University of Sussex
Business, Management & Economics

Economics Department Working Paper Series
No. 26-2011

Explaining the Diversification Path of Exporters in Brazil: How Similar and Sophisticated are New Products?[1]

Xavier Cirera
IDS and CARIS, University of Sussex

Anabel Marin
SPRU, University of Sussex and Conicet, Argentina

Ricardo Markwald
FUNCEX, Brazil

Abstract: A stylised fact of the economic literature suggests that export diversification is good for economic growth and is associated with economic development. In addition, there is evidence suggesting that the level of sophistication of countries' exports "matters" for growth and development. This paper contributes to this literature by analysing two unexplored dimensions of export diversification: the degree of relatedness (similarity) and sophistication of new products in relation to existing ones. The objective of this paper is to understand the mechanisms through which firms are able to diversify to less related and more sophisticated activities. We do so using a unique dataset that links data on exports, innovation and firms' characteristics at the firm level in Brazil. The main findings suggest that i) diversification occurs in very closely related activities, where firms have some core competences, ii) most diversification occurs in new products with lower level of sophistication than existing exports, iii) the degree of diversification and innovativeness of the production basket, and the position that the firm has developed in the domestic market appear to matter for diversification towards more or less distant products.

JEL Classification: **F14; L25**

Key Words: Diversification; Relatedness; Sophistication; Trade; Innovation; Brazil

[1] Financial support from the Economic and Social Research Council (ESRC) is gratefully acknowledged. We would like to thank Azim Essaji, seminar participants at the University of Sussex, FUNCEX and the Workshop "Firm, Trade and Development" at Aix-en-Provence for comments on earlier versions of the paper.

1 Introduction

One of the main objectives of economic development policy is to achieve export diversification. A widely accepted empirical result establishes that at least until relatively high levels of per capita income are reached, economic development is associated with the diversification of production into a progressively wider array of new types of industries and exported products (Imbs and Wacziarg, 2003). Diversification is crucial for achieving economic development for several reasons. It reduces vulnerability with respect to external shocks (Haddad *et al*. 2009), decreases the incidence of trade shocks (Ghosh and Ostry 1994), and creates learning opportunities. More importantly, it is clearly correlated with high rates of growth (Al-Marhubi 2000, Herzer and Nowak-Lehnmann 2006, Heiko Hesse 2009).

While the benefits of diversification are clear, it is less evident how to achieve it. There are, moreover, indications that breaking into new export markets is becoming increasingly difficult. One of the major realities of the growth paths of many developing and emerging economies is the difficulties in reaching a higher stage of diversification that would enable them to sustain growth and development. Their composition of production and exports still involves relatively high levels of concentration on 'natural' resource-based activities, with slow rates of diversification away from this concentration.

Recent evidence suggests that not only diversification is important for growth, but also that the diversification path matters. Hidalgo et al. (2007) show that the typical path followed by countries in the process of diversification occurs in products that are related in the "product space" to the pre-existent ones, and that the type of products that this process allows is crucial for growth (Hausmann et al., 2007).

The objective of this paper is to characterise the process of diversification regarding *relatedness* and *sophistication* and explore which firm level characteristics can be associated with the different paths identified. These two features of diversification paths are key for firms from developing countries, given the characteristics of their export baskets; typically heavily concentrated in a few commodities linked to natural resources and with low value added

Brazil constitutes an excellent case of study for this research for two reasons. First, because despite being one of the most diversified economies in Latin America, Brazil still lags well behind advanced economies and other emerging economies like China and Mexico regarding diversification (Hummels and Klenow 2005). Second, because even though its export basket is still heavily dependent on natural resources, Brazil has developed a few competitive manufacturing sectors, which makes the country an interesting case study.

In order to identify firm' level determinants of the path of export diversification we use five groups of firm level determinants identified in the economic, business, and innovation literature affecting the decision to diversify (Cirera et al., 2011). These are: (i) the structural characteristics of the firm, (ii) its position in the domestic market, (iii) characteristics of the firm's production basket, (iv) characteristics of the firm production processes, and (v) firm learning and innovative efforts.

Overall we find that: i) most diversification occurs in new products with lower level of sophistication than existing exports, ii) most diversification occurs in new products with lower level of sophistication than existing exports, iii) only the degree of diversification and innovativeness of the production basket, and the position that the firm has developed in the domestic market appear to matter for diversification towards more or less distant products, and; iv) none of these elements seem to matter for diversification towards more or less sophisticated products..

The paper is organised as follows. Section 2 describes the set of firm level determinants used to explain relatedness and diversification in export diversification. Section 3 describes the dataset and methodology used in the paper. Section 4 characterises the path of firm export diversification in Brazil. Section 5 characterises the path of firm diversification along two dimensions: relatedness and sophistication, and explores the association between firm level determinants to the process of diversification and diversification paths. The last section concludes.

2 Determinants of the Export Diversification Path: Relatedness and Sophistication

Most of the papers analysing firm diversification to new activities and how related or similar these new products are to existing ones come from the business literature. In particular, the "resource-based" literature stresses the importance of related capabilities and resources explaining which new products will be produced and exported in the context of multi product firms. While firms will target demand dynamic sectors for diversification, diversification will only occur to those sectors which are similar and can be handled by existing resources and capabilities (Lien and Klein 2010).

This idea has been largely explored by Hidalgo and Hausmann (2009), who proved this point at the country level. They built a network representing the product space based on co-occurrences of countries exporting the same product. The authors showed that diversification occurs in countries by moving to similar and closer products in the product space (Hidalgo *et al.* 2007) where they have capabilities.

Focusing at the firm or sector level, Neffke and Svensson Henning (2008) measure relatedness as co-occurrence of firm production portfolios at the plant level. The assumption is that the more plants produce the same pair of products the more similar the capabilities required to produce them. Bryce and Winter (2009) stress the role of experience, size, assets, sector complexity and R&D investments. Finally, Fan and Lang (2000) analyse firm performance according to different measures of relatedness and find that firms that have more vertically related activities are not necessarily better performers

The main implication of this literature is, therefore, that the diversification path is not random and follows "feasible" paths along proximity or relatedness to existing products produced or exported by the firm.

Relatedness raises two important questions for developing countries, often specialised in the lower end of the value added chain regarding exports. The first question is how to

diversify into related new exports of higher levels of technological sophistication in the mix exported. The second question is under what factors it is possible to diversify to less related activities, which are of higher value added.

While the business literature emphasise the role of capabilities and relatedness, there is very little understanding of the process through which firms diversify towards more or less sophisticated activities. As a result, it is uncertain when diversification will be more likely to occur into more or less sophisticated activities.

One area, however, where we have more information is on the determinants of the diversification process. Cirera et al. (2011b) review the business, innovation and economics literature, and find several candidates of firm level determinants that may impact the decision of a firm to diversify. These micro-determinants can be divided into five main groups:

1) **Structural characteristics of the firm** such as size, ownership global engagement, etc. Our main hypothesis here are that size and global engagement (reflected in higher FDI participation and involvement with foreign clients or enterprises of the group) are better prepared to engage in the process of export diversification. However, as we discussed before, these association do not always hold, since other variables may be more important. It might happen for instance that domestic firms that are more dependent on specific clients or buyers may find it more difficult to introduce new products for exporting.

2) **Position of the firm in the domestic market**. Firms in better positions in the domestic market, which have been using the domestic market to improve products, increase quality and gain better positions relative to their competitors, are more likely to introduce new products.

3) **Characteristics of the product basket of the firm**. Firms with production baskets which are more diversified, that introduce innovations in products and that are less concentrated in value are more likely to have the required capabilities to introduce new products to export, and for this diversification to be more capable to reach less related new products.

4) **Characteristics of the process of production**. Firms that are more efficient, and that have introduced improvements in their processes of production are more likely to be able to gain the capabilities that allow them to introduce new products for exporting.

5) **Learning efforts of the firm**. Firms that have invested in R&D and other efforts to change the characteristics of their products and processes, which have highly skilled personnel and have invested in effort to market their products, are also more likely to succeed in gaining the necessary capabilities to diversify exports, and are likely to be more capable to reach more unrelated and sophisticated products.

Cirera et al. (2011) find that these five sets of determinants play an important role explaining the decision of Brazilian firms to export. However, it is unclear which of these sets of determinants are more conducive to more/less related/sophisticated paths.

Understanding the effect of these firm level determinants on the path of diversification, therefore, can contribute to deepening our understanding in this area where very little is known. As a result, in addition to documenting how related and sophisticated is the diversification path for Brazilian firms, a key objective of the paper is also to analyse which of these firm level determinants is more or less conducive to unrelated and sophisticated diversification.

3 Data and Methodology

3.1 Data sources

In order to analyse firm export diversification we use a unique dataset that links production, trade and innovation data for Brazilian firms (Cirera et al., 2011). Concretely, we link the universe of exporters at the firm level and HS-8 digits classification, with the annual production and firm surveys for manufacturing firms, and the innovation survey available in 2000, 2003 and 2005 (See Appendix 1 for a detailed description).

3.2 Methodology

3.2.1 Measuring export diversification

Cirera et al. (2011) analyse extensively the process of export diversification in Brazil. One critical element when analysing diversification is to correctly identify episodes where new products are introduced for exporting. There are two challenges in doing this. First, we only observe exports flows for the period of our sample, so we cannot determine whether a product was introduced before this period. As a result, in our sample we cannot consider a new export a product that was exported in 2000, since we do not know whether it was exported in 1999. Second, most export flows tend to be short lived (Besedes and Prusa 2006, Martincus and Carballo, 2009), and, therefore any meaningful methodology to identify firm export diversification needs to consider some degree of time sustainability; otherwise we would identify an extraordinary number of cases of firm export diversification.

In order to address these issues, we follow Cirera et al. (2011) and identify a case of export diversification when a firm introduces a new product and:

- The new product is not exported in 2000, and once introduced is exported continuously until 2008; if introduced in 2007, also exported or in 2008 and 2009; or,
- The new product not exported in 2000, and once introduced exported at least 5 years; or exported three consecutive years at the end of the period (2006, 2007 and 2008; or 2007, 2008 and 2009); or,
- The new product not exported before 2002 and exported at least three years after.

3.2.2 Measuring and explaining relatedness

Measuring relatedness is a complex issue since products are more similar or dissimilar depending on the dimension that one wants to analyse. The key element that we want to capture is relatedness in terms of firm capabilities to produce products. Since firms' capabilities are very difficult to measure, the existing literature suggests different approximations, ranging from categorical measures to SIC classification distances, input ratios, commodity flows or co-occurrence measures (Lien and Klein 2009).

The crudest measure of relatedness looks at sector relatedness by focusing on industry or trade classification similarities; namely, whether pairs of products are within the same classification category in SIC or SITC classification at 3 or 4 digits measure. This type of measures, while simple to calculate, fail to capture the fact that certain products within the same sector at 3 or 4 digits of aggregation may require very different capabilities for their production.

Another set of measures is based on similarity in input use or commodity flows across. These measures provide a proxy of similarity in the production process across sectors. The idea is that products that require similar inputs have similar technologies and capabilities.

A final measure of relatedness is based on co-occurrence. Here, rather than assuming that similarity is based on belonging to the same sector or using the same technology, the assumption is that two products require similar capabilities when is likely that firms and countries tend to produce or export these products. This co-occurrence is then used as metric reflecting relatedness. Hidalgo and Hausmann (2009) build a network representing the product space based on co-occurrences of countries exporting the same product. They show that diversification occurs in countries by moving to similar and closer products in the product space (Hidalgo *et al.* 2007). Neffke and Svensson Henning (2008) measure co-occurrence at the plant level using firm production portfolios, under the assumption that the more plants produce the same pair of products the more similar the capabilities required to produce them.

In order to accommodate these different dimensions of relatedness in our analysis, we use the following three types of measures:

1 **Correlation based on the input use of the input-output matrix in 2005**. We calculate the correlation in terms of the Leontief input use between the 55 national account sectors, and then map the correlations from sectors to activities (CNAE 1.0) and then to the HS-8 product level of the MERCOSUR nomenclature (NCM). For each firm and year, we calculate the correlation between each product exported in t and the new product exported in t+1. Then, we take the maximum correlation as the measure of relatedness. If one of the products exported in t is in the same HS-4 digits sector than the new product introduced has a correlation value of one and, therefore, is highly related to existing exports.

2 **Correlation based on the product space (Hidalgo et al., 2007**). The authors develop a methodology where SITC-4 sectors are related in terms of co-occurrences defined by the conditional probability that any given pair of SITC-4

products is exported by countries in the world. [1] We then convert SITC-4 into HS-4 sectors using concordance tables and replace the correlation to unity when two products belong to the same HS-4. Then, we map the correlation between any pair of HS-4 sectors to any pair of HS-8 products. Again, we use the maximum correlation between all the products exported in t and the new product in t+1.

3 **Classification based measures.** We calculate the minimum difference between the existing and new product for different levels of aggregation and classifications. Concretely, we use HS-4 sector level and HS-2 sector level for trade classifications; and CNAE-2 and CNAE-3 for industrial classification. All product codes at HS-8 are mapped to each of this classifications and the different between new and existing product pairs is computed. A difference of zero implies that the new product is in the same classification sector than at least one of the existing products.

Since firms are normally multiproduct for both, domestic production and export, it is important to define the reference product for calculating relatedness. We first look at relatedness of new products in relation to existing exports in t. Since firms tend to export more than one product, we compute the different measures focusing on the distance between the new product in t+1 and the more similar product in the export basket in t. This measures how unrelated the new product is in relation to the closest product in the export basket. Then, we look at an alternative measure of relatedness in relation to the core production activity for the firm, measured by the most important product, exported or not, in terms of largest sales to the domestic market in t.

In some cases firms introduce more than one new product for exporting in the same year. In these cases we select the five new products with the largest sales, compute the distances to exports in t and core product in t, and select as a measure of relatedness the more dissimilar value.

Once we have calculated the different relatedness distances, corresponding to the different dimensions and measures, we proceed to analyse the impact of the different firm characteristics and efforts identified above on relatedness distances. We first create an index D_{it} with value zero if the firm *i* in period *t* is an exporter that do not diversify; with value one if it is an exporter who diversifies to a totally related activity, and; with value two if it is an exporter that diversifies to a totally unrelated activity. Table 1 shows for each measure when diversification is considered related or unrelated

Table 1 Definitions of relatedness in diversification

Measure	Dimension	Definition
Input use	Relatedness in terms of sector input use according the Leontief input-output matrix	Related if correlation=1 Unrelated if correlation≠1
Hidalgo et al	Relatedness according to capabilities required to export two products according map of product space at country level	Related if correlation=1 Unrelated if correlation≠1
HS-2 diff.	Relatedness according to the same HS-2 sector	Related if difference=0 Unrelated if difference≠0
HS-2 diff.	Relatedness according to the same HS-4 sector	Related if difference=0 Unrelated if difference≠0
CNAE2 diff.	Relatedness with core product according to the same CNAE2 sector	Related if difference=0 Unrelated if difference≠0
CNAE3 diff.	Relatedness with core product according to the same	Related if difference=0

[1] A country is considered to export a given product if it has a revealed comparative advantage larger than one.

	CNAE3 sector	Unrelated if difference≠0

Source: Author's own elaboration

Using D_{it} as dependent variable we estimate equation (2) using a multinomial logit estimator for all different sets of measures, for relatedness vis-a-vis exports and vis-a-vis core production activity. We use lagged dependent variables to avoid endogeneity problems in the decision to diversify. As explanatory variables X_{it-1}, we use the set of variables identified by the literature in section 2 (See Table 2 below). We also use year *T*, sector *S* and regional dummies *R* to control for year effects, sector demand factors and the large correlation between certain regions and export.

$$D_{it} = \alpha_0 + \sum_k \beta_k X_{kit-1} + \sum_{i=1}^{t} T_{t-1} + \sum_{j=1}^{j} S_j + \sum_{n=1}^{N} R_n + u_{it} \quad (1)$$

One problem with the formulation of equation (1) is the narrow definition of relatedness captured by D_{it}. In this index, we impose a restrictive definition of related diversification, only occurring when there is very high relatedness (correlation one or same sector). In order to better consider the degree of relatedness, we also add a new set of estimates that replace the index D_{it} with a variable Re_{it}, that uses the correlations and absolute value distances computed in section 4. Although these new dependent variables are continuous, they are truncated; correlations between -1 and 1, and the absolute value difference of product classifications truncated at zero. As a result, we use a random effects tobit estimator that allows us to handle the truncation of the dependent variable.

3.2.3 Measuring and explaining sophistication

A final element that this paper examines when looking at the path of diversification is the degree of sophistication. The question we want to formulate here is whether new exports in t+1 are or higher/lower sophistication than the most sophisticated product in t. This is a measure of whether firms use diversification to upgrade their export basket.

The definition of sophistication is clearly problematic, since it can be defined along several dimensions: quality, value added, technological content or conducive to higher country growth. We focus in two dimensions of sophistication suggested in the literature, sophistication conducive to growth and technology intensity. Concretely, we use the following two measures:

- PRODY – We use the measure of sophistication introduced by Hausmann *et al*. (2007) and Lall *et al.* (2006). Using the BACI dataset from CEPII that includes COMTRADE HS-6 trade data, we calculate for each product and year from 2001 to 2007, the measure defined in (2) below. This measure is a weighted average of the GDP per capita of the countries that export a specific product k, weighted by the respective export shares in relation to the sum of exports shares for that product and year.

$$PRODY_{kt} = \sum_i \frac{x_{ikt} / X_{kt}}{\sum_i x_{ikt} / X_{kt}} GDPcap_{it} \quad (2)$$

Once PRODY is calculated we re-scale the measure as the ratio with the mean PRODY on that specific year. Then we use the ratio to compare existing exports in t with the sophistication measure of the new exported product in t+1. We calculate the change in sophistication ratio from the most sophisticated product in the export basket to the new product. This sophistication change is zero when the new product introduced has the same sophistication level or it is within the same HS-6 code.

- OECD classification (Hatzichronoglou 1999) – we use the technological content sophistication index from the OECD. This classification groups products according to the following rankings: 1) not industrial products; 2) low technological intensity; 3) low/medium intensity; 4) medium/high intensity and 5) high technological intensity.

  Once we have grouped existing exports in these groups, we use the existing highest technological group and calculate the difference with the technological group of the new exported product. When the new product is within the same HS-6 digit group than an existing product in t, we set the difference to zero,

For the cases of diversification in several products we use the same approach than for relatedness indices. We use the five new products with the largest export shares, compute the differences in sophistication/technological content, and use the largest difference.

Once we identify differences in sophistication and technological content between new exports in t-1 and the closest export in t, we create a dichotomous variable T_{it}. This variable has value -1 if the new export implies a sophistication/technology level below the maximum existing in t-1; value zero if implies the same level and 1 if it implies a higher level of sophistication/technology index.

We estimate equation (3) using a multinomial logit estimator. We use lagged dependent variables to avoid endogeneity problems in the decision to diversify. As explanatory variables X_{it-1}, we use as before the available information on firm characteristics, productivity and size, market position, scope of production and exports, innovation and other variables of interest.

$$T_{it} = \alpha_0 + \sum_k \beta_k X_{kit-1} + \sum_{i=1}^{t} T_{t-1} + \sum_{j=1}^{j} S_j + \sum_{n=1}^{N} R_n + u_{it} \qquad (3)$$

We also use year *T*, sector *S* and regional dummies *R* to control for year effects, sector demand factors and the large correlation between certain regions and exports. Equation (4) is estimated only for the sample of exporters that diversify, so there is a significant reduction in the number of observations.

