \right|$

Table A2: Heterogeneity of Effect of Factor Endowments on Export Diversification over Time

The dependent variable is log of export diversification indicator

	(1) Coefficients	(2) Standard Errors
Human Capital x		
2000	0.068***	(0.006)
2001	0.068***	(0.005)
2002	0.074***	(0.004)
2003	0.047***	(0.004)
2004	0.059***	(0.004)
2005	0.061***	(0.004)
2006	0.057***	(0.004)
2007	0.053***	(0.004)
2008	0.069***	(0.005)
2009	0.059***	(0.005)
2010	0.066***	(0.005)
2011	0.072***	(0.005)
2012	0.073***	(0.005)
2013	0.075***	(0.005)
Physical Capital x		
2000	0.043***	(0.007)
2001	0.043***	(0.005)
2002	0.066***	(0.004)
2003	0.083***	(0.004)
2004	0.069***	(0.004)
2005	0.068***	(0.004)
2006	0.064***	(0.004)
2007	0.084***	(0.005)
2008	0.055***	(0.005)
2009	0.051***	(0.005)
2010	0.036***	(0.005)
2011	0.028***	(0.005)
2012	0.028***	(0.004)
2013	0.031***	(0.005)
Land Endowments x		
2000	0.030***	(0.007)
2001	0.003	(0.005)
2002	-0.026***	(0.004)
2003	-0.025***	(0.004)
2004	-0.016***	(0.004)
2005	-0.022***	(0.004)
2006	-0.022***	(0.004)
2007	-0.027***	(0.004)
2008	-0.020***	(0.004)
2009	-0.016***	(0.004)
2010	-0.012***	(0.004)
2011	-0.022***	(0.003)
2012	-0.017***	(0.003)
2013	-0.017***	(0.003)
R2	0.205	
N	1,020,257	

The estimations show standardised beta coefficients with robust standard errors in column (2). These coefficients were obtained using Stata 13 SE. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The estimations include fixed effect for firms, products and time. The dependent variable, XDijkt, is total number of varieties exported by a firm (Nijkt) in the given year by sector relative to the average of the same for all firm across all markets ($\bar{N}$ijkt).

The International Growth Centre (IGC) aims to promote sustainable growth in developing countries by providing demand-led policy advice based on frontier research.

Find out more about
our work on our website
www.theigc.org

For media or communications
enquiries, please contact
mail@theigc.org

Subscribe to our newsletter
and topic updates
www.theigc.org/newsletter

Follow us on Twitter
@the_igc

Contact us
International Growth Centre,
London School of Economic
and Political Science,
Houghton Street,
London WC2A 2AE

IGC
International
Growth Centre

DIRECTED BY
LSE
UNIVERSITY OF OXFORD

FUNDED BY
UKaid
from the British people

Designed by soapbox.co.uk