3.3 Explaining the Diversification path

Based on the set of firm determinants of diversification identified in Section 2 and the information available on the dataset (see Cirera et al. (2011) for more details), we use the following proxies for the empirical analysis:

Table 2 Proxies used in empirical analysis

Variables	Type	Description
	Related to:	
Size	**Structural characteristics of the firm**	Natural log of employment
Multinationality		Dummy with value 1 when firms have a share of foreign capital higher than 10%
Global group Integration		Dummy variable with value 1 if the firm is strongly linked to other firms in their group operating in foreign countries
Global value chain integration		Dummy variable with value 1 if the firm is strongly linked with clients operating in foreign countries
Quality of firm's products	**Position of the firm in the domestic market**	Ratio between the unit value of the firm's product and the average unit value for that product for all firms
Position in the domestic market		Firm market share in its main product
Concentration in the value of production	**Characteristics of the product basket of the firm**	Herfindahl index of production
Concentration in production		Distance CNAE (two digits) of the main products produced by the firm[2]
Innovative output - Product		Categorical variable that assumes the value of 1 if the firm has introduced a product innovation in the last 3 years
Innovative output- Process	**Characteristics of the process of production**	Categorical variable that assumes the value of 1 if the firm has introduced a product innovation in the last 3 years
Total Factor Productivity		TFP calculated using the methodology proposed by Levinsohn and Petrin
Geographical Distance	**Learning efforts of the firm**	The average geographical distance of all export destinations for the firm

R&D Innovation efforts		Dummy variable that assumes the value of 1 if the firm engages in R&D
Other innovation efforts different from R&D		Dummy variable that assumes the value of 1 when the firm engages investments in machinery, and in setting up plants.
Marketing efforts		Dummy variable that assumes the value of 1 if the firm engages in marketing expenditures
Firm's skills		Ratio between firm and sector average wage

[1] When the enterprise is multiproduct, the average unit value of the company is used

[2] This is a different dimension of concentration from the Herfindahl. While the Herfindahl shows the concentration of the firm revenue in terms of products, the difference in sector composition gives an idea of the production scope of the firm.

[3] This is confirmed in our data where we found that the value fob of exports is positively related with the mean distance of exports at firm level (see next section). We regress the normalized fob value on a set of product fixed effects, year dummies and the logarithm of the average distance to all the destination markets for a given export flow (product/firm) in a year. The coefficient on distance is 0.31 and statistically significant at 99% confidence level, suggesting that average distance increases the size of the flow. Larger flows are exported to more distant markets. However, we cannot differentiate whether this is due to higher prices or higher volumes, or both.

Source: Author's own elaboration

4 Stylised Facts on Firms' Diversification Path in Brazil

This section characterises the export diversification path in relation to relatedness and sophistication followed by Brazilian firms during the period 2000 to 2009.

4.1 Relatedness

Most diversification occurs in related activities, but unrelated diversification is not uncommon

In order to identify how related is the diversification path, we implement the methodologies described in section 3 to our dataset and determine how new products relate to existing products. We use two different set of products as reference. First, we compare new products exported in t with the closest product exported in t-1. Second, we use as a reference product the core activity for domestic production in t-1.

Figure 1 plots the probability distribution functions for the values of the different measures of relatedness. The first column refers to relatedness of new exports in relation to existing exports, while the second column computes the measures in relation

to the core production activity. In general we find that diversification tends to occur with higher likelihood in relatively related or similar products.

We define *related diversification* as the introduction of new products that are in the same product classification (zero distance) as the reference product or that have correlation one under the input use or Hidalgo‘s product space measures. Focusing on diversification vis-a-vis the export basket in t-1 (left column), we observe that related diversification occurs in 70 per cent and 49 per cent of the cases for input use and Hidalgo’s measure; while when using classification distances, diversification in the same HS-2 chapter occurs in 77 per cent of cases and within the same HS-4 group in 36 per cent of the cases. Clearly, the HS-4 measure and Hidalgo’s product space correlation appear to be more stringent measures of related diversification, than input use and HS-2 chapter.

When looking at the extent of unrelated diversification, different measures provide different pictures. Measures of relatedness based on HS classifications show long left tails. However, when looking in more detail to the HS-2 chapter based measure we obtained a small second mode between 40 and 60 HS-2 chapters’ difference. In most cases this appears to be the result of firms exporting products that can be part of different stages of a vertical value chain (i.e. inputs and final products), indicating that unrelated diversification may not be that unrelated if we account for different stages of the product cycle (see below).

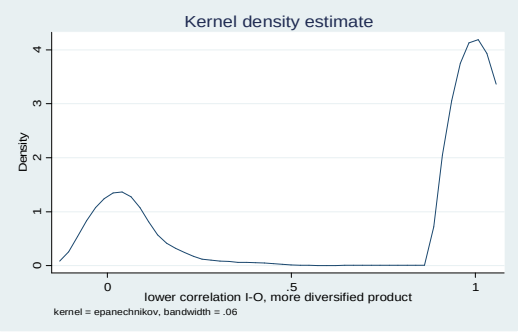
On the other hand, correlation based measures show a mixed picture. The measure based on correlation of input use suggests more polarised diversification, where unrelated diversification tends to occur in highly unrelated activities. On the other hand, Hidalgo’s product space correlation measure shows a more evenly distributed diversification across the correlation spectrum, with a large mode around 0.35 and a small second mode at around 0.65.

The second column in Figure 1 focuses in relatedness vis-a-vis the core production activity. By definition, related correlation is less frequent than before, since now we are comparing new products with the single activity with largest domestic sales. This gives us some measure of relatedness related to firm core competences. In this case, due to the fact that we compare products with domestic production, we use changes in CNAE (industrial classification) activity sectors and input use correlation.[2] The probability distribution functions plotted in the second column of Figure 1 show much lower prevalence of related diversification. Again, the sector input use correlation shows a large number of cases where diversification is somehow unrelated to core activities. The degree of unrelatedness or distance, however, is much lower when using changes in CNAE classification. These results suggest that in general in the context of multiproduct firms, export capabilities span well beyond any core activity.

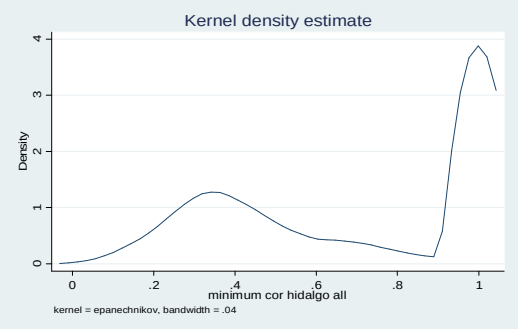
[2] The Hidalgo measure is based on correlations in the export product space using SITC classification at 4 digits.

Figure 1 Relatedness in Diversification

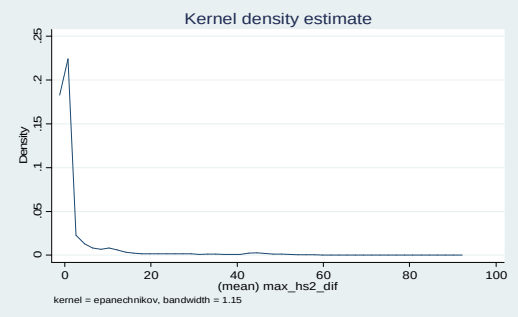
Relatedness vis-a-vis exports



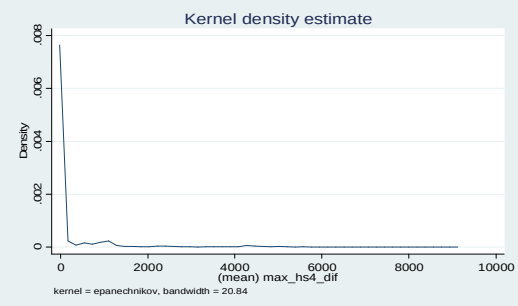
2(a) Input use similarity



2(b) Hidalgo et al. similarity

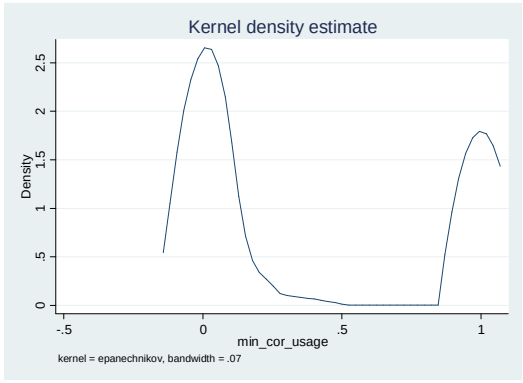


2 (c) HS-2 differences

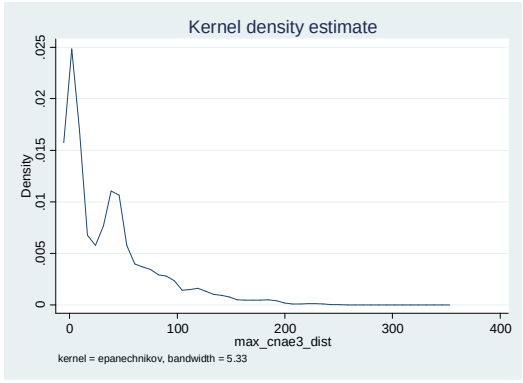


2(d) HS-4 difference

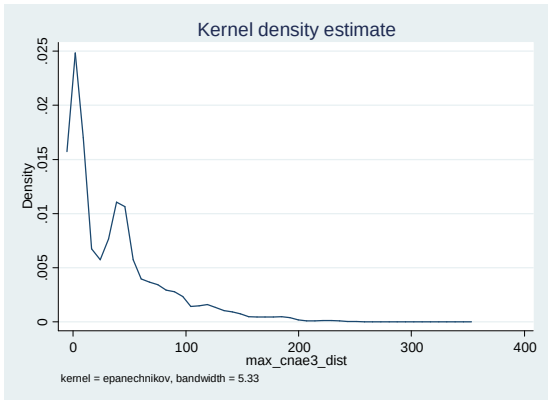
Relatedness vis-a-vis core product



2(e) Input use similarity



2(f) CNAE-2 difference



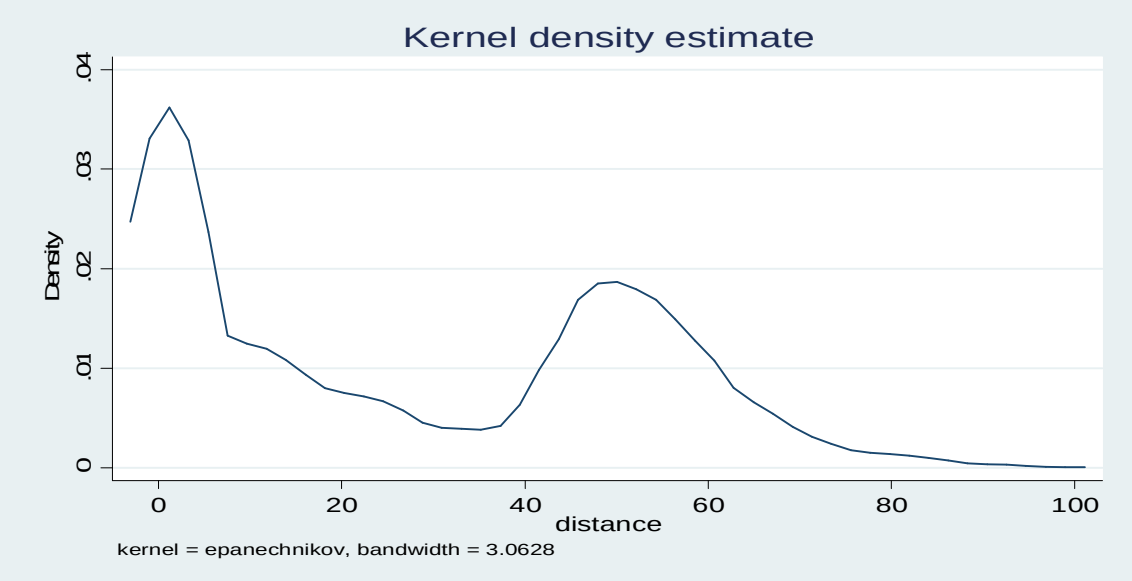
2(f) CNAE-3 difference

Unrelated diversification may still occur within a product cycle or a value chain

We further look at the issue of unrelated diversification path described above by taking advantage of the fact that in our sample some firms introduce more than one new product at the same time. This allows us to compare how related are new products. In this case we use a simple similarity index based on HS-2 chapter classification. We compute the maximum and minimum HS-2 code for all identified new products under alternative classifications for each firm. For only 8.4 per cent of the cases the difference between products is within the same HS-4 digits sector and 23 per cent within the same HS-2 chapter.

Figure 2 shows the probability distribution function of the calculated distances. Interestingly, there are two modes in the distribution. The first main mode occurs for observations with zero distance, where the new products introduce by the firm are in the same HS-2 chapter. There is, however, a second mode on the right of the distribution between a distance of 40 and 60 HS2 chapters. Interestingly, the large majority of firms that diversify more than one product with distances in this second cluster are firms that introduce a product from the plastic and rubber sectors as well as products from the machinery and transportation sectors. Therefore, while most firms that diversify in more than one product do so in products of the same sector, in a significant amount of cases where diversification occurs in products of different sectors, these products appear to be part of different stages of the same value chain.

Figure 2 Probability distribution function HS-2 distances – New products basket



The implication of this result is twofold. First, classification distances trying to measure relatedness and similarity fail to capture that some firms are able to produce and export in different stages of the value chain. As a result, while distances between products may appear to be large, capabilities within the same value chain may be similar. Second, multiproduct firms, where firms may produce in different sectors, face a wider array of diversification possibilities in different sectors, all requiring very similar capabilities.

Unrelated diversification occurs across different sectors

A final characterisation of the degree of relatedness in the diversification path is to look at the sector composition. We classify each firm that diversifies by sector according to the HS-2 chapter of the main export in t-1 and the main CNAE-2 production sector in t-1. Appendix 2 tabulates for each sector and relatedness method the number of firms that carry out related and unrelated diversification. Table A2.1 focuses on unrelated diversification between export activities. It includes agriculture and commodity HS-2 chapters corresponding to firms that export manufactures but their main export activity is non-manufacturing. Focusing on Hidalgo's relatedness measure, the sectors where non-related diversification is more prevalent are: 84 metal machinery, 85 electrical machinery, 44 wood and articles of wood, 90 optical and medical equipment, and 39 plastic and articles of plastic. These sectors account for most firms with unrelated diversification paths, and all have more than 39 per cent of diversification cases corresponding to unrelated activities.

Table A2.2 shows similar tabulations but in relation to relatedness to core activities. In this case, and focusing on input use correlations measures, there are two sectors that show large prevalence of unrelated diversification: sector 24 non-metallic mineral and 29 machines and equipments.

In general, it is difficult to interpret the sector decomposition of related and unrelated diversification, since multi-product firms have core competences that go beyond their core business activity for both export and production. Nevertheless, firms in machinery sectors seem to be more likely to diversify to less related activities. In addition, and contrary to some common misperceptions, firms with a core activity in some natural resources such as minerals are also able to diversify beyond their core activity sectors with high frequency.

These results tend to confirm the resource based approach to firm diversification, where due to capabilities constraints new exported products tend to be related to existing products. Nevertheless, even when we consider the fact that firms are multi-product and multi-export, unrelated diversification occurs with some frequency. Therefore, an interesting question is whether the processes to more or less related diversification require different firm dynamics. In other words, how do firms acquire the capabilities that allow them to jump to less related activities?

4.3 Sophistication/technological content

Low diversification to more sophisticated exports

While looking at relatedness in the diversification process is useful for analysing the scope of firms to diversify along the extensive margin, one important question is whether diversification occurs towards more sophisticated products or products with higher technological content. While firms will prioritise profitability of new activities for the given set of capabilities they have, it is important to analyse whether these diversification paths are conducive to products with larger value added or technological content. The extent to which diversification occurs in more

sophisticated activities gives an indication of the capacity of firms to use exports as a vehicle for upgrading sophistication.

In order to characterise the sophistication of the diversification path we use two measures (see section 3). First, we use the Hausmann *et al.* (2007) PRODY measure, which quantifies sophistications as the weighted average of the GDP per capita of the countries exporting the particular product. The assumption is that richer countries export more sophisticated products. In order to look at technology issues, the second measure is based on the OECD proposed index classifying sectors according to low, medium-low, medium-high and high technological intensity (Hatzichronoglou, 1999).

We compute the differences between the most sophisticated or higher technology product in the export basket in t-1 and the more sophisticated/higher technology content new product exported in t. The measures are also calculated in relation to the core production product. Then, when the differences between new and existing products are positive, we define the diversification path as diversification upgrading.

Table 3 below shows the results. In the case of the PRODY index, in 64 per cent of cases new exports are less sophisticated, 4 per cent of cases diversification occurs at the same level of sophistication (same HS-6) and in 32 per cent of cases there is diversification upgrading. For the OECD index, the fact that the measure is based on an index with 5 levels implies more prevalence of same level of technological content. In 60 per cent of cases diversification occurs in the same technological content level, 34 per cent in a lower technological content product and only 6% indicates diversification upgrading.

The measures are re-calculated in relation to the core production activity. As expected, the fact that we are comparing multiple products with only the core production yields much larger share of diversification upgrading. However, this share is very large, ranging from 85 to 91 per cent, which indicates that new exports tend to be of higher sophistication that the main core production activity.

Table 3 Sophistication/technological content changes in Diversification

change in	Exports		Core production	
sophistication/technology content	PRODY	OECD	PRODY	OECD
lower level	63.70%	33.47%	6.67%	2.70%
same level	4.28%	60.13%	1.88%	12.20%
higher level	32.03%	6.40%	91.45%	85.10%

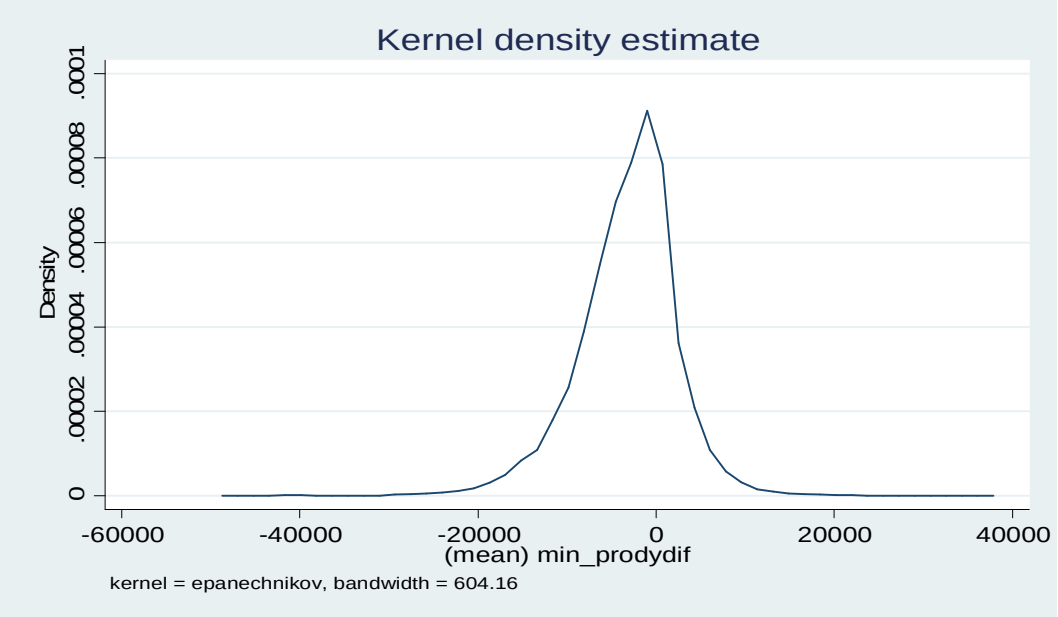
Source: Author's own elaboration

In order to look at the size of upgrading/downgrading we plot the probability distribution function of the PRODY index with regards differences between existing and new exported products (Figure 3). The figure shows the larger prevalence of new products been of lower sophistication. It also shows a larger tail on the left indication that positive changes in PRODY, when occur; tend to be lower in size (absolute value) than negative changes.

The table in Appendix 3 focuses on decomposing diversification upgrading by HS-2 core sector. The two main sectors with a very large share of diversification upgrading

and a significant number of product lines are sectors: 41 raw hides and skins and 44 wood and its articles.

Figure 3 Distribution of change in sophistication in product diversification (PRODY)



The findings regarding the degree of sophistication of the diversification path are sensitive to the index used. In general, we tend to find that most diversification occurs towards the same or lower sophistication/technological content products. However, there are a significant number of cases when using the PRODY index where diversification upgrading occurs, although this upgrading is small in size. Finally, when comparing with the core firm activity, new exports are largely of higher sophistication/technological content than the core production activity. This implies that lower sophistication core activities support financially the diversification and upgrading activities of the firm.

5 Explaining the Export Diversification Paths of Brazilian Firms

The degree of relatedness and sophistication of the diversification path are key for firms from developing countries, given the characteristics of their export baskets; typically heavily concentrated in a few commodities linked to natural resources and with low value added, The first matters to understand how capable are firms to expand to new activities, potentially in new sectors away from traditional exports. The second is important to determine how capable firms are to move to higher value added activities. Countries with firms more capable to move to unrelated and sophisticated new activities are likely to be able to move faster through the product space, and, therefore, likely to experience higher growth.

This section attempts to explain what types of firm determinants identified in Table 2 are more conducive to achieve diversification paths that are more unrelated and sophisticated.

5.1 Explaining Relatedness

5.1.1 Relatedness vis-a-vis exports

Based on the methodology described in section 3, we estimate equation (2) using a multinomial logit model. The results of the estimations are summarised in Table 4 The base category is exporters that do not diversify $D_{it}=0$. Therefore all coefficients, for exporters that diversify to highly related activities $D_{it}=1$ and exporters that diversify to unrelated activities $D_{it}=2$, need to be interpreted vis-a-vis exporters that do not diversify. We show the results for the Leontief input use measure, Hidalgo's correlation measure and HS-2 classification distance.

The Pseudo R^2 is around 0.25 and the chi-squared test of joint non-significance is rejected for all the specifications. We focus on the relative odds of each of the diversification options with regards to exporters that do not diversify.

We use two specifications. The first specification only uses as dependent variable the structural characteristics of the firm and its processes. Productivity and size increase the probability of a firm diversifying. The relative contributions of these two variables to both categories, related and unrelated diversification, depend on the measure used. The impact of an increase in productivity and size on the probabilities of diversifying is larger in unrelated diversification only when Hidalgo's measure is used. In the case of input use and HS-2 difference measures, productivity and size have a larger impact on the odds of related diversification.

Foreign ownership increases the probability of diversification and the odds are larger for related than unrelated diversification. This suggests a positive role of foreign control on diversification, but more importantly towards related diversification. Dependency on a parent company also increases diversification in general, and the size of the odds depends once again on the measure of relatedness used. Learning from linked firms can become an important vehicle to diversify to unrelated activities, but the results suggest that this is important for both types of diversification. Finally, the coefficient on dependency on clients is not statistically significant.

Regarding the position of the firm in the market, the quality level of the firm products proxied by the unit value ratio suggests a positive and statistically significant effect for unrelated diversification for Leontief and Hidalgo measures, and not statistically significant for related diversification. This implies that exporters with higher quality products are more likely to diversify to unrelated activities. However, when we measure relatedness wit HS-2 classification distances, we obtain that the quality coefficient is only positive for the probability of related diversification. In addition, firms' market power, proxied by its market share across products, increases the probability of diversification in general, either related or unrelated. Only in the case of Hidalgo's measure, the coefficient is negative but not statistically significant for related diversification.

Table 4 Multinomial logit estimates related diversification vis-a-vis exports

	Leontief		Leontief		Hidalgo		Hidalgo		HS2 difference		HS2 difference	
	related	unrelated	related	unrelated	related	unrelated	Related	unrelated	related	unrelated	related	unrelated
TFP	0.1758***	0.1463**	0.1639***	0.1471**	0.1180**	0.2340***	0.1194**	0.2152***	0.1698***	0.1476**	0.1553***	0.1539**
	(0.0467)	(0.0607)	(0.0475)	(0.0619)	(0.0591)	(0.0564)	(0.0603)	(0.0574)	(0.0460)	(0.0638)	(0.0468)	(0.0651)
log employment	0.3621***	0.3004***	0.3559***	0.3332***	0.2579***	0.4446***	0.2804***	0.4329***	0.3882***	0.1912***	0.3869***	0.2132***
	(0.0351)	(0.0454)	(0.0361)	(0.0470)	(0.0446)	(0.0410)	(0.0459)	(0.0422)	(0.0346)	(0.0490)	(0.0356)	(0.0503)
ratio unit value to product average	0.0324	0.0534*	0.0311	0.0525*	-0.0310	0.0565**	-0.0342	0.0548**	0.0437*	0.0214	0.0421*	0.0219
	(0.0263)	(0.0302)	(0.0266)	(0.0304)	(0.0385)	(0.0273)	(0.0392)	(0.0276)	(0.0248)	(0.0377)	(0.0251)	(0.0378)
Firm market share	0.4523**	0.5456**	0.5861***	0.6421**	-0.0833	0.5496**	0.0460	0.6897***	0.3360*	0.8987***	0.4722**	0.9738***
	(0.2000)	(0.2531)	(0.2022)	(0.2549)	(0.2617)	(0.2240)	(0.2643)	(0.2272)	(0.1970)	(0.2722)	(0.1992)	(0.2736)
distance CNAE 2 digits divisions	0.0109	0.0460***	-0.0010	0.0405***	0.0016	0.0272***	-0.0079	0.0155*	0.0102	0.0460***	-0.0012	0.0408***
	(0.0083)	(0.0099)	(0.0086)	(0.0103)	(0.0106)	(0.0091)	(0.0110)	(0.0094)	(0.0081)	(0.0104)	(0.0085)	(0.0109)
(mean) dist	0.0000	0.0000	0.0000	0.0000	0.0000*	0.0000**	0.0000*	0.0000**	0.0000**	0.0000	0.0000**	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
R& D dummy	0.2951***	0.2358**			0.2640**	0.2694***			0.3281***	0.1559		
	(0.0876)	(0.1134)			(0.1115)	(0.1000)			(0.0853)	(0.1257)		
Product innovation	0.2847***	0.3984***	0.3914***	0.4489***	0.2215**	0.4467***	0.2844***	0.5372***	0.2943***	0.3700***	0.4042***	0.4041***
	(0.0838)	(0.1102)	(0.0747)	(0.0980)	(0.1069)	(0.0979)	(0.0949)	(0.0875)	(0.0819)	(0.1200)	(0.0731)	(0.1061)
foreign	0.3681***	0.3640***	0.3860***	0.3768***	0.5117***	0.4342***	0.5324***	0.4469***	0.4082***	0.2115	0.4254***	0.2249
	(0.0954)	(0.1226)	(0.0962)	(0.1233)	(0.1179)	(0.1074)	(0.1191)	(0.1081)	(0.0927)	(0.1396)	(0.0935)	(0.1402)
Group dependency	0.2229*	0.2856*	0.2064*	0.2822*	0.0811	0.3581***	0.0688	0.3430***	0.2963**	0.2560	0.2794**	0.2558
	(0.1237)	(0.1526)	(0.1255)	(0.1542)	(0.1529)	(0.1302)	(0.1547)	(0.1318)	(0.1197)	(0.1812)	(0.1215)	(0.1824)
Client dependency	0.0531	-0.2092	0.0963	-0.2209	-0.0345	0.0465	-0.0317	0.0846	0.0731	-0.4120*	0.1105	-0.4101*

	(0.1510)	(0.2054)	(0.1521)	(0.2069)	(0.2002)	(0.1640)	(0.2019)	(0.1650)	(0.1466)	(0.2482)	(0.1479)	(0.2488)
Other Innovation			0.0000	0.0000			0.0000	0.0000			0.0000	0.0000
			(0.0000)	(0.0000)			(0.0000)	(0.0000)			(0.0000)	(0.0000)
marketing			0.1332**	0.2254***			0.1821**	0.1025			0.1680**	0.1699*
			(0.0677)	(0.0877)			(0.0857)	(0.0783)			(0.0662)	(0.0962)
skills			3.2854**	4.8176***			4.6001***	2.9386**			3.6881***	3.9831**
			(1.4240)	(1.4391)			(1.4213)	(1.4936)			(1.3308)	(1.6202)
university			-0.1039	-0.0488			-0.0303	-0.0778			-0.0919	-0.0657
			(0.0692)	(0.0906)			(0.0877)	(0.0806)			(0.0678)	(0.0986)
Herfindahl concentration normalised of production			-0.5659***	-0.2823**			-0.4535***	-0.5991***			-0.5525***	-0.2492*
			(0.0946)	(0.1255)			(0.1201)	(0.1119)			(0.0926)	(0.1365)
information			-0.1168	-0.1335			-0.2153**	-0.0414			-0.0991	-0.1589
			(0.0834)	(0.1081)			(0.1086)	(0.0932)			(0.0815)	(0.1201)
_cons	-26.3476	-39.5079	-25.5899	-39.2741	-37.5721	-28.7479	-36.9714	-28.0728	-39.5352	-26.1591	-38.8054	-25.8951
	.	.	.	.	.		.	.			.	.
N	9103		9103		9103		9103		9103		9103	
Log-likelihood	-4881.21		-4853.64		-4328.2		-4303.09		-4732.46		-4707.38	
Pseudo R2	0.243		0.2472		0.2422		0.2466		0.253		0.257	

Exporters that do not diversify are the base category *** significant at 1% confidence level, ** significant at 5% confidence level and * significant at 10% confidence level. Year, region and sector dummy coefficient omitted from the table.

.

Looking at the characteristics of the production mix of the firm, a more clear result is provided by the coefficient on the variable that measures the extent of sector diversification of the firm production basket. The coefficients on CNAE-2 distance are positive and statistically significant only for unrelated diversification, confirming Teece *et al.* (1994) findings of the importance of evolution and path dependent firm strategies that tend to diversify. Developing more cross-sector capabilities allow firms to diversify to unrelated exports with more likelihood. Also, firms that have more concentrated business structures and whose sales depend in fewer products, proxied by the normalised Herfindahl, are less likely to diversify. Finally, firms that carry out product innovations increase the probability of both types of diversification

In relation to learning efforts, the level of acquired skilled staff of the firm increases the probability of diversification, but for both related and unrelated diversification, with odds that vary according to the measure used. Marketing efforts also increase the probability of diversification, and the effect of the distance of destination markets is positive on diversification in general, although some coefficients are not statistically significant. In this case, it is unclear whether firms that diversify to more distant markets have also more capabilities to do so in unrelated activities. Regarding innovation efforts, firms that engage in R&D increase the probability of both types of diversification. Other innovation activities, however, do not have a statistically significant coefficient and, therefore, the results do not suggest that they increase the probability of diversification for exporters.

In general the results suggest that only the quality level of the firm, proxied by unit values and, specially, the degree of diversification of the production structure, have a significant larger impact explaining unrelated diversification. Higher quality investing firms may have more leverage and capabilities to expand to new activities that are less related. More importantly, existing diversified production capabilities facilitate jumps in the product space and also having firm strategies more conducive to introducing unrelated activities.

One interesting result when comparing relatedness methodologies is the fact that the coefficients that are statistically significant tend to be larger coefficients in related diversification for input use and classification based methodologies, and larger coefficients for unrelated diversification under Hidalgo's product space methodology. This suggests some degree of similarity between the technology dimension of relatedness measures using input use and classification based measures.

In order to better explain the differences in related and unrelated diversification paths, we re-estimate equation (2) using as dependent variable the correlation and classification distances, the Re_{it}. These are continues measures of relatedness. As a result, we can use only the sample of firms that introduce a new export in t and that have a defined distance measure.

The tobit RE estimates are shown in Table 5 below. We need to be careful when interpreting the signs. For correlation based measures, a positive coefficient implies higher correlation and, therefore, more relatedness. On the other hand, for classification based measures, a positive coefficient implies an increase in distance and, therefore, higher sector unrelatedness.

Table 5 Tobit RE estimates related diversification vis-a-vis exports

	Leontief correlation		hidalgo correlation		HS2 difference		HS4 difference	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
TFP using Levinsohn and Petrin and value added	0.0270	0.0210	-0.0190	-0.0190	-0.4210	-0.1830	-54.6800	-52.6470
	-(0.0403)	-(0.0412)	(0.0183)	(0.0181)	(0.8255)	(0.8318)	(41.4242)	(41.7833)
log employment	0.0270	0.0060	-0.0380**	-0.0300*	-2.0440**	-1.7390**	-22.3770	-11.2450
	-(0.0303)	-(0.0316)	-(0.0130)	-(0.0135)	-(0.6428)	-(0.6637)	(30.6534)	(32.1286)
ratio unit value to product average	-0.0100	-0.0100	-0.0140	-0.0140	-0.4650	-0.4270	36.2450	37.4550
	(0.0227)	(0.0217)	(0.0095)	(0.0095)	(0.5407)	(0.5338)	-(23.5357)	-(23.4094)
firm market share by product	-0.0910	-0.0730	-0.1780*	-0.1760*	4.2630	3.7400	-8.7620	-10.7690
	(0.1717)	(0.1738)	-(0.0739)	-(0.0746)	-(3.6436)	-(3.6311)	(175.2400)	(179.4833)
Herfindahl concentration normalised of production		-0.1620		0.0070		3.1580		33.3140
		(0.0848)		-(0.0350)		-(1.7842)		-(85.4205)
distance CNAE 2 digits divisions	-0.0220**	-0.0250**	-0.0070*	-0.0070*	0.4990**	0.5670**	19.2510**	20.2790**
	-(0.0067)	-(0.0069)	-(0.0028)	-(0.0031)	(0.1378)	(0.1428)	(6.8754)	(7.0413)
(mean) dist	0.0000	0.0000	0.0000	0.0000	0.0010*	0.0010*	0.0220	0.0230
	(0.0000)	(0.0000)	(0.0000	(0.0000)	(0.0004)	(0.0004)	-(0.0127)	-(0.0126)
dummy for product innovation	-0.0620	-0.0100	-0.0730*	-0.0710*	0.8550	-0.5160	116.5310	28.3100
	(0.0775)	(0.0714)	-(0.0332)	-(0.0290)	-(1.6442)	(1.3946)	-(79.8158)	-(69.0488)
Other_Innov_Exp		0.0000		0.0000		0.0000		-0.0010
		(0.0000)		(0.0000)		(0.0000)		(0.0011)
RDdummy	0.0600		0.0170		-2.7050		-106.4320	
	-(0.0789)		-(0.0333)		(1.6801)		(81.2458)	
significant marketing changes		-0.0690		0.0030		0.4320		100.4180
		(0.0585)		-(0.0214)		-(1.2343)		-(59.7726)
number of high skill technical staff		-1.0690		0.5700		-3.8510		1576.3100
		(0.9137)		-(0.4524)		(20.2684)		-(1037.0461)

group_dep1	-0.0570	-0.0590	-0.0730	-0.0760*	-1.8470	-1.5220	37.6660	38.4800
	(0.0966)	(0.0952)	(0.0376)	-(0.0380)	(2.1729)	(2.1743)	-(96.5795)	-(98.6667)
client_dep1	0.1540	0.1790	-0.0130	-0.0160	-6.4480*	-6.9000*	-251.5530*	-277.5840*
	-(0.1283)	-(0.1288)	(0.0481)	(0.0500)	-(2.9443)	-(2.9362)	-(128.3434)	-(128.5111)
high information from university		-0.0360		0.0100		0.3940		19.8580
		(0.0600)		-(0.0250)		-(1.2710)		-(62.0563)
foreign_cap1	0.0290	0.0260	0.0270	0.0280	-1.7030	-1.6770	-62.0620	-53.7840
	-(0.0829)	-(0.0839)	-(0.0342)	-(0.0341)	(1.8117)	(1.8228)	(85.0164)	(85.3714)
independent or group		0.0290		-0.0270		-2.2380		-83.8260
		-(0.0707)		(0.0290)		(1.5329)		(71.6462)
Constant	6.4200	6.6690	0.7880	0.8010	2.6040	-0.8630	469.5590	327.4780
	-(321.0000)	-(222.3000)	-(0.5019)	-(0.5070)	-(23.6727)	(21.5750)	-(1381.0559)	-(1364.4917)
Log-likelihood	-1883.01	-1878.85	-1301.74	-1300.53	-3142.10	-3139.01	-14187.65	-14184.82
rho	0.1165	0.1005	0.1153	0.1145	0.2154	0.2045	0.0293	0.0170
	0.0670	0.0678	0.0685	0.0687	0.0818	0.0824	0.0676	0.0449
Observations	2214	2214	1788	1788	2240	2240	2240	2240
Number of group (firm_)	1719	1719	1407	1407	1741	1741	1741	1741

*** significant at 1% confidence level, ** significant at 5% confidence level and * significant at 10% confidence level. Year, region and sector dummy coefficient omitted from the table.

The parameter rho is the panel-level variance component. It is statistically significant in most cases, but for the case of HS-4 differences rho is not statistically significant and the panel estimator does not explain a larger part of the variance than the pooled estimator.

We focus on statistically significant coefficients. The main variable that appears to explain unrelated diversification is as suggested above the degree of diversification of the production structure; reflecting, the importance of existing capabilities and business strategies opting for unrelated diversification. This result is consistent across specifications.

In the case of the Hidalgo correlation measure, larger firms, firms with more market power and firms that introduce product innovations tend to diversify towards less related export products. While firm size may be important to acquire capabilities for new exports, the result on market power is less obvious. On the one hand, firms with more market power may have more leverage for introducing new unrelated exported products. On the other hand, business concentration on fewer products in terms of business value may be reflecting a narrow a business strategy that focuses on expanding a few set of existing products. The results suggest the former been more important than the latter. Finally, product improvements in t-1 are likely to result on more likelihood of introducing new unrelated products.

Regarding, unrelated diversification measured by changes in HS classification, an interesting result is the fact that the higher the degree of dependence from clients and buyers, the more related is the export diversification path. Clients or buyers appear to encourage diversification but in related activities within the same sector. This could be the result of the interest of these buyers in being supplied in similar goods, or the lack of incentives for firms to develop different export products and finding buyers outside existing clients.

Again comparing the different classification measures, we observe more similarities between input use and classification based measures in the sign of the coefficients, which may indicate that sector use of inputs is similar within classification categories.

5.1.2 Relatedness vis-a-vis core production activity

Relatedness can also be expressed in relation to the core production activity of the firm in terms of sales. The core production activity is a good approximation to the core competences of the firm. In addition, it is possible that for some firms this core product identified using the manufacturing survey (PIA) is not exported. For these cases, explaining the distance between the core product and the new exported product may shed some light about how firms build the specific capabilities required for exporting, such as trade networks, information or marketing.

As suggested by the left hand side panel in Figure 1, unrelated diversification is more prevalent when considering relatedness towards core activities. In general, we observe that in the context of multiproduct firms, export capabilities go beyond any core activity.

Table 6 Multinomial Logit estimates related diversification vis-a-vis core production activity

	Leontief input use correlation				CNAE 2 difference				CNAE 3 difference			
	related	**unrelated**	**Related**	**unrelated**	**related**	**Unrelated**	**related**	**Unrelated**	**related**	**Unrelated**	**related**	**unrelated**
TFP	0.1260**	0.2529***	0.1190*	0.2385***	0.1049	0.2412***	0.1022	0.2263***	0.1305**	0.2583***	0.1190**	0.2484***
	(0.0606)	(0.0549)	(0.0617)	(0.0558)	(0.0675)	(0.0514)	(0.0687)	(0.0523)	(0.0570)	(0.0574)	(0.0580)	(0.0583)
log employment	0.2978***	0.3909***	0.3101***	0.3916***	0.2645***	0.3937***	0.2960***	0.3853***	0.2797***	0.4140***	0.2874***	0.4173***
	(0.0467)	(0.0397)	(0.0478)	(0.0409)	(0.0531)	(0.0375)	(0.0545)	(0.0386)	(0.0440)	(0.0412)	(0.0452)	(0.0423)
ratio unit value to product average	-0.0021	0.0430	-0.0046	0.0423	0.0181	0.0330	0.0187	0.0307	-0.0088	0.0474*	-0.0143	0.0472*
	(0.0403)	(0.0268)	(0.0406)	(0.0271)	(0.0433)	(0.0264)	(0.0434)	(0.0268)	(0.0393)	(0.0270)	(0.0401)	(0.0272)
Firm market share	-0.1478	0.8275***	0.0001	0.9520***	-0.2982	0.7737***	-0.1572	0.9211***	-0.1991	0.9734***	-0.0429	1.0759***
	(0.2805)	(0.2134)	(0.2825)	(0.2160)	(0.3208)	(0.2042)	(0.3202)	(0.2072)	(0.2577)	(0.2226)	(0.2611)	(0.2248)
Herfindahl concentration normalised of production			-0.4396***	-0.4948***			-0.1420	-0.6256***			-0.5536***	-0.3932***
			(0.1242)	(0.1088)			(0.1409)	(0.1024)			(0.1188)	(0.1121)
distance CNAE 2 digits divisions	-0.0531***	0.0396***	-0.0625***	0.0299***	-0.0522***	0.0297***	-0.0532***	0.0171**	-0.0465***	0.0420***	-0.0581***	0.0342***
	(0.0135)	(0.0087)	(0.0140)	(0.0090)	(0.0153)	(0.0084)	(0.0157)	(0.0087)	(0.0124)	(0.0089)	(0.0129)	(0.0092)
(mean) dist	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000**	0.0000	0.0000*	0.0000	0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Product innovation	0.1030	0.4851***	0.2640***	0.5786***	0.0777	0.4447***	0.2043*	0.5560***	0.1414	0.4763***	0.3033***	0.5622***
	(0.1130)	(0.0952)	(0.0989)	(0.0849)	(0.1288)	(0.0894)	(0.1123)	(0.0798)	(0.1074)	(0.0980)	(0.0947)	(0.0872)
Other Innovation			0.0000	0.0000			0.0000*	0.0000			0.0000	0.0000*
			(0.0000)	(0.0000)			(0.0000)	(0.0000)			(0.0000)	(0.0000)
R& D dummy	0.4031***	0.2584***			0.3920***	0.2781***			0.4019***	0.2394**		
	(0.1191)	(0.0969)			(0.1363)	(0.0915)			(0.1120)	(0.1004)		

marketing			0.2320***	0.0223			0.2222**	0.0568			0.1850**	0.0448
			(0.0888)	(0.0766)			(0.1008)	(0.0722)			(0.0847)	(0.0788)
skills			3.5486**	2.9687**			4.2146**	2.7711**			4.0543***	2.6878*
			(1.6938)	(1.3412)			(1.7876)	(1.3184)			(1.5560)	(1.4072)
Group dependency	-0.2345	0.2059*	-0.2478	0.1895	-0.3242	0.1947	-0.3386	0.1818	0.0589	0.1474	0.0469	0.1284
	(0.1969)	(0.1242)	(0.1982)	(0.1259)	(0.2311)	(0.1212)	(0.2326)	(0.1230)	(0.1656)	(0.1304)	(0.1671)	(0.1320)
Client dependency	0.0759	-0.1930	0.0968	-0.1542	-0.0598	-0.1022	-0.0553	-0.0614	-0.0099	-0.1598	0.0162	-0.1253
	(0.2054)	(0.1712)	(0.2067)	(0.1718)	(0.2444)	(0.1615)	(0.2459)	(0.1627)	(0.1999)	(0.1747)	(0.2017)	(0.1750)
university			-0.1019	-0.0745			-0.0087	-0.1161			-0.1206	-0.0618
			(0.0918)	(0.0786)			(0.1037)	(0.0742)			(0.0876)	(0.0808)
foreign	0.0029	0.5525***	0.0198	0.5616***	0.0420	0.4691***	0.0554	0.4809***	0.1015	0.5204***	0.1165	0.5306***
	(0.1374)	(0.1015)	(0.1382)	(0.1021)	(0.1559)	(0.0976)	(0.1568)	(0.0983)	(0.1266)	(0.1053)	(0.1275)	(0.1058)
information			-0.1343	-0.1412			-0.2417*	-0.1028			-0.1150	-0.1627*
			(0.1129)	(0.0921)			(0.1313)	(0.0870)			(0.1059)	(0.0952)
N	9103		9103		9103		9103		9103		9103	
Log-likelihood	-4209.65		-4191.52		-4141.63		-4118.10		-4245.97		-4227.55	
Pseudo R2	0.2830		0.2861		0.2764		0.2805		0.2822		0.2853	

Exporters that do not diversify are the base category *** significant at 1% confidence level, ** significant at 5% confidence level and * significant at 10% confidence level. Year, region and sector dummy coefficient omitted from the table.

Table 6 shows the results of the multinomial logit estimations for relatedness vis-a-vis core production activities. Again the results need to be interpreted as the odds in relation to the base category: exporters that do not diversify. Since the Hidalgo product space correlation is defined mainly for exports, we only use the Leontief input use measure, and industry based classifications, CNAE 2 digits and CNAE 3 digits.

The pseudo R^2 suggests an overall fit of around 0.28. As in the case of relatedness with exports, both size and productivity increases the odds of exporters to diversify. In this case, however, the coefficient suggests that productivity and size increases more the probability of diversifying towards unrelated activities than to related activities. The degree of quality of the products produced, proxied by the unit value, is not statistically significant in most cases, only increasing the odds of unrelated diversification within CNAE-3 sectors. Firm value concentration represented by the Herfindahl index is again negative, indicating that firms with a more concentrated business in few activities are less likely to diversify and remain exporting the same products. The size of the odds depends on the relatedness measure used.

The extent of sector diversification in production is again the more robust determinant of the diversification path. The more diversified production is the less likely for exporters to diversify to related activities and the more likely is to diversify to unrelated sectors. This implies that when we look at core capabilities, more diversified exporters are more likely to diversify to unrelated activities and less likely to diversify to very similar activities. Product innovation, however, is more important for unrelated diversification, and not always statistically significant for related diversification.

In terms of learning efforts, the distance to main destination export markets is in this case mainly statistically not significant. Skills and marketing increase the probability of an exporter to diversify to related exports. In the case of skills it also increases the odds of unrelated diversification for an exporter, but with lower probability than related diversification. Regarding innovation variables, engaging in R&D increases the odds of an exporter diversifying, and there is larger probability for related diversification. Finally, marketing innovation increases the probability only of related diversification, while foreign owned firms are more likely to engage in unrelated diversification.

In general, when we compare the results of the determinants of related diversification paths between relatedness to exports and to core firm activity a few important issues emerge. The quality of products and distance to export destinations are not important for diversification in relation to the core activity. In particular, the quality of products proxied by their unit value appears to matter only for unrelated diversification in relation to exports; which may signal that only higher quality firms can expand to unrelated activities in international markets. Foreign ownership is more important when explaining unrelated diversification in relation to core activities, while in the previous case of relatedness in relation to exports, it was important for both types of diversification. The impact of skills and marketing efforts increase the probability of diversification, but for core production relatedness the odds for related diversification are higher. Finally, the effects of dependency on a group or clients are not statistically significant explaining relatedness in diversification to the core activity.

Table 7 Tobit RE estimates related diversification vis-a-vis core production activity

	Leontief correlation		cnae3 difference		cnae2 difference	
	(1)	(2)	(3)	(4)	(5)	(6)
TFP using Levinsohn and Petrin and value added	-0.0300	-0.0290	4.1700*	3.8920*	0.4870*	0.4530*
	(0.0203)	(0.0207)	(1.7090)	(1.7145)	(0.2108)	(0.2117)
log employment	-0.0300	-0.0240	3.2810*	2.4070	0.3950*	0.3240
	(0.0160)	(0.0166)	(1.3124)	-(1.3599)	(0.1626)	-(0.1679)
ratio unit value to product average	-0.0050	-0.0050	0.2000	0.2560	0.0600	0.0690
	(0.0122)	(0.0114)	-(1.0000)	-(0.9846)	-(0.1154)	-(0.1169)
firm market share by product	-0.2670**	-0.2690**	16.1890*	15.2530*	2.0010*	1.8580*
	-(0.0905)	-(0.0909)	(7.4261)	(7.4044)	(0.9179)	(0.9153)
Herfindahl concentration normalised of production firm year		0.0260		4.3070		0.9780*
		-(0.0456)		-(3.6812)		(0.4549)
distance CNAE 2 digits divisions	-0.0220**	-0.0210**	1.7780**	1.8290**	0.2200**	0.2360**
	-(0.0037)	-(0.0038)	(0.2993)	(0.3079)	(0.0364)	(0.0375)
mean) dist	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
dummy for product innovation	-0.1000*	-0.0710*	7.9560*	5.0230	0.9790*	0.5660
	-(0.0391)	-(0.0341)	(3.2741)	-(2.8219)	(0.4029)	-(0.3472)
RDdummy	0.0500		-5.8820		-0.7800	
	-(0.0397)		(3.3232)		(0.4084)	
Other_Innov_Exp		0.0000		0.0000		0.0000
		(0.0000)		(0.0000)		(0.0000)
significant marketing changes		0.0320		-1.4550		-0.1940
		-(0.0291)		(2.4661)		(0.3031)
number of high skill technical staff		-0.0670		-20.3550		-1.9430

		(0.5154)		(42.4063)		(5.1132)
group_dep1	-0.0510	-0.0490	1.3210	1.7000	0.0210	0.0750
	(0.0455)	(0.0458)	-(3.8853)	-(3.8636)	-(0.4200)	-(0.4688)
client_dep1	0.0780	0.0780	-5.1780	-5.2800	-0.6870	-0.7290
	-(0.0624)	-(0.0624)	(5.2303)	(5.2277)	(0.6361)	(0.6395)
high information from university		-0.0210		0.0730		0.1600
		(0.0309)		-(2.4333)		-(0.3137)
foreign_cap1	-0.1310**	-0.1320**	5.4090	5.3170	0.6500	0.6340
	(-0.0428)	-(0.0429)	-(3.5123)	-(3.5212)	-(0.4305)	-(0.4284)
independent or group		-0.0160		1.9250		0.1060
		(0.0348)		-(2.9615)		-(0.3655)
Constant	4.6880	4.9760	-71.6870	-72.0800	-9.6140	-9.8390
	-52.0889	-165.8667	52.7110	53.0000	6.2429	6.3071
Observations	1884	1884	1884	1884	1884	1884
rho	0.4345	0.4344	0.2614	0.2570	0.3002	0.2929
	0.0556	0.0560	0.0549	0.0550	0.0577	0.0579
Log-likelihood	-1604.41	-1604.11	-7761.08	-7760.66	-4036.99	-4035.02

*** significant at 1% confidence level, ** significant at 5% confidence level and * significant at 10% confidence level. Year, region and sector dummy coefficient omitted from the table

The main common finding between both types of relatedness, in relation to exports and to core activity, is the role of the degree of sector diversification of production increasing the likelihood of unrelated diversification.

Again we re-estimate equation (2) using the computed distance measures and the random effects tobit model (see Table 7 above). Positive signs increase relatedness diversification of the Leontief input use measure and decreases within industry classification relatedness. The rho coefficient suggests panel-level variance of around 0.25-0.43, indicating that panel estimates capture a large proportion of the variance in comparison to pooled estimates.

Size and productivity increase unrelated diversification; however, the coefficients are not statistically significant for the input use measure. Market power and the sector diversification of the product array also increase unrelated diversification. Finally, product innovation and foreign ownership also increases the degree of unrelated diversification, however, the coefficients for industry classification based measures are not always statistically significant.

The results are very similar to the ones in Table 5. The main differences lie in the statistical significance of some of the coefficients. In general, we find that larger and more productive firms, with larger market power and foreign owned, that have more diversified capabilities, are more likely to diversify to unrelated activities. In addition, product innovation is more likely to lead to unrelated diversification.

5.1.3 Robustness

Given the low survival rates of exports we imposed in our methodology that firm where required to be able to sustain exports for some years in order to be selected as diversification cases. This is potentially problematic if our control group, firms that do not diversify, includes a large number of firms that close down. If this is the case, factors explaining firm closure and survival may be affecting the decision to diversify.

In order to control for this potential problem, we create a new sample, the survivor sample, which contains only firms that survive our whole sample period. This sample, therefore, does not contain closing down firms, and minimises the risk of "closure" factors affecting the decision to diversify to more or less related activities.

We re-estimate equation (2) again for all relatedness measures and in relation to exports and the core activity using the survivor sample. The estimations produce very similar results to the ones described above. [3] This is mainly explained by the fact that the final survivor sample contains a large proportion of the normal sample, which implies that it is very unlikely that survivor factors may be important factors affecting the estimations.

[3] The results are available upon request.

5.2 Sophistication

A final dimension that we explore in this paper in relation to the diversification path is the degree of sophistication and technological content. In section 4 we defined *diversification upgrading* as the introduction of a new exported product of higher sophistication or technological content than the previous exported products or core production activity. The larger the capacity of firms to diversify towards more sophisticated activities, the more firms use the extensive margin of trade to gain value added, and, more importantly, the larger the expected impact of exports on economic growth and development (Hausmann *et al.* 2007).

Section 4 showed that there is larger prevalence of introducing new products with lower sophistication and same or lower technological content than existing exports. However, when compared with the core production activity, there is very high prevalence of diversification upgrading, although this positive upgrading tends to be small in size.

The objective of this section is to identify which of the group of determinants reviewed above are more conducive to one or another type of diversification. In order to do so we estimate equation (4) for both, the PRODY measures and the OECD technology index.

5.2.1 Sophistication vis-a-vis exports

The results of the estimates regarding the sophistication measure in relation to exports are summarised in Table 8 below. The dependent variable is the dichotomous variable T_{it}. This variable has value -1 if the new export implies a sophistication/technology level below the maximum existing in t-1; value zero if implies the same level and 1 if it implies a higher level of sophistication/technology index. The base category in this case is diversification with the same degree of PRODY sophistication. This is almost equivalent to same HS-4 diversification, since only in these cases it is likely that two different products have the same PRODY value. The Pseudo-R^2 is low, around 0.09, also likely the result of significantly reducing the sample size for observations where this measure could be computed.

Since in the case of sophistication the characteristics of the process of production are critical, we also report alternative specifications using a dummy for process innovation.

We focus on statistically significant results robust across specifications. Most of the coefficients are statistically not significant at 95 per cent confidence level. There are two main variables that are consistently significant. Size increases the probability of both upgrading and downgrading diversification, but is larger and statistically significant only for downgrading diversification. Elements that were important explaining relatedness, such as market power, quality and more importantly, the diversification of the production structure, are now not statistically significant.

Table 8 Multinomial logit estimates on the determinants of diversification upgrading (PRODY index) vis-a vis exports

	(1) downgrade	(1) upgrade	(2) downgrade	(2) upgrade	(3) downgrade	(3) upgrade	(4) downgrade	(4) upgrade
TFP using Levinsohn and Petrin and value added	0.0151	-0.1139	-0.0123	-0.1336	0.0024	-0.1265	-0.0367	-0.1605
	(0.1948)	(0.1987)	(0.1978)	(0.2016)	(0.1965)	(0.2003)	(0.1996)	(0.2034)
log employment	0.3979***	0.2144	0.3296**	0.1549	0.4448***	0.2575	0.3728**	0.1862
	(0.1499)	(0.1538)	(0.1505)	(0.1544)	(0.1538)	(0.1576)	(0.1545)	(0.1583)
ratio unit value to product average	-0.0075	-0.0608	-0.0028	-0.0565	-0.0120	-0.0652	-0.0064	-0.0606
	(0.1089)	(0.1143)	(0.1090)	(0.1144)	(0.1089)	(0.1143)	(0.1089)	(0.1143)
firm market share by product	-0.2785	-0.0537	-0.2274	-0.0081	-0.3297	-0.1056	-0.3143	-0.1166
	(0.7378)	(0.7625)	(0.7413)	(0.7657)	(0.7364)	(0.7607)	(0.7364)	(0.7601)
Herfindahl concentration normalised of production firm year	-0.4471	-0.1642	-0.4494	-0.1656	-0.4829	-0.1960	-0.4954	-0.2090
	(0.3770)	(0.3881)	(0.3775)	(0.3886)	(0.3774)	(0.3884)	(0.3775)	(0.3885)
distance CNAE 2 digits divisions	-0.0078	-0.0094	-0.0070	-0.0093	-0.0106	-0.0123	-0.0117	-0.0145
	(0.0311)	(0.0321)	(0.0312)	(0.0323)	(0.0313)	(0.0323)	(0.0314)	(0.0324)
(mean) distance	0.0001**	0.0002***	0.0001**	0.0002***	0.0001**	0.0002***	0.0002***	0.0002***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
dummy for product innovation	-0.3404	-0.4200	-0.5307*	-0.6625**				
	(0.3583)	(0.3686)	(0.3163)	(0.3248)				
dummy for process innovation					-0.5306*	-0.4809	-0.7499**	-0.7358**
					(0.2935)	(0.3017)	(0.3103)	(0.3191)
R& D dummy	-0.1197	-0.2056			-0.1959	-0.3435		
	(0.3731)	(0.3853)			(0.3189)	(0.3289)		
Other_Innovation dummy			0.0000	0.0000			0.0000	0.0000
			(0.0000)	(0.0000)			(0.0000)	(0.0000)
significant marketing changes	0.3915	0.4856	0.3202	0.4148	0.4304	0.5011	0.3380	0.4063
	(0.4809)	(0.5026)	(0.4865)	(0.5076)	(0.4788)	(0.5004)	(0.4861)	(0.5069)
number of high skill technical staff	-0.0160	-0.1860	-0.0295	-0.2179	0.0404	-0.1438	0.0226	-0.1830
	(0.6328)	(0.6557)	(0.6347)	(0.6582)	(0.6344)	(0.6571)	(0.6361)	(0.6592)
group_dep1	0.3013	-0.2075	0.3156	-0.1933	0.2890	-0.2141	0.3128	-0.1846
	(0.3717)	(0.3889)	(0.3733)	(0.3903)	(0.3712)	(0.3884)	(0.3730)	(0.3898)
client_dep1			0.1861	0.2761			0.2769	0.3372
			(0.2800)	(0.2884)			(0.2845)	(0.2935)
high information from university			0.2141	0.1398			0.2367	0.1464

			(0.2653)	(0.2737)			(0.2656)	(0.2740)
foreign_cap1	1.1796	3.4634	0.4972	2.8526	0.7946	3.1059	-0.4528	1.5244
	(5.4833)	(5.6027)	(5.2310)	(5.3489)	(5.4111)	(5.5311)	(4.8757)	(5.0018)
independent or group	-0.0984	-0.1443	-0.1039	-0.1402	-0.1204	-0.1662	-0.1324	-0.1790
	(0.3230)	(0.3344)	(0.3259)	(0.3372)	(0.3228)	(0.3341)	(0.3256)	(0.3368)
N	2240		2240		2240		2240	
log likelihood	-1431.922		-1429.164		-1430.902		-1428.420	
Pseudo R2	0.0879		0.0897		0.0886		0.0902	

Exporters that diversify to same level of sophistication are the base category *** significant at 1% confidence level, ** significant at 5% confidence level and * significant at 10% confidence level. Year, region and sector dummy coefficient omitted from the table.

Table 9 Multinomial logit estimates on the determinants of diversification upgrading (OECD technology index) vis-a-vis exports

	(1) downgrade	(1) upgrade	(2) downgrade	(2) Upgrade	(3) downgrade	(3) Upgrade	(4) downgrade	(4) upgrade
TFP using Levinsohn and Petrin and value added	0.1748**	0.0175	0.1780**	0.0273	0.1766**	0.0133	0.1847***	0.0231
	(0.0695)	(0.1212)	(0.0698)	(0.1226)	(0.0695)	(0.1205)	(0.0698)	(0.1221)
log employment	0.0664	0.0797	0.0940*	0.1155	0.0667	0.0876	0.1059**	0.1097
	(0.0518)	(0.0978)	(0.0518)	(0.0982)	(0.0523)	(0.0984)	(0.0523)	(0.0987)
ratio unit value to product average	0.0179	-0.0571	0.0141	-0.0591	0.0190	-0.0600	0.0160	-0.0618
	(0.0378)	(0.0951)	(0.0378)	(0.0947)	(0.0378)	(0.0952)	(0.0378)	(0.0950)
firm market share by product	0.4405	-0.1033	0.4404	-0.0657	0.4478	-0.1088	0.4765*	-0.0892
	(0.2812)	(0.5877)	(0.2813)	(0.5873)	(0.2809)	(0.5889)	(0.2802)	(0.5876)
Herfindahl concentration normalised of production firm year	-0.0200	-0.1436	-0.0275	-0.1603	-0.0255	-0.1546	-0.0321	-0.1635
	(0.1404)	(0.2675)	(0.1406)	(0.2681)	(0.1404)	(0.2682)	(0.1405)	(0.2685)
distance CNAE 2 digits divisions	0.0414***	0.0289	0.0412***	0.0294	0.0415***	0.0288	0.0419***	0.0292
	(0.0114)	(0.0200)	(0.0114)	(0.0199)	(0.0114)	(0.0200)	(0.0114)	(0.0199)
mean) dist	0.0000	0.0001	0.0000	0.0001	0.0000	0.0001	0.0000	0.0001
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
dummy for product innovation	0.1146	-0.4076	0.2330**	-0.1991				
	(0.1291)	(0.2575)	(0.1109)	(0.2069)				
dummy for process innovation					-0.0038	-0.1697	0.0737	-0.1188
					(0.1069)	(0.1996)	(0.1083)	(0.2082)
R& D dummy	0.1714	0.3535			0.2384**	0.1324		
	(0.1345)	(0.2729)			(0.1159)	(0.2198)		
Other_Innovation dummy			0.0000	0.0000			0.0000	0.0000
			(0.0000)	(0.0000)			(0.0000)	(0.0000)
significant marketing changes	0.0366	-0.5349	0.0385	-0.5319	0.0574	-0.5702	0.0690	-0.5579
	(0.1513)	(0.4250)	(0.1517)	(0.4249)	(0.1508)	(0.4246)	(0.1512)	(0.4242)
number of high skill technical staff	0.3818*	-0.3309	0.3914*	-0.3076	0.3880*	-0.3254	0.4044**	-0.3204
	(0.2049)	(0.4935)	(0.2056)	(0.4936)	(0.2053)	(0.4947)	(0.2057)	(0.4940)
group_dep1	0.3810***	-0.1016	0.3787***	-0.1030	0.3732***	-0.0878	0.3595***	-0.0921
	(0.1319)	(0.2861)	(0.1319)	(0.2860)	(0.1317)	(0.2863)	(0.1315)	(0.2859)

client_dep1			-0.0558	-0.0686			-0.0318	-0.0607
			(0.1001)	(0.1891)			(0.1017)	(0.1988)
high information from university			-0.1035	0.0784			-0.0870	0.0641
			(0.0969)	(0.1849)			(0.0967)	(0.1844)
foreign_cap1	-2.0527	0.9411	-1.3140	1.6355	-2.0600	0.6393	-0.7201	1.1541
	(1.8760)	(2.6111)	(1.8171)	(2.5917)	(1.8746)	(2.6349)	(1.7697)	(2.5850)
independent or group	0.1263	0.1662	0.1300	0.1693	0.1243	0.1554	0.1389	0.1582
	(0.1146)	(0.2224)	(0.1151)	(0.2237)	(0.1146)	(0.2224)	(0.1149)	(0.2232)
N	2240		2240		2240		2240	
log likelihood	-1797.63		-1797.22		-1799.26		-1799.94	
Pseudo R2	0.0844		0.0846		0.0836		0.0832	

Exporters that diversify to same level of technology are the base category *** significant at 1% confidence level, ** significant at 5% confidence level and * significant at 10% confidence level. Year, region and sector dummy coefficient omitted from the table

Probably the most important variables for explaining sophistication are the average distance of exports and quality levels. In line with the Allen-Alchien hypothesis, we should expect higher quality levels in exports going to more distant destinations. This is the result of higher per-unit transport costs, but also in the case of Brazil, can be the result of facing more demanding markets in the US and the EU, which are distant markets. The more demanding the new markets, the higher the incentive for upgrading the product exported. Interestingly, we observe positive coefficients for both, diversification upgrading and downgrading. Although the odds for upgrading are larger, we cannot conclude that there is a clear effect on upgrading of firms that export to tougher markets. This may be explained by two factors. First, exports to Asia may make distance a bad proxy of the toughness of market demand. Second, the Brazilian export pattern is biased towards primary commodities and processed natural resources exported to developed and Asian markets, while manufactures are exported to the region. Therefore, there may be little scope for quality differentiation in distant markets.

A puzzling result, however, is the negative sign on both product and process innovation for both types of diversification, although some of the coefficients are not statistically significant.

Table 9 shows the estimates when using the technological content index. Now the control group, diversification at the same technological level, is around 60 per cent of the observations. We obtain statistically significant results only for the downgrade categories. Concretely, more productive, more diversified, with higher technical staff and with larger dependence on a group of firms have larger probability of technology downgrading in the new product introduced for export.

5.2.2 Sophistication vis-a-vis core activity

Table A4.1 in Appendix 4 shows the estimates of the determinants of sophistication upgrading and downgrading when this is calculated vis-a-vis the main core production activity. Now, around 91 per cent of cases imply some degree of diversification upgrading, and only 2 per cent stay in the same sophistication level. The Pseudo R^2 is still low, and if we focus on statistically significant coefficients we find that both, quality and market power, increase the probability of both upgrading and downgrading, and dependency on a client or buyer increases the probability of upgrading.

These results contrast with Table A4.2, where we look at the OECD technology index. Paradoxically now, market power decreases the probability of upgrading, while business concentration, proxied by the Herfindahl index, increases the probability of upgrading.

One of the problems of the methodology used so far is the significant loss of observations when looking at diversification cases only. This is due to the fact that we focus only on those firms that diversify, and from these cases we use only observations where we can map the sophistication or technology index. As a result the sample size is significantly reduced to slightly more than 2000 observations. More importantly, it is possible that the sample that we are using is non-random, since we are omitting all exporters that do not diversify, and variables that explain diversification and, therefore, being in the sample, also explain the different types of diversification.

5.2.3 Relatedness and Sophistication as a joint decision

One way of addressing this potential sample selection problem and to learn more about the potential determinants of sophistication is to analyse relatedness and sophistication jointly. In fact, one can understand the firm's efforts towards specific diversification paths in term of relatedness and sophistication as a single decision. In other words, it is the amount of available capabilities, innovation efforts to gain new capabilities and firm characteristics which jointly determine the diversification path in relation to both, sophistication and relatedness.

The advantage of estimating relatedness and sophistication jointly is the fact that we can use the entire sample, including those exporters that do not diversify. In order to do so, we first construct a new index to measure the diversification path, RS_{it}. The index has value 0 if the firm *i* is an exporter in *t* but it does not diversify; value 1 if the firm diversifies towards an unrelated product of lower technology/sophistication level; value 2 if the product is highly related and of higher technological content; value 3 if the firm diversifies towards an unrelated product of lower sophistication/technological content, and; value 4 if diversification occurs towards an unrelated and higher technology/sophistication product. Then we estimate equation *(4)* using the RS_{it} index and a multinomial Logit estimator.

Tables A4.3 to A4.10 in Appendix 4 show the results for all the combinations of Leontief input use measure and HS-2 sector differences as methodologies for relatedness, and the OECD index and PRODY for sophistication and technology. We also calculate the indexes in relation to exports and to core production activity.

Given the very large amount of information we focus on those variables that are consistently statistically significant across specifications and that tend to explain the differences in upgrading vs downgrading. The estimates, however, do not identify any clear candidates for explaining upgrading/downgrading. For example, group dependency explains downgrading when using the Leontief and the OECD index, and in relation to exports, but the results is not robust across specifications. In general, we find that when estimating relatedness and sophistication jointly, the results are very sensitive to the type of index and specification used, and there is no clear indication of what measures are more conducive towards upgrading/downgrading diversification.

For the PRODY measures, we also re-estimate the model using a fixed effects estimator on the PRODY difference measure. None of the results show any variable statistically significant influencing upgrading or downgrading.

Summing up, the results of the estimates seem to suggest that the level of sophistication of the diversification path is likely to be determined by other factors not included in our model.

6 Conclusions and Policy Implications

An increasing number of papers support the idea that the type of products that a country export matters for economic growth and economic development. However, we know very little about the capacity of firms to diversify to different and more sophisticated products, and more importantly, what type of firm activities and characteristics are more conducive to achieve this type of diversification. This is especially important in developing countries, given the characteristics of their export baskets; typically heavily concentrated in a few commodities linked to natural resources and with low value added.

This paper has contributed to fill the gap in this area by analysing the degree of relatedness and sophistication of firms export diversification paths. The analysis used a unique dataset that links export, production and innovation data at the firm level in Brazil, for the period 2000-2009.

A first characterisation of the export diversification path of Brazilian firms along these two dimensions suggests the following stylised facts:

- Most diversification activity occurs in highly related or similar products to existing exports, but they tend to be unrelated to the core competences of the firm;
- Some of the cases of unrelated diversification are potentially the result of firms exporting in different stages of the same value chain.
- In most cases diversification occurs to less sophisticated or technology intensive products than existing ones, and to more sophisticated or technological intensive products than the core competences of the firm.

More importantly, the findings of the paper contribute to improving our understanding of the main firm determinants of the diversification path. Concretely, the econometric analysis showed some interesting conclusions. We found that the most important variable explaining the ability of firms to achieve unrelated diversification is the existing scope of diversification in production. This reflects path dependency and evolution on business strategies geared towards expanding the range on unrelated exports of the firm.

As suggested above, however, for some cases the degree of unrelatedness corresponds to products of different sectors that can be part of the same value chain. It is, therefore, possible that some cases of unrelated diversification could be considered as related diversification if we assume that activities within a value chains may require similar capabilities. In these cases, the diversification of the production base represents firms that have opted to acquire capabilities in several stages of the value chain.

Other findings suggest that firms with larger market power are more likely to diversify to unrelated activities, especially in relation to core activities; and that firms that introduced product innovations are also more likely to diversify to unrelated activities. For trade classification based measures, client dependency reduces the extent of unrelated diversification, suggesting greater incentives to diversify within the same trade sector. The results on size, productivity and foreign ownership vary according to the classification used and the coefficients are not always statistically significant.

Comparing the determinants of related diversification in relation to exports or in relation to the core production activity suggest that the quality of products proxied by their unit value appears to matter only for unrelated diversification in relation to exports. This may signal that only higher quality firms can expand to unrelated activities in international markets. The effect of group dependency also seems to matter for relatedness in relation to exports. This suggests that these factors may be important to acquire specific capabilities that allow unrelated diversification in international markets.

We have also attempted to explain the determinants of upgrading/downgrading in sophistication and technological content. The results, however, do not point out to any of the five groups of firm determinants robustly explaining upgrading/downgrading. The level of sophistication of the diversification path is likely to be determined by other factors not captured by our firm determinants.

In general, the main policy implication of these findings is the fact that domestic firm production dynamics, such as market power or sector diversification, are important determinants of the diversification path in relation to how related are new activities to existing ones. This suggests that export diversification support policies should also consider the extent to which firms diversify and improve their production base domestically prior to exporting.

The paper raises some important questions for further research. First, our focus has been mainly on acquiring capabilities via innovation efforts and certain learning activities. More work is required to understand the role of other forms of acquiring capabilities such as technology diffusion, FDI links or trade networks. Second, a better understanding of the specific processes through which firms are able to diversify to more sophisticated products is needed.

References

Aitken, B., Hanson, G. H., and Harrison, A. E. (1997). 'Spillovers, Foreign Investment, and Export Behavior', *Journal of International Economics*, 43.1-2: 103-32

Al-Marhubi, F. (2000). 'Export Diversification and Growth: An Empirical Investigation', *Applied Economics Letters*, 7.9: 559-62

Alchian and Allen. (1964) University Economics. Belmont, CA: Wadsworth Publishing Company, 1964

Besedes, T and Prusa, T. (2006) 'Ins, Outs, and the Duration of Trade', *Canadian Journal of Economics*, 39.1: 266-95.

Bryce, D. J and Winter, S. G. (2009) 'A General Interindustry Relatedness Index', *Management Science*, 55.9: 1570-85.

Cirera, X., Marin, A. and R. Markwald (2011) "Firm Behaviour and the Introduction of New exports: Evidence from Brazil" *IDS Working Paper*

Fan, J. P. H and Lang, L. H. P. (2000) 'The Measurement of Relatedness: An Application to Corporate Diversification', *Journal of Business*, 73.4: 629-60.

Ghosh and Ostry. (1994) "Export Instability and the External Balance in Developing Countries," *IMF Working Paper* No. 94/8, International Monetary Fund, Washington DC.

Haddad, M. E., Lim, J. J and Saborowski, C. (2009) 'Trade Openness Reduces Growth Volatility When Countries Are Well Diversified', *mimeo.*

Hatzichronoglou, T. (1999) *The Globalisation of Industry in the OECD Countries*, OECD Science, Technology and Industry Working Papers, 1999/02, OECD Publishing.

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

Hausmann, R and Rodrik, D. (2003) 'Economic Development as Self-Discovery', *Journal of Development Economics*, 72.2: 603-33.

Herzer, D and Nowak-Lehnmann, F. (2006) 'What Does Export Diversification Do for Growth? An Econometric Analysis', *Applied Economics*, 38.15: 1825-38.

Hesse, H. (2009) 'Export Diversication and Economic Growth', in Newfarmer, R. S., Shaw, W and Walkenhorst, P (eds), *Breaking into New Markets: Emerging Lessons for Export Diversification,* Washington, DC: The World Bank.

Hidalgo, C. A., Klinger, B., Barabasi, A. L and Hausmann, R. (2007) 'The Product Space Conditions the Development of Nations', *Science*, 317.

Hidalgo, C. A and Hausmann, R. (2009) 'The Building Blocks of Economic Complexity', *Quantative Finance Papers 0909.3890,* arXiv.org.

Imbs and Wacziarg. (2003) "Stages of Diversification," *American Economic Review,* vol. 93(1), pages 63-86

Lall, S., Weiss, J and Zhang, J. (2006) 'The Sophistication of Exports: A New Trade Measure', *World Development*, 34.2: 222-37.

Lien, L. B and Klein, P. G. (2009) 'Using Competition to Measure Relatedness', *Journal of Management*, 35.4: 1078-107.

____. 2010. *Why There? Decomposing the Choice of Target Industry*, SSRN eLibrary.

Martincus, C. V and Carballo, J. (2009) *Survival of New Exporters in Developing Countries: Does It Matter How They Diversify?*, IADB Working Paper Series 140.

Neffke, F and Svensson Henning, M. (2008). *Revealed Relatedness: Mapping Industry Space*, Papers in Evolutionary Economic Geography 0819, Utrecht University, Section of Economic Geography.

Teece, D. J., Rumelt, R., Dosi, G and Winter, S. (1994) 'Understanding Corporate Coherence: Theory and Evidence', *Journal of Economic Behavior & Organization*, 23.1: 1-30.

Volker, N. (2009) 'Die Another Day: Duration in German Import Trade', *Review of World Economics (Weltwirtschaftliches Archiv)*, 145.1: 133-54.

Appendix 1 Data Sources

We use the following databases:

PIA (Pesquisa Industrial Anual)

PIA is a firm survey for manufacturing and mining sectors conducted annually by IBGE (Instituto Brasileiro de Geografia e Estatistica). PIA has two different modules, *PIA empresa*, which focus on firm characteristics, and *PIA produto*, which describes the production and sales portfolio for each firm.

It surveys firms in the formal sector with tax identification number, and with a core activity in manufacturing or mining. Firms with 30 or more employees are included in the sample, while smaller firms up to 29 workers are included randomly in the sample. In total PIA covers more than 40,000 firms. PIA produto is based on the PIA empresa sample. However, before 2004 only the largest firms from PIA empresa were included.

PINTEC (Pesquisa de Inovação Tecnológica)

PINTEC is an innovation survey based on the CIS-4 surveys of the European Union. It provides detailed information on R&D expenditure and innovation processes for a sample of firms. Firms with more than 500 workers are automatically included in the sample, while firms from 5 to 499 workers are included randomly. PINTEC is available for 2000, although with a different questionnaire, as well as 2003 and 2005.

SECEX (Secretaria Comercio Exterior)

SECEX provides the universe of registered trade flows at the firm level, by HS-8 product and market destination. The dataset used aggregates export fob values per year, product and destination.

Due to its most restrictive sampling methodology, estimations are based on the sample of firms surveyed in PINTEC. However, the overall dataset includes all the data available. When merging PIA and PINTEC, 73 per cent of observations from PINTEC are matched with PIA data. Interestingly, all exporters from SECEX have been surveyed by PIA or PINTEC, and they represent 17 per cent of the overall sample.

Appendix 2 Relatedness in Diversification and Sector Composition

Table A2.1 Export diversification towards dissimilar products by HS-2 chapter (number of firms)

Core HS2	des	total	HS-2 differ	% new products in HS-2 chapter	% of all dissimilar new products	Input use dissimilarity	% new products in HS-2 chapter	% of all dissimilar new products	hidalgo dissimilarity	% new products in HS-2 chapter	% of all dissimilar new products
02	MEAT AND EDIBLE MEAT OFFAL	113	28	24.78%	0.69%	6	5.31%	0.12%	18	15.93%	0.26%
03	FISH AND CRUSTACEANS, MOLLUSCS AND OTHER AQUATIC INVERTEBRATES	68	4	5.88%	0.10%	7	10.29%	0.14%	25	36.76%	0.36%
04	DAIRY PRODUCE; BIRDS' EGGS; NATURAL HONEY; EDIBLE PRODUCTS OF ANIMAL ORIGIN, NOT	27	9	33.33%	0.22%	8	29.63%	0.16%	5	18.52%	0.07%
05	PRODUCTS OF ANIMAL ORIGIN, NOT ELSEWHERE SPECIFIED OR INCLUDED	32	6	18.75%	0.15%	5	15.63%	0.10%	4	12.50%	0.06%
07	EDIBLE VEGETABLES AND CERTAIN ROOTS AND TUBERS	5		0.00%	0.00%		0.00%	0.00%		0.00%	0.00%
08	EDIBLE FRUIT AND NUTS; PEEL OF CITRUS FRUITS OR MELONS	37	15	40.54%	0.37%	6	16.22%	0.12%	11	29.73%	0.16%
09	COFFEE, TEA, MAT+ AND SPICES	32	14	43.75%	0.35%	14	43.75%	0.28%	13	40.63%	0.19%
10	CEREALS	24	9	37.50%	0.22%	6	25.00%	0.12%	11	45.83%	0.16%
11	PRODUCTS OF THE MILLING INDUSTRY; MALT; STARCHES; INULIN; WHEAT GLUTEN	18	6	33.33%	0.15%		0.00%	0.00%	5	27.78%	0.07%
12	OIL SEEDS AND OLEAGINOUS FRUITS; MISCELLANEOUS GRAINS, SEEDS AND FRUIT; INDUSTRI	40	24	60.00%	0.59%	20	50.00%	0.40%	11	27.50%	0.16%
13	LAC; GUMS, RESINS AND OTHER VEGETABLE SAPS AND EXTRACTS	33	11	33.33%	0.27%	11	33.33%	0.22%	9	27.27%	0.13%
14	VEGETABLE PLAITING MATERIALS; VEGETABLE PRODUCTS NOT ELSEWHERE SPECIFIED OR INCL	4		0.00%	0.00%		0.00%	0.00%		0.00%	0.00%
15	ANIMAL OR VEGETABLE FATS AND OILS AND THEIR CLEAVAGE PRODUCTS; PREPARED EDIBLE F	78	33	42.31%	0.82%	22	28.21%	0.43%	26	33.33%	0.38%
16	PREPARATIONS OF MEAT, OF FISH OR OF CRUSTACEANS, MOLLUSCS OR OTHER AQUATIC INVER	21	9	42.86%	0.22%	8	38.10%	0.16%	7	33.33%	0.10%
17	SUGARS AND SUGAR CONFECTIONERY	132	73	55.30%	1.80%	56	42.42%	1.11%	34	25.76%	0.49%
18	COCOA AND COCOA PREPARATIONS	24	12	50.00%	0.30%	4	16.67%	0.08%	8	33.33%	0.12%
19	PREPARATIONS OF CEREALS, FLOUR, STARCH OR	77	15	19.48%	0.37%	6	7.79%	0.12%	20	25.97%	0.29%

	MILK; PASTRYCOOKS' PRODUCTS										
20	PREPARATIONS OF VEGETABLES, FRUIT, NUTS OR OTHER PARTS OF PLANTS	116	33	28.45%	0.82%	16	13.79%	0.32%	26	22.41%	0.38%
21	MISCELLANEOUS EDIBLE PREPARATIONS	89	46	51.69%	1.14%	19	21.35%	0.38%	24	26.97%	0.35%
22	BEVERAGES, SPIRITS AND VINEGAR	94	44	46.81%	1.09%	37	39.36%	0.73%	32	34.04%	0.46%
23	RESIDUES AND WASTE FROM THE FOOD INDUSTRIES; PREPARED ANIMAL FODDER	88	40	45.45%	0.99%	27	30.68%	0.53%	30	34.09%	0.43%
24	TOBACCO AND MANUFACTURED TOBACCO SUBSTITUTES	15	3	20.00%	0.07%		0.00%	0.00%	8	53.33%	0.12%
25	SALT; SULPHUR; EARTHS AND STONE; PLASTERING MATERIALS, LIME AND CEMENT	73	29	39.73%	0.72%	31	42.47%	0.61%	6	8.22%	0.09%
26	ORES, SLAG AND ASH	7		0.00%	0.00%		0.00%	0.00%	3	42.86%	0.04%
27	MINERAL FUELS, MINERAL OILS AND PRODUCTS OF THEIR DISTILLATION; BITUMINOUS SUBST	56	17	30.36%	0.42%	15	26.79%	0.30%	12	21.43%	0.17%
28	INORGANIC CHEMICALS; ORGANIC OR INORGANIC COMPOUNDS OF PRECIOUS METALS, OF RARE-	109	34	31.19%	0.84%	32	29.36%	0.63%	31	28.44%	0.45%
29	ORGANIC CHEMICALS	159	40	25.16%	0.99%	41	25.79%	0.81%	82	51.57%	1.19%
30	PHARMACEUTICAL PRODUCTS	204	37	18.14%	0.91%	50	24.51%	0.99%	42	20.59%	0.61%
31	FERTILISERS	38	12	31.58%	0.30%	12	31.58%	0.24%	18	47.37%	0.26%
32	TANNING OR DYEING EXTRACTS; TANNINS AND THEIR DERIVATIVES; DYES, PIGMENTS AND OT	262	67	25.57%	1.66%	105	40.08%	2.07%	126	48.09%	1.82%
33	ESSENTIAL OILS AND RESINOIDS; PERFUMERY, COSMETIC OR TOILET PREPARATIONS	161	41	25.47%	1.01%	44	27.33%	0.87%	73	45.34%	1.06%
34	SOAP, ORGANIC SURFACE-ACTIVE AGENTS, WASHING PREPARATIONS, LUBRICATING PREPARATI	92	31	33.70%	0.77%	34	36.96%	0.67%	50	54.35%	0.72%
35	ALBUMINOIDAL SUBSTANCES; MODIFIED STARCHES; GLUES; ENZYMES	65	27	41.54%	0.67%	24	36.92%	0.47%	25	38.46%	0.36%
36	EXPLOSIVES; PYROTECHNIC PRODUCTS; MATCHES; PYROPHORIC ALLOYS; CERTAIN COMBUSTIBL	14	9	64.29%	0.22%	5	35.71%	0.10%	3	21.43%	0.04%
37	PHOTOGRAPHIC OR CINEMATOGRAPHIC GOODS	27	10	37.04%	0.25%	11	40.74%	0.22%	11	40.74%	0.16%
38	MISCELLANEOUS CHEMICAL PRODUCTS	332	105	31.63%	2.59%	132	39.76%	2.61%	140	42.17%	2.02%
39	PLASTICS AND ARTICLES THEREOF	755	175	23.18%	4.32%	228	30.20%	4.50%	297	39.34%	4.29%
40	RUBBER AND ARTICLES THEREOF	196	60	30.61%	1.48%	71	36.22%	1.40%	59	30.10%	0.85%
41	RAW HIDES AND SKINS (OTHER THAN FURSKINS) AND LEATHER	215	22	10.23%	0.54%	27	12.56%	0.53%	64	29.77%	0.92%

42	ARTICLES OF LEATHER; SADDLERY AND HARNESS; TRAVEL GOODS, HANDBAGS AND SIMILAR CO	74	31	41.89%	0.77%	26	35.14%	0.51%	26	35.14%	0.38%
43	FURSKINS AND ARTIFICIAL FUR; MANUFACTURES THEREOF	11	4	36.36%	0.10%	4	36.36%	0.08%	3	27.27%	0.04%
44	WOOD AND ARTICLES OF WOOD; WOOD CHARCOAL	883	47	5.32%	1.16%	62	7.02%	1.22%	346	39.18%	5.00%
45	CORK AND ARTICLES OF CORK	5		0.00%	0.00%	3	60.00%	0.06%		0.00%	0.00%
47	PULP OF WOOD OR OF OTHER FIBROUS CELLULOSIC MATERIAL; RECOVERED (WASTE AND SCRAP	3	3	100.00%	0.07%		0.00%	0.00%		0.00%	0.00%
48	PAPER AND PAPERBOARD; ARTICLES OF PAPER PULP, OF PAPER OR OF PAPERBOARD	243	65	26.75%	1.61%	77	31.69%	1.52%	112	46.09%	1.62%
49	PRINTED BOOKS, NEWSPAPERS, PICTURES AND OTHER PRODUCTS OF THE PRINTING INDUSTRY;	69	18	26.09%	0.44%	19	27.54%	0.38%	18	26.09%	0.26%
50	SILK	7		0.00%	0.00%		0.00%	0.00%	3	42.86%	0.04%
51	WOOL, FINE OR COARSE ANIMAL HAIR; HORSEHAIR YARN AND WOVEN FABRIC	12	5	41.67%	0.12%		0.00%	0.00%	5	41.67%	0.07%
52	COTTON	163	24	14.72%	0.59%	11	6.75%	0.22%	30	18.40%	0.43%
53	OTHER VEGETABLE TEXTILE FIBRES; PAPER YARN AND WOVEN FABRICS OF PAPER YARN	9	4	44.44%	0.10%		0.00%	0.00%		0.00%	0.00%
54	STRIP AND THE LIKE OF MAN-MADE TEXTILE MATERIALS	123	33	26.83%	0.82%	12	9.76%	0.24%	58	47.15%	0.84%
55	MAN-MADE STAPLE FIBRES	35	13	37.14%	0.32%	5	14.29%	0.10%	18	51.43%	0.26%
56	WADDING, FELT AND NONWOVENS; SPECIAL YARNS; TWINE, CORDAGE, ROPES AND CABLES AND	66	18	27.27%	0.44%	21	31.82%	0.41%	19	28.79%	0.27%
57	CARPETS AND OTHER TEXTILE FLOOR COVERINGS	17	4	23.53%	0.10%		0.00%	0.00%	6	35.29%	0.09%
58	SPECIAL WOVEN FABRICS; TUFTED TEXTILE FABRICS; LACE; TAPESTRIES; TRIMMINGS; EMBR	31	12	38.71%	0.30%	11	35.48%	0.22%	11	35.48%	0.16%
59	IMPREGNATED, COATED, COVERED OR LAMINATED TEXTILE FABRICS; TEXTILE ARTICLES OF A	91	39	42.86%	0.96%	29	31.87%	0.57%	42	46.15%	0.61%
60	KNITTED OR CROCHETED FABRICS	86	13	15.12%	0.32%	10	11.63%	0.20%	21	24.42%	0.30%
61	ARTICLES OF APPAREL AND CLOTHING ACCESSORIES, KNITTED OR CROCHETED	521	83	15.93%	2.05%	91	17.47%	1.80%	242	46.45%	3.50%
62	ARTICLES OF APPAREL AND CLOTHING ACCESSORIES, NOT KNITTED OR CROCHETED	369	78	21.14%	1.93%	66	17.89%	1.30%	147	39.84%	2.12%
63	OTHER MADE-UP TEXTILE ARTICLES; SETS; WORN CLOTHING AND WORN TEXTILE ARTICLES; R	101	36	35.64%	0.89%	22	21.78%	0.43%	47	46.53%	0.68%
64	FOOTWEAR, GAITERS AND THE LIKE; PARTS OF SUCH ARTICLES	676	92	13.61%	2.27%	104	15.38%	2.05%	80	11.83%	1.16%

65	HEADGEAR AND PARTS THEREOF	16	10	62.50%	0.25%	10	62.50%	0.20%	5	31.25%	0.07%
68	ARTICLES OF STONE, PLASTER, CEMENT, ASBESTOS, MICA OR SIMILAR MATERIALS	312	92	29.49%	2.27%	106	33.97%	2.09%	54	17.31%	0.78%
69	CERAMIC PRODUCTS	105	50	47.62%	1.24%	38	36.19%	0.75%	36	34.29%	0.52%
70	GLASS AND GLASSWARE	93	22	23.66%	0.54%	33	35.48%	0.65%	49	52.69%	0.71%
71	NATURAL OR CULTURED PEARLS, PRECIOUS OR SEMI-PRECIOUS STONES, PRECIOUS METALS, M	213	62	29.11%	1.53%	68	31.92%	1.34%	60	28.17%	0.87%
72	IRON AND STEEL	112	38	33.93%	0.94%	36	32.14%	0.71%	24	21.43%	0.35%
73	ARTICLES OF IRON OR STEEL	411	142	34.55%	3.51%	172	41.85%	3.40%	161	39.17%	2.33%
74	COPPER AND ARTICLES THEREOF	61	24	39.34%	0.59%	23	37.70%	0.45%	15	24.59%	0.22%
76	ALUMINIUM AND ARTICLES THEREOF	141	48	34.04%	1.19%	48	34.04%	0.95%	48	34.04%	0.69%
78	LEAD AND ARTICLES THEREOF	6		0.00%	0.00%		0.00%	0.00%		0.00%	0.00%
79	ZINC AND ARTICLES THEREOF	4	3	75.00%	0.07%	3	75.00%	0.06%		0.00%	0.00%
80	TIN AND ARTICLES THEREOF	6	3	50.00%	0.07%		0.00%	0.00%	3	50.00%	0.04%
81	OTHER BASE METALS; CERMETS; ARTICLES THEREOF	13	8	61.54%	0.20%	8	61.54%	0.16%	5	38.46%	0.07%
82	TOOLS, IMPLEMENTS, CUTLERY, SPOONS AND FORKS, OF BASE METAL; PARTS THEREOF OF BA	209	63	30.14%	1.56%	80	38.28%	1.58%	82	39.23%	1.19%
83	MISCELLANEOUS ARTICLES OF BASE METAL	146	59	40.41%	1.46%	60	41.10%	1.19%	64	43.84%	0.92%
84	NUCLEAR REACTORS, BOILERS, MACHINERY AND MECHANICAL APPLIANCES; PARTS THEREOF	3,508	632	18.02%	15.62%	1,184	33.75%	23.39%	1,710	48.75%	24.71%
85	ELECTRICAL MACHINERY AND EQUIPMENT AND PARTS THEREOF; SOUND RECORDERS AND REPROD	1,247	245	19.65%	6.05%	427	34.24%	8.44%	570	45.71%	8.24%
86	RAILWAY OR TRAMWAY LOCOMOTIVES, ROLLING-STOCK AND PARTS THEREOF; RAILWAY OR TRAM	28	3	10.71%	0.07%	5	17.86%	0.10%	10	35.71%	0.14%
87	VEHICLES OTHER THAN RAILWAY OR TRAMWAY ROLLING-STOCK, AND PARTS AND ACCESSORIES	683	198	28.99%	4.89%	280	41.00%	5.53%	295	43.19%	4.26%
88	AIRCRAFT, SPACECRAFT, AND PARTS THEREOF	42	13	30.95%	0.32%	19	45.24%	0.38%	27	64.29%	0.39%
89	SHIPS, BOATS AND FLOATING STRUCTURES	3		0.00%	0.00%		0.00%	0.00%		0.00%	0.00%
90	OPTICAL, PHOTOGRAPHIC, CINEMATOGRAPHIC, MEASURING, CHECKING, PRECISION, MEDICAL	637	135	21.19%	3.34%	221	34.69%	4.37%	305	47.88%	4.41%
91	CLOCKS AND WATCHES AND PARTS THEREOF	10	4	40.00%	0.10%	3	30.00%	0.06%		0.00%	0.00%
92	MUSICAL INSTRUMENTS; PARTS AND ACCESSORIES OF SUCH ARTICLES	33	11	33.33%	0.27%	12	36.36%	0.24%	10	30.30%	0.14%

93	ARMS AND AMMUNITION; PARTS AND ACCESSORIES THEREOF	20	13	65.00%	0.32%	13	65.00%	0.26%	11	55.00%	0.16%
94	FURNITURE; BEDDING, MATTRESSES, MATTRESS SUPPORTS, CUSHIONS AND SIMILAR STUFFED	720	188	26.11%	4.65%	254	35.28%	5.02%	236	32.78%	3.41%
95	TOYS, GAMES AND SPORTS REQUISITES; PARTS AND ACCESSORIES THEREOF	115	35	30.43%	0.86%	42	36.52%	0.83%	50	43.48%	0.72%
96	MISCELLANEOUS MANUFACTURED ARTICLES	110	44	40.00%	1.09%	48	43.64%	0.95%	49	44.55%	0.71%
99	OTHER PRODUCTS	50	30	60.00%	0.74%	8	16.00%	0.16%	5	10.00%	0.07%
	UNKNOWN	744	100	13.44%	2.47%	148	19.89%	2.92%	432	58.06%	6.24%
	TOTAL	17325	4047	23.36%		5062	29.22%		6919	39.94%	

Table A2.2 Export diversification and relatedness vis-a-vis core production activity – by CNAE-2 sector (number of firms)

Core CNAE-2 sector	Description	input usage		CNAE-2 distance		CNAE-3 distance	
		Similar	Dissimilar	Similar	Dissimilar	Similar	Dissimilar
11	FABRICAÇÃO DE BEBIDAS		12		14		14
13	FABRICAÇÃO DE PRODUTOS DO FUMO		6		10		11
14	FABRICAÇÃO DE PRODUTOS TÊXTEIS	25	18	25	29	23	31
15	CONFECÇÃO DE ARTIGOS DO VESTUÁRIO E ACESSÓRIOS	462	221	461	356	336	481
16	PREPARAÇÃO DE COUROS E FABRICAÇÃO DE ARTEFATOS DE COURO, ARTIGOS PARA VIAGEM E CALÇADOS	3	14	3	14	3	14
17	FABRICAÇÃO DE PRODUTOS DE MADEIRA	394	194	394	241	249	386
18	FABRICAÇÃO DE CELULOSE, PAPEL E PRODUTOS DE PAPEL	333	142	332	188	279	241
19	IMPRESSÃO E REPRODUÇÃO DE GRAVAÇÕES	139	338	139	358	117	380
20	FABRICAÇÃO DE COQUE, DE PRODUTOS DERIVADOS DO PETRÓLEO E DE BIOCOMBUSTÍVEIS	108	38	108	233	76	265
21	FABRICAÇÃO DE PRODUTOS QUÍMICOS	62	84	62	92	38	116
22	FABRICAÇÃO DE PRODUTOS FARMOQUÍMICOS E FARMACÊUTICOS	15	44	15	56	8	63
23	FABRICAÇÃO DE PRODUTOS DE BORRACHA E DE MATERIAL PLÁSTICO	9	28	10	34	10	34
24	FABRICAÇÃO DE PRODUTOS DE MINERAIS NÃO-METÁLICOS	256	728	609	509	237	881
25	METALURGIA	142	312	147	348	124	371
26	FABRICAÇÃO DE PRODUTOS DE METAL, EXCETO MÁQUINAS E EQUIPAMENTOS	136	205	138	233	124	247
27	FABRICAÇÃO DE EQUIPAMENTOS DE INFORMÁTICA, PRODUTOS ELETRÔNICOS E ÓPTICOS	87	158	104	176	81	199
28	FABRICAÇÃO DE MÁQUINAS, APARELHOS E MATERIAIS ELÉTRICOS	176	331	177	408	123	462
29	FABRICAÇÃO DE MÁQUINAS E EQUIPAMENTOS	546	1,355	572	1,520	271	1,821
30	FABRICAÇÃO DE VEÍCULOS AUTOMOTORES, REBOQUES E CARROCERIAS	17	53	17	71	15	73
31	FABRICAÇÃO DE OUTROS EQUIPAMENTOS DE TRANSPORTE, EXCETO VEÍCULOS AUTOMOTORES	158	262	158	344	98	404
32	FABRICAÇÃO DE MÓVEIS	40	143	40	204	26	218
33	FABRICAÇÃO DE PRODUTOS DIVERSOS	124	220	124	288	99	313

34	MANUTENÇÃO, REPARAÇÃO E INSTALAÇÃO DE MÁQUINAS E EQUIPAMENTOS	42	540	43	618	41	620
35	ELETRICIDADE, GÁS E OUTRAS UTILIDADES	10	76	10	91	10	91
36	CAPTAÇÃO, TRATAMENTO E DISTRIBUIÇÃO DE ÁGUA	89	314	89	403	86	406
37	ESGOTO E ATIVIDADES RELACIONADAS		3		3		3

Appendix 3 Diversification upgrading by sector

Core hs2	des		PRODY upgrade				OECD upgrade		
		total	firms	% upgrade	% total	firms	% upgrade	% total
2	MEAT AND EDIBLE MEAT OFFAL	113	40	35.40%	0.72%			
3	FISH AND CRUSTACEANS, MOLLUSCS AND OTHER AQUATIC INVERTEBRATES	68	22	32.35%	0.40%			
4	DAIRY PRODUCE; BIRDS' EGGS; NATURAL HONEY; EDIBLE PRODUCTS OF ANIMAL ORIGIN, NOT	27	13	48.15%	0.23%	3	11.11%	0.28%
5	PRODUCTS OF ANIMAL ORIGIN, NOT ELSEWHERE SPECIFIED OR INCLUDED	32	9	28.13%	0.16%	6	18.75%	0.56%
8	EDIBLE FRUIT AND NUTS; PEEL OF CITRUS FRUITS OR MELONS	37	16	43.24%	0.29%	5	13.51%	0.47%
9	COFFEE, TEA, MAT+ AND SPICES	32	15	46.88%	0.27%	8	25.00%	0.74%
10	CEREALS	24	8	33.33%	0.14%	5	20.83%	0.47%
11	PRODUCTS OF THE MILLING INDUSTRY; MALT; STARCHES; INULIN; WHEAT GLUTEN	18	6	33.33%	0.11%			
12	OIL SEEDS AND OLEAGINOUS FRUITS; MISCELLANEOUS GRAINS, SEEDS AND FRUIT; INDUSTRI	40	19	47.50%	0.34%	13	32.50%	1.21%
13	LAC; GUMS, RESINS AND OTHER VEGETABLE SAPS AND EXTRACTS	33	6	18.18%	0.11%	3	9.09%	0.28%
15	ANIMAL OR VEGETABLE FATS AND OILS AND THEIR CLEAVAGE PRODUCTS; PREPARED EDIBLE F	78	35	44.87%	0.63%	8	10.26%	0.74%
16	PREPARATIONS OF MEAT, OF FISH OR OF CRUSTACEANS, MOLLUSCS OR OTHER AQUATIC INVER	21	8	38.10%	0.14%			
17	SUGARS AND SUGAR CONFECTIONERY	132	61	46.21%	1.10%	31	23.48%	2.89%
18	COCOA AND COCOA PREPARATIONS	24	8	33.33%	0.14%			
19	PREPARATIONS OF CEREALS, FLOUR, STARCH OR MILK; PASTRYCOOKS' PRODUCTS	77	38	49.35%	0.69%			
20	PREPARATIONS OF VEGETABLES, FRUIT, NUTS OR OTHER PARTS OF PLANTS	116	51	43.97%	0.92%	6	5.17%	0.56%
21	MISCELLANEOUS EDIBLE PREPARATIONS	89	23	25.84%	0.42%	10	11.24%	0.93%
22	BEVERAGES, SPIRITS AND VINEGAR	94	34	36.17%	0.61%	8	8.51%	0.74%
23	RESIDUES AND WASTE FROM THE FOOD INDUSTRIES; PREPARED ANIMAL FODDER	88	29	32.95%	0.52%	12	13.64%	1.12%
24	TOBACCO AND MANUFACTURED TOBACCO SUBSTITUTES	15	3	20.00%	0.05%			
25	SALT; SULPHUR; EARTHS AND STONE; PLASTERING MATERIALS, LIME AND CEMENT	73	41	56.16%	0.74%	23	31.51%	2.14%

27	MINERAL FUELS, MINERAL OILS AND PRODUCTS OF THEIR DISTILLATION; BITUMINOUS SUBST	56	31	55.36%	0.56%	7	12.50%	0.65%
28	INORGANIC CHEMICALS; ORGANIC OR INORGANIC COMPOUNDS OF PRECIOUS METALS, OF RARE-	109	35	32.11%	0.63%	5	4.59%	0.47%
29	ORGANIC CHEMICALS	159	40	25.16%	0.72%	10	6.29%	0.93%
30	PHARMACEUTICAL PRODUCTS	204	41	20.10%	0.74%	9	4.41%	0.84%
31	FERTILISERS	38	17	44.74%	0.31%			
32	TANNING OR DYEING EXTRACTS; TANNINS AND THEIR DERIVATIVES; DYES, PIGMENTS AND OT	262	58	22.14%	1.05%	9	3.44%	0.84%
33	ESSENTIAL OILS AND RESINOIDS; PERFUMERY, COSMETIC OR TOILET PREPARATIONS	161	45	27.95%	0.81%			
34	SOAP, ORGANIC SURFACE-ACTIVE AGENTS, WASHING PREPARATIONS, LUBRICATING PREPARATI	92	27	29.35%	0.49%			
35	ALBUMINOIDAL SUBSTANCES; MODIFIED STARCHES; GLUES; ENZYMES	65	12	18.46%	0.22%	4	6.15%	0.37%
36	EXPLOSIVES; PYROTECHNIC PRODUCTS; MATCHES; PYROPHORIC ALLOYS; CERTAIN COMBUSTIBL	14	4	28.57%	0.07%			
37	PHOTOGRAPHIC OR CINEMATOGRAPHIC GOODS	27	3	11.11%	0.05%	3	11.11%	0.28%
38	MISCELLANEOUS CHEMICAL PRODUCTS	332	98	29.52%	1.77%	15	4.52%	1.40%
39	PLASTICS AND ARTICLES THEREOF	755	232	30.73%	4.19%	57	7.55%	5.31%
40	RUBBER AND ARTICLES THEREOF	196	69	35.20%	1.25%	24	12.24%	2.23%
41	RAW HIDES AND SKINS (OTHER THAN FURSKINS) AND LEATHER	215	190	88.37%	3.43%	4	1.86%	0.37%
42	ARTICLES OF LEATHER; SADDLERY AND HARNESS; TRAVEL GOODS, HANDBAGS AND SIMILAR CO	74	31	41.89%	0.56%	6	8.11%	0.56%
43	FURSKINS AND ARTIFICIAL FUR; MANUFACTURES THEREOF	11	4	36.36%	0.07%			
44	WOOD AND ARTICLES OF WOOD; WOOD CHARCOAL	883	588	66.59%	10.63%	23	2.60%	2.14%
48	PAPER AND PAPERBOARD; ARTICLES OF PAPER PULP, OF PAPER OR OF PAPERBOARD	243	112	46.09%	2.02%	29	11.93%	2.70%
49	PRINTED BOOKS, NEWSPAPERS, PICTURES AND OTHER PRODUCTS OF THE PRINTING INDUSTRY;	69	28	40.58%	0.51%	8	11.59%	0.74%
51	WOOL, FINE OR COARSE ANIMAL HAIR; HORSEHAIR YARN AND WOVEN FABRIC	12	4	33.33%	0.07%			
52	COTTON	163	30	18.40%	0.54%	3	1.84%	0.28%
54	STRIP AND THE LIKE OF MAN-MADE TEXTILE MATERIALS	123	39	31.71%	0.70%	5	4.07%	0.47%
55	MAN-MADE STAPLE FIBRES	35	10	28.57%	0.18%			
56	WADDING, FELT AND NONWOVENS; SPECIAL YARNS; TWINE, CORDAGE, ROPES AND CABLES AND	66	28	42.42%	0.51%	7	10.61%	0.65%

57	CARPETS AND OTHER TEXTILE FLOOR COVERINGS	17	6	35.29%	0.11%			
58	SPECIAL WOVEN FABRICS; TUFTED TEXTILE FABRICS; LACE; TAPESTRIES; TRIMMINGS; EMBR	31	10	32.26%	0.18%			
59	IMPREGNATED, COATED, COVERED OR LAMINATED TEXTILE FABRICS; TEXTILE ARTICLES OF A	91	29	31.87%	0.52%	12	13.19%	1.12%
60	KNITTED OR CROCHETED FABRICS	86	51	59.30%	0.92%			
61	ARTICLES OF APPAREL AND CLOTHING ACCESSORIES, KNITTED OR CROCHETED	521	135	25.91%	2.44%	16	3.07%	1.49%
62	ARTICLES OF APPAREL AND CLOTHING ACCESSORIES, NOT KNITTED OR CROCHETED	369	104	28.18%	1.88%	11	2.98%	1.02%
63	OTHER MADE-UP TEXTILE ARTICLES; SETS; WORN CLOTHING AND WORN TEXTILE ARTICLES; R	101	30	29.70%	0.54%	3	2.97%	0.28%
64	FOOTWEAR, GAITERS AND THE LIKE; PARTS OF SUCH ARTICLES	676	174	25.74%	3.14%	31	4.59%	2.89%
65	HEADGEAR AND PARTS THEREOF	16	3	18.75%	0.05%	3	18.75%	0.28%
68	ARTICLES OF STONE, PLASTER, CEMENT, ASBESTOS, MICA OR SIMILAR MATERIALS	312	118	37.82%	2.13%	12	3.85%	1.12%
69	CERAMIC PRODUCTS	105	43	40.95%	0.78%	11	10.48%	1.02%
70	GLASS AND GLASSWARE	93	29	31.18%	0.52%	7	7.53%	0.65%
71	NATURAL OR CULTURED PEARLS, PRECIOUS OR SEMI-PRECIOUS STONES, PRECIOUS METALS, M	213	94	44.13%	1.70%	42	19.72%	3.91%
72	IRON AND STEEL	112	49	43.75%	0.89%	12	10.71%	1.12%
73	ARTICLES OF IRON OR STEEL	411	142	34.55%	2.57%	56	13.63%	5.21%
74	COPPER AND ARTICLES THEREOF	61	21	34.43%	0.38%	10	16.39%	0.93%
76	ALUMINIUM AND ARTICLES THEREOF	141	42	29.79%	0.76%	8	5.67%	0.74%
80	TIN AND ARTICLES THEREOF	6	3	50.00%	0.05%			
81	OTHER BASE METALS; CERMETS; ARTICLES THEREOF	13	5	38.46%	0.09%	4	30.77%	0.37%
82	TOOLS, IMPLEMENTS, CUTLERY, SPOONS AND FORKS, OF BASE METAL; PARTS THEREOF OF BA	209	43	20.57%	0.78%	26	12.44%	2.42%
83	MISCELLANEOUS ARTICLES OF BASE METAL	146	44	30.14%	0.80%	14	9.59%	1.30%
84	NUCLEAR REACTORS, BOILERS, MACHINERY AND MECHANICAL APPLIANCES; PARTS THEREOF	3,508	765	21.81%	13.82%	137	3.91%	12.76%
85	ELECTRICAL MACHINERY AND EQUIPMENT AND PARTS THEREOF; SOUND RECORDERS AND REPROD	1,247	407	32.64%	7.35%	62	4.97%	5.77%
86	RAILWAY OR TRAMWAY LOCOMOTIVES, ROLLING-STOCK AND PARTS THEREOF; RAILWAY OR TRAM	28	7	25.00%	0.13%			
87	VEHICLES OTHER THAN RAILWAY OR TRAMWAY ROLLING-STOCK, AND PARTS AND ACCESSORIES	683	184	26.94%	3.32%	13	1.90%	1.21%

88	AIRCRAFT, SPACECRAFT, AND PARTS THEREOF	42	12	28.57%	0.22%			
90	OPTICAL, PHOTOGRAPHIC, CINEMATOGRAPHIC, MEASURING, CHECKING, PRECISION, MEDICAL	637	162	25.43%	2.93%	34	5.34%	3.17%
91	CLOCKS AND WATCHES AND PARTS THEREOF	10	4	40.00%	0.07%			
92	MUSICAL INSTRUMENTS; PARTS AND ACCESSORIES OF SUCH ARTICLES	33	10	30.30%	0.18%	4	12.12%	0.37%
93	ARMS AND AMMUNITION; PARTS AND ACCESSORIES THEREOF	20	6	30.00%	0.11%			
94	FURNITURE; BEDDING, MATTRESSES, MATTRESS SUPPORTS, CUSHIONS AND SIMILAR STUFFED	720	288	40.00%	5.20%	83	11.53%	7.73%
95	TOYS, GAMES AND SPORTS REQUISITES; PARTS AND ACCESSORIES THEREOF	115	38	33.04%	0.69%	25	21.74%	2.33%
96	MISCELLANEOUS MANUFACTURED ARTICLES	110	29	26.36%	0.52%	16	14.55%	1.49%
99	OTHER PRODUCTS	50	39	78.00%	0.70%	28	56.00%	2.61%

Appendix 4 Sophistication estimates

Table A4.1 Multinomial logit estimates on the determinants of diversification upgrading (PRODY index) vis-a vis core activity

	(1) downgrade	(1) upgrade	(2) downgrade	(2) upgrade	(3) downgrade	(3) upgrade	(4) downgrade	(4) upgrade
TFP using Levinsohn and Petrin and value added	-0.3945	-0.3940	-0.4032	-0.4106*	-0.4039	-0.4082*	-0.4135	-0.4224*
	(0.2623)	(0.2421)	(0.2651)	(0.2446)	(0.2634)	(0.2432)	(0.2675)	(0.2471)
log employment	0.2990	0.2494	0.2653	0.2845	0.3101	0.2726	0.3408	0.3727*
	(0.2101)	(0.1956)	(0.2086)	(0.1937)	(0.2115)	(0.1966)	(0.2094)	(0.1943)
ratio unit value to product average	0.4966*	0.5077*	0.5027*	0.5079*	0.4800*	0.4866*	0.4764*	0.4801*
	(0.2924)	(0.2884)	(0.2922)	(0.2881)	(0.2902)	(0.2862)	(0.2861)	(0.2818)
firm market share by product	3.4329**	2.3636*	3.4599**	2.4078*	3.4726**	2.3667*	3.6662**	2.6182*
	(1.4664)	(1.4199)	(1.4608)	(1.4140)	(1.4583)	(1.4125)	(1.4698)	(1.4237)
Herfindahl concentration normalised of production firm year	-0.7703	-0.6425	-0.7301	-0.5939	-0.8397	-0.6952	-0.7439	-0.6079
	(0.5561)	(0.5155)	(0.5610)	(0.5203)	(0.5538)	(0.5135)	(0.5582)	(0.5173)
distance CNAE 2 digits divisions	0.0660	0.0848	0.0744	0.0936	0.0655	0.0835	0.0755	0.0943
	(0.0678)	(0.0649)	(0.0686)	(0.0656)	(0.0682)	(0.0653)	(0.0688)	(0.0658)
mean) dist	0.0000	-0.0001	0.0000	-0.0001	-0.0001	-0.0001	0.0000	-0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
dummy for product innovation	0.6192	0.3805	0.5572	0.5660				
	(0.5202)	(0.4856)	(0.4479)	(0.4157)				
dummy for process innovation					-0.3066	-0.4608	-0.3290	-0.4221
					(0.4179)	(0.3872)	(0.4505)	(0.4214)
R& D dummy	-0.1178	0.3889			0.3821	0.7967*		
	(0.5785)	(0.5453)			(0.5025)	(0.4725)		
Other_Innovation dummy			0.0000	0.0000			0.0000	0.0000
			(0.0000)	(0.0000)			(0.0000)	(0.0000)
significant marketing changes			0.6402	0.3693			0.7292*	0.4655

			(0.4039)	(0.3783)			(0.4058)	(0.3803)
number of high skill technical staff	6.9963	2.0015	6.6389	3.5914	6.2474	1.2506	11.4840	8.5053
	(10.7598)	(10.5644)	(10.5939)	(10.3898)	(10.8415)	(10.6463)	(11.6181)	(11.4247)
group_dep1	0.6275	0.4210	0.7004	0.4526	0.8469	0.6215	0.9402	0.7057
	(0.7523)	(0.7232)	(0.7562)	(0.7271)	(0.7449)	(0.7162)	(0.7475)	(0.7181)
client_dep1	20.2831	20.2130***	19.6375	19.6022***	20.0693	19.9914***	20.2787	20.2480***
	.	(0.3170)	.	(0.3171)	.	(0.3162)	.	(0.3164)
high information from university			-0.2262	-0.0974			0.0662	0.2150
			(0.4051)	(0.3765)			(0.4352)	(0.4087)
foreign_cap1	-0.2304	0.0023	-0.2336	-0.0203	-0.2314	0.0102	-0.2977	-0.0895
	(0.5321)	(0.4962)	(0.5379)	(0.5025)	(0.5296)	(0.4941)	(0.5321)	(0.4965)
independent or group	-0.3128	-0.3231	-0.3704	-0.3342	-0.3495	-0.3625	-0.3553	-0.3176
	(0.4745)	(0.4430)	(0.4777)	(0.4462)	(0.4749)	(0.4438)	(0.4766)	(0.4452)
N	2027		2027		2027		2027	
log likelihood	-715.30		-715.73		-715.26		-716.04	
Pseudo R2	0.1260		0.1255		0.1260		0.1251	

Exporters that diversify to same level of sophistication are the base category *** significant at 1% confidence level, ** significant at 5% confidence level and * significant at 10% confidence level. Year, region and sector dummy coefficient omitted from the table.

Table A4.2 Multinomial logit estimates on the determinants of diversification upgrading (OECD technology index) vis-a-vis core activity

	(1) downgrade	(1) Upgrade	(2) downgrade	(2) upgrade	(3) downgrade	(3) upgrade	(4) Downgrade	(4) Upgrade
TFP using Levinsohn and Petrin and value added	-0.0103	-0.0319	0.0073	-0.0293	-0.0047	-0.0321	0.0044	-0.0313
	(0.1795)	(0.0867)	(0.1809)	(0.0868)	(0.1805)	(0.0870)	(0.1816)	(0.0871)
log employment	0.0907	0.0582	0.0257	0.0482	0.0889	0.0714	0.0149	0.0711
	(0.1407)	(0.0670)	(0.1393)	(0.0672)	(0.1431)	(0.0682)	(0.1414)	(0.0684)
ratio unit value to product average	-0.0536	0.0566	-0.0481	0.0587	-0.0523	0.0539	-0.0500	0.0549
	(0.1195)	(0.0519)	(0.1198)	(0.0518)	(0.1192)	(0.0518)	(0.1201)	(0.0518)
firm market share by product	-0.6636	-0.8772**	-0.7125	-0.8787**	-0.6606	-0.8775**	-0.7830	-0.8741**
	(0.7809)	(0.3538)	(0.7843)	(0.3542)	(0.7791)	(0.3537)	(0.7798)	(0.3535)
Herfindahl concentration normalised of production firm year	0.4630	0.3750**	0.4238	0.3641*	0.4663	0.3685**	0.4455	0.3638*
	(0.3901)	(0.1862)	(0.3905)	(0.1864)	(0.3895)	(0.1862)	(0.3901)	(0.1865)
distance CNAE 2 digits divisions	0.0362	0.0260	0.0309	0.0244	0.0359	0.0252	0.0303	0.0238
	(0.0354)	(0.0174)	(0.0353)	(0.0175)	(0.0355)	(0.0175)	(0.0354)	(0.0175)
mean) dist	0.0001	0.0000	0.0001	0.0000	0.0001	0.0000	0.0001*	0.0000
	(0.0001)	(0.0000)	(0.0001)	(0.0000)	(0.0001)	(0.0000)	(0.0001)	(0.0000)
dummy for product innovation	0.0861	0.0719	-0.2374	-0.0131				
	(0.3425)	(0.1711)	(0.3028)	(0.1473)				
dummy for process innovation					0.0285	-0.1571	-0.0852	-0.1992
					(0.2985)	(0.1410)	(0.3017)	(0.1434)
R& D dummy	-0.6948*	-0.1509			-0.6558**	-0.0764		
	(0.3646)	(0.1782)			(0.3237)	(0.1529)		
Other_Innovation dummy			0.0000	0.0000			0.0000	0.0000
			(0.0000)	(0.0000)			(0.0000)	(0.0000)
significant marketing changes			-0.1178	-0.0994			-0.1253	-0.0841
			(0.2687)	(0.1263)			(0.2684)	(0.1263)
number of high skill technical staff	-4.0529	-3.1496	-6.2140	-3.4548*	-4.0940	-3.2243	-6.9595	-3.3124
	(4.4669)	(2.1429)	(4.8383)	(2.0825)	(4.5139)	(2.1437)	(5.0433)	(2.0581)

group_dep1	-0.4165	0.0454	-0.3824	0.0404	-0.4142	0.0772	-0.4162	0.0605
	(0.3858)	(0.1937)	(0.3850)	(0.1942)	(0.3828)	(0.1925)	(0.3826)	(0.1931)
client_dep1	-0.1597	-0.0427	-0.2174	-0.0453	-0.1591	-0.0286	-0.2331	-0.0302
	(0.5852)	(0.2628)	(0.5865	(0.2638	(0.5850)	(0.2629)	(0.5849)	(0.2639)
high information from university			0.0435	0.0973			0.0218	0.1407
			(0.2804	(0.1328			(0.2848)	(0.1349)
foreign_cap1	0.5005	-0.0636	0.5081	-0.0665	0.4979	-0.0707	0.5281	-0.0715
	(0.3338)	(0.1719)	(0.3327	(0.1720	(0.3328)	(0.1717)	(0.3313)	(0.1716)
independent or group	0.4006	-0.0181	0.3678	-0.0109	0.4062	-0.0225	0.3728	-0.0085
	(0.3011)	(0.1519)	(0.3017	(0.1524	(0.3014)	(0.1519)	(0.3016)	(0.1524)
N	2027		2027		2027		2027	
log likelihood	-1121.29		-1122.42		-1120.60		-1121.74	
Pseudo R2	0.0961		0.0952		0.0966		0.0957	

Exporters that diversify to same level of sophistication are the base category *** significant at 1% confidence level, ** significant at 5% confidence level and * significant at 10% confidence level. Year, region and sector dummy coefficient omitted from the table.

Table A4.3 Multinomial logit estimates relatedness (Leontief input use) and technological content (OECD) – exports

	(1)				(2)			
	Related No upgrade	**Related Upgrade**	**Unrelated No upgrade**	**Unrelated Upgrade**	**Related No upgrade**	**Related Upgrade**	**Unrelated No upgrade**	**Unrelated Upgrade**
TFP using Levinsohn and Petrin and value added	0.1584***	0.8277*	0.1791**	0.0426***	0.1604***	0.7991*	0.1792***	0.0400
	(0.0476)	(0.4350)	(0.0690)	(0.1260)	(0.0475)	(0.4337)	(0.0690)	(0.1260)
log employment	0.3530***	0.6941**	0.3165	0.3869	0.3836***	0.8326**	0.3348***	0.3877***
	(0.0362)	(0.3275)	(0.0514)	(0.0984)	(0.0364)	(0.3237)	(0.0518)	(0.0991)
ratio unit value to product average	0.0282	0.2176*	0.0654***	-0.0727***	0.0287	0.1886	0.0705**	-0.0716
	(0.0270)	(0.1285)	(0.0314)	(0.1049)	(0.0270)	(0.1282)	(0.0313)	(0.1050)
firm market share by product	0.5893***	0.6224	0.6680	0.6199***	0.6364***	1.1019	0.7304***	0.6212
	(0.2033)	(1.5005)	(0.2735)	(0.5826)	(0.2022)	(1.4825)	(0.2716)	(0.5816)
Herfindahl concentration normalised of production firm year	-0.5683***	0.1909	-0.1993***	-0.6045***	-0.5875***	0.1361	-0.2214	-0.6093**
	(0.0950)	(0.8895)	(0.1376)	(0.2679)	(0.0948)	(0.8807)	(0.1374)	(0.2678)
distance CNAE 2 digits divisions	-0.0012	-0.0358	0.0442**	0.0264***	-0.0007	-0.0331	0.0450***	0.0260
	(0.0086)	(0.0928)	(0.0114)	(0.0197)	(0.0086)	(0.0909)	(0.0114)	(0.0197)
mean) dist	0.0000	0.0001	0.0000***	0.0000***	0.0000**	0.0001	0.0000	0.0000
	(0.0000)	(0.0001)	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0000)	(0.0000)
dummy for product innovation	0.3923***	0.4505	0.5375	0.1273***				
	(0.0749)	(0.7079)	(0.1082)	(0.2050)				
dummy for process innovation					-0.0082	-1.2817*	0.2550**	0.0782
					(0.0772)	(0.6561)	(0.1097)	(0.2159)
Other_Innovation dummy	0.0000	0.0000	0.0000***	0.0000***	0.0000*	0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
significant marketing changes	0.1384**	-0.0201	0.2240**	0.2745***	0.1811***	0.2088	0.2576***	0.2809
	(0.0680)	(0.6150)	(0.0960)	(0.1844)	(0.0676)	(0.6141)	(0.0956)	(0.1832)
number of high skill technical staff	3.1877**	0.1201	4.2299	6.6813	4.6435***	2.5953	5.7393***	7.3648***

	(1.4344)	(11.5387)	(1.5515)	(2.3763)	(1.4407)	(10.6378)	(1.5713)	(2.4289)
group_dep1	0.2374*	-0.9920	0.3277**	-0.1912***	0.3110**	-0.8398	0.3908**	-0.1809
	(0.1261)	(1.2346)	(0.1608)	(0.4399)	(0.1258)	(1.2469)	(0.1603)	(0.4386)
client_dep1	0.0920	0.2377	-0.1572***	-0.6215***	0.1473	0.4540	-0.1300	-0.6208
	(0.1528)	(1.0981)	(0.2191)	(0.5311)	(0.1529)	(1.1226)	(0.2194)	(0.5306)
high information from university	-0.1068	0.3502	-0.0343***	-0.1192***	-0.0218	0.7025	-0.0028	-0.1226
	(0.0695)	(0.6141)	(0.0994)	(0.1904)	(0.0735)	(0.6218)	(0.1031)	(0.2071)
foreign_cap1	0.3934***	-0.4217	0.4190	0.1856***	0.3717***	-0.6185	0.3971***	0.1876
	(0.0964)	(0.8279)	(0.1320)	(0.2896)	(0.0962)	(0.8490)	(0.1316)	(0.2894)
independent or group	-0.1113	-0.6759	-0.1883***	0.1229***	-0.0830	-0.6157	-0.1634	0.1308
	(0.0837)	(0.7331)	(0.1185)	(0.2233)	(0.0833)	(0.7259)	(0.1183)	(0.2224)
Observations	9077				9077			
Log-likelihood	-5181.58				-5197.44			
Pseudo R2	0.2445				0.2422			

Exporters that do not diversify are the base category *** significant at 1% confidence level, ** significant at 5% confidence level and * significant at 10% confidence level. Year, region and sector dummy coefficient omitted from the table.

Table A4.4 Multinomial logit estimates relatedness (HS-2 difference) and technological content (OECD) – exports

	(1)				(2)			
	Related No upgrade	**Related Upgrade**	**Unrelated No upgrade**	**Unrelated Upgrade**	**Related No upgrade**	**Related Upgrade**	**Unrelated No upgrade**	**Unrelated Upgrade**
TFP using Levinsohn and Petrin and value added	0.1016	0.0627	0.1589***	0.1235	0.1004	0.0818	0.1598***	0.1220
	(0.1372)	(0.4206)	(0.0436)	(0.1266)	(0.1373)	(0.4104)	(0.0435)	(0.1265)
log employment	0.2329**	0.2630	0.3410***	0.4127***	0.2401**	0.3940	0.3677***	0.4173***
	(0.1060)	(0.3082)	(0.0334)	(0.0979)	(0.1065)	(0.3067)	(0.0336)	(0.0983)
ratio unit value to product average	0.0660	0.0881	0.0384	-0.0194	0.0677	0.0579	0.0406*	-0.0199
	(0.0671)	(0.2203)	(0.0240)	(0.0877)	(0.0672)	(0.2362)	(0.0239)	(0.0883)
firm market share by product	1.8025***	-1.2871	0.5356***	0.7044	1.8217***	-1.2458	0.5862***	0.7068
	(0.5096)	(2.5569)	(0.1881)	(0.5558)	(0.5080)	(2.5287)	(0.1867)	(0.5555)
Herfindahl concentration normalised of production firm year	-0.6316**	-1.3787	-0.4646***	-0.4407*	-0.6415**	-1.3825	-0.4846***	-0.4453*
	(0.2900)	(0.9484)	(0.0874)	(0.2633)	(0.2899)	(0.9496)	(0.0871)	(0.2631)
distance CNAE 2 digits divisions	0.0194	0.0527	0.0096)	0.0199	0.0196	0.0576	0.0102	0.0197
	(0.0242)	(0.0661)	(0.0078	(0.0201)	(0.0242)	(0.0667)	(0.0078)	(0.0201)
mean) dist	-0.0001	0.0000	0.0000	0.0000	-0.0001	0.0000	0.0000**	0.0000
	(0.0000)	(0.0001)	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0000)	(0.0000)
dummy for product innovation	0.2207	0.5920	0.4381***	0.0821				
	(0.2256)	(0.6824)	(0.0688)	(0.2028)				
dummy for process innovation					0.0926	-1.0168	0.0777	0.0001
					(0.2300)	(0.7348)	(0.0712)	(0.2112)
Other_Innovation dummy	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
significant marketing changes	0.2303	-0.1947	0.1578**	0.2951	0.2474	-0.0724	0.1979***	0.3013*
	(0.2011)	(0.6351)	(0.0626)	(0.1823)	(0.2005)	(0.6323)	(0.0622)	(0.1812)
number of high skill technical staff	1.1771	-19.0403	3.6578***	6.4427***	2.0638	-5.7103	5.0879***	7.0622***

	(4.8194)	(59.2441)	(1.2749)	(2.3012)	(4.7597)	(39.0607)	(1.3090)	(2.3590)
group_dep1	0.4171	0.0395	0.2866**	-0.2904	0.4475	0.2838	0.3551***	-0.2764
	(0.3565)	(1.2578)	(0.1168)	(0.4389)	(0.3555)	(1.2673)	(0.1165)	(0.4379)
client_dep1	-0.0465	0.6432	0.0305	-0.6861	-0.0261	1.0836	0.0773	-0.6744
	(0.4853)	(1.1496)	(0.1429)	(0.5296)	(0.4853)	(1.1499)	(0.1430)	(0.5291)
high information from university	-0.0176	-0.6661	-0.0899	-0.0366	-0.0006	-0.1521	-0.0255	-0.0159
	(0.2095)	(0.6549)	(0.0640)	(0.1875)	(0.2185)	(0.6957)	(0.0677)	(0.2028)
foreign_cap1	0.0577	0.9300	0.4145***	0.0035	0.0524	0.9492	0.3923***	0.0056
	(0.2915)	(0.8495)	(0.0888)	(0.2878)	(0.2915)	(0.8477)	(0.0886)	(0.2877)
independent or group	0.0447	-0.7593	-0.1356*	0.0985	0.0595	-0.6987	-0.1070	0.1046
	(0.2417)	(0.8357)	(0.0777)	(0.2204)	(0.2414)	(0.8317)	(0.0773)	(0.2195)
Observations	9103				9103			
Log-likelihood	-4461.12				-4479.99			
Pseudo R2	0.2677				0.2646			

Exporters that do not diversify are the base category *** significant at 1% confidence level, ** significant at 5% confidence level and * significant at 10% confidence level. Year, region and sector dummy coefficient omitted from the table.

Table A4.5 Multinomial logit estimates relatedness (Leontief input use) and technological content (OECD) – core production

	(1)				(2)			
	Related No upgrade	**Related Upgrade**	**Unrelated No upgrade**	**Unrelated Upgrade**	**Related No upgrade**	**Related Upgrade**	**Unrelated No upgrade**	**Unrelated Upgrade**
TFP using Levinsohn and Petrin and value added	0.2569**	0.0662	0.1718*	0.2539***	0.2511**	0.0644	0.1744*	0.2544***
	(0.1158)	(0.0699)	(0.1006)	(0.0611)	(0.1159)	(0.0699)	(0.1007)	(0.0609)
log employment	0.0515	0.3662***	0.3437***	0.3718***	0.0334	0.3784***	0.3677***	0.4076***
	(0.0904)	(0.0543)	(0.0725)	(0.0445)	(0.0918)	(0.0547)	(0.0731)	(0.0446)
ratio unit value to product average	-0.0972	0.0243	0.0226	0.0450	-0.0920	0.0274	0.0338	0.0485*
	(0.0937)	(0.0434)	(0.0507)	(0.0289)	(0.0937)	(0.0432)	(0.0502)	(0.0287)
firm market share by product	0.2255	-0.0805	1.3258***	0.8329***	0.2036	-0.0542	1.4040***	0.8972***
	(0.5219)	(0.3239)	(0.3618)	(0.2363)	(0.5205)	(0.3229)	(0.3599)	(0.2344)
Herfindahl concentration normalised of production firm year	-0.4773***	-0.4439***	-0.8233	-0.4140***	-0.4817*	-0.4540***	-0.8494***	-0.4390***
	(0.2487)	(0.1389)	(0.2000)	(0.1182)	(0.2487)	(0.1389)	(0.1996)	(0.1178)
distance CNAE 2 digits divisions	-0.0813	-0.0576	0.0177	0.0332	-0.0812***	-0.0579***	0.0182	0.0340***
	(0.0310)	(0.0153)	(0.0178)	(0.0095)	(0.0311)	(0.0153)	(0.0179)	(0.0095)
mean) dist	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000**
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
dummy for product innovation	0.1744	0.3085***	0.6970***	0.5587***				
	(0.1964)	(0.1109)	(0.1605)	(0.0923)				
dummy for process innovation					0.3133	0.1187	0.3636**	0.1322
					(0.1994)	(0.1159)	(0.1553)	(0.0935)
Other_Innovation dummy	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000*
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
significant marketing changes	0.3775**	0.2087**	0.0874	0.0297	0.3745**	0.2351**	0.1305	0.0763
	(0.1754)	(0.0995)	(0.1365)	(0.0830)	(0.1748)	(0.0992)	(0.1360)	(0.0825)
number of high skill technical staff	5.3290**	2.4210	3.5650*	3.5026**	6.1055***	3.5470	5.2506***	5.0946***

	(2.1563)	(2.5997)	(2.0689)	(1.5078)	(2.1969)	(2.5224)	(2.0208)	(1.4958)
group_dep1	-0.1257	-0.2462	0.1260	0.2437*	-0.1356	-0.2060	0.2003	0.3298**
	(0.3519)	(0.2295)	(0.1958)	(0.1369)	(0.3504)	(0.2295)	(0.1950)	(0.1364)
client_dep1	-0.3517	0.1783	0.0534	-0.2331	-0.3654	0.2116	0.0892	-0.1807
	(0.4906)	(0.2224)	(0.2717)	(0.1888)	(0.4880)	(0.2222)	(0.2716)	(0.1888)
high information from university	-0.1687	-0.0890	-0.1158	-0.0652	-0.2442	-0.0671	-0.0879	-0.0044
	(0.1845)	(0.1026)	(0.1419)	(0.0852)	(0.1927)	(0.1100)	(0.1439)	(0.0884)
foreign_cap1	-0.0161	0.0257	0.7125***	0.5084***	-0.0012	0.0174	0.6760***	0.4759***
	(0.2496)	(0.1587)	(0.1738)	(0.1119)	(0.2491)	(0.1587)	(0.1729)	(0.1115)
independent or group	0.1334	-0.2329*	-0.2008	-0.1201	0.1443	-0.2099	-0.1839	-0.0902
	(0.2112)	(0.1283)	(0.1603)	(0.0993)	(0.2114)	(0.1279)	(0.1600)	(0.0989)
Observations	8960				8960			
Log-likelihood	-5013.47				-5034.76			
Pseudo R2	0.2674				0.2643			

Exporters that do not diversify are the base category *** significant at 1% confidence level, ** significant at 5% confidence level and * significant at 10% confidence level. Year, region and sector dummy coefficient omitted from the table.

Table A4.6 Multinomial logit estimates relatedness (HS-2 differences) and technological content (OECD) – core production

	(1)				(2)			
	Related No upgrade	**Related Upgrade**	**Unrelated No upgrade**	**Unrelated Upgrade**	**Related No upgrade**	**Related Upgrade**	**Unrelated No upgrade**	**Unrelated Upgrade**
TFP using Levinsohn and Petrin and value added	0.1371	0.0536	0.1568**	0.1462***	0.2629	0.1377	0.1780**	0.1720***
	(0.1678)	(0.1067)	(0.0716)	(0.0464)	(0.3074)	(0.1426)	(0.0803)	(0.0500)
log employment	0.0196	0.3457***	0.2091***	0.3330***	0.1491	0.4427***	0.2567***	0.3903***
	(0.1194)	(0.0822)	(0.0539)	(0.0356)	(0.1979)	(0.1042)	(0.0608)	(0.0381)
ratio unit value to product average	-0.0934	-0.0442	0.0096	0.0474**	-0.5587*	-0.0720	0.0193	0.0524**
	(0.1013)	(0.0753)	(0.0383)	(0.0229)	(0.3033)	(0.0950)	(0.0449)	(0.0260)
firm market share by product	1.5649***	0.7016	0.5273*	0.2856	2.9823***	0.9933*	0.8457***	0.5180**
	(0.5813)	(0.4735)	(0.2817)	(0.1946)	(0.9591)	(0.5737)	(0.3139)	(0.2072)
Herfindahl concentration normalised of production firm year	-0.5866*	-0.2384	-0.5496***	-0.3480***	-1.3836**	-0.3668	-0.6665***	-0.4397***
	(0.3241)	(0.2139)	(0.1453)	(0.0930)	(0.5901)	(0.2696)	(0.1646)	(0.0990)
distance CNAE 2 digits divisions	0.0603*	0.0163	-0.0250*	0.0042	0.0815*	0.0182	-0.0235	0.0036
	(0.0308)	(0.0196)	(0.0145)	(0.0081)	(0.0446)	(0.0242)	(0.0164)	(0.0086)
mean) dist	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0001**	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0000)	(0.0000)	(0.0000)
dummy for product innovation	0.4353*	0.4810***	0.3405***	0.3705***				
	(0.2590)	(0.1719)	(0.1160)	(0.0736)				
dummy for process innovation					-0.1023	0.0065	0.3608***	0.1332*
					(0.4262)	(0.2170)	(0.1302)	(0.0802)
Other_Innovation dummy	0.0000	0.0000	0.0000**	0.0000*	0.0000	0.0000	0.0000*	0.0000*
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
significant marketing changes	-0.3560	0.2062	0.1822*	0.0485	-0.2614	0.3073*	0.2565**	0.1237*
	(0.2302)	(0.1497)	(0.1017)	(0.0662)	(0.4036)	(0.1860)	(0.1138)	(0.0702)
number of high skill technical staff	-1.7026	6.8552***	6.5987***	2.8298*	1.6033	6.0211***	5.9862***	4.4365***

	(3.2283)	(2.0637)	(1.7227)	(1.7148)	(5.5214)	(2.1307)	(1.7134)	(1.4854)
group_dep1	-0.1114	0.2447	0.1133	0.1542	0.2441	0.3315	0.1520	0.2269*
	(0.3914)	(0.3042)	(0.1606)	(0.1202)	(0.7017)	(0.3425)	(0.1785)	(0.1279)
client_dep1	0.8739**	-0.0901	-0.0408	-0.0441	0.8880	-0.0191	-0.0636	-0.0345
	(0.4234)	(0.3920)	(0.2237)	(0.1492)	(0.6892)	(0.4560)	(0.2520)	(0.1607)
high information from university	0.0539	-0.0568	-0.1749*	-0.0798	0.2393	0.0454	-0.1841	-0.0376
	(0.2275)	(0.1575)	(0.1055)	(0.0675)	(0.4082)	(0.2044)	(0.1223)	(0.0761)
foreign_cap1	-0.2734	0.1565	0.5128***	0.3046***	-0.4857	0.2253	0.4930***	0.3292***
	(0.3162)	(0.2448)	(0.1347)	(0.0948)	(0.5537)	(0.2822)	(0.1489)	(0.1000)
independent or group	-0.1694	-0.1557	-0.0863	-0.1598**	-0.9072	-0.1726	-0.0321	-0.1287
	(0.2663)	(0.1992)	(0.1206)	(0.0804)	(0.5761)	(0.2415)	(0.1346)	(0.0859)
Observations	8960				8960			
Log-likelihood	-5142.27				-4555.30			
Pseudo R2	0.1653				0.2606			

Exporters that do not diversify are the base category *** significant at 1% confidence level, ** significant at 5% confidence level and * significant at 10% confidence level. Year, region and sector dummy coefficient omitted from the table.

Table A4.7 Multinomial logit relatedness (input use) and sophistication (PRODY) exports

	(1)				(2)			
	Related No upgrade	**Related Upgrade**	**Unrelated No upgrade**	**Unrelated Upgrade**	**Related No upgrade**	**Related Upgrade**	**Unrelated No upgrade**	**Unrelated Upgrade**
TFP using Levinsohn and Petrin and value added	0.2139***	0.0601	0.1883**	0.0864	0.2154***	0.0645	0.1891**	0.0827
	(0.0548)	(0.0779)	(0.0747)	(0.0972)	(0.0547)	(0.0778)	(0.0747)	(0.0973)
log employment	0.3954***	0.2414***	0.3803***	0.2510***	0.4256***	0.2779***	0.4062***	0.2447***
	(0.0406)	(0.0610)	(0.0555)	(0.0762)	(0.0408)	(0.0615)	(0.0559)	(0.0770)
ratio unit value to product average	0.0385	-0.0093	0.0612*	0.0294	0.0391	-0.0086	0.0658*	0.0323
	(0.0286)	(0.0542)	(0.0344)	(0.0524)	(0.0287)	(0.0540)	(0.0341)	(0.0525)
firm market share by product	0.5737**	0.5905*	0.5440*	0.8478**	0.6220***	0.6469*	0.6181**	0.8501**
	(0.2245)	(0.3460)	(0.2999)	(0.4031)	(0.2234)	(0.3442)	(0.2978)	(0.4023)
Herfindahl concentration normalised of production firm year	-0.6311***	-0.3794**	-0.3804**	-0.0782	-0.6489***	-0.4036**	-0.4057***	-0.0800
	(0.1067)	(0.1618)	(0.1484)	(0.2057)	(0.1065)	(0.1618)	(0.1481)	(0.2057)
distance CNAE 2 digits divisions	0.0003	-0.0053	0.0450***	0.0307*	0.0009	-0.0045	0.0456***	0.0304*
	(0.0095)	(0.0151)	(0.0119)	(0.0169)	(0.0095)	(0.0151)	(0.0119)	(0.0169)
mean) dist	0.0000*	0.0000	0.0000	0.0000	0.0000**	0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
dummy for product innovation	0.3939***	0.4111***	0.5708***	0.1998				
	(0.0838)	(0.1286)	(0.1162)	(0.1611)				
dummy for process innovation					0.0063	-0.1043	0.2094*	0.2371
					(0.0856)	(0.1327)	(0.1170)	(0.1677)
Other_Innovation dummy	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
significant marketing changes	0.1522**	0.0950	0.2579**	0.1760	0.1930**	0.1490	0.2992***	0.1790
	(0.0757)	(0.1164)	(0.1025)	(0.1451)	(0.0753)	(0.1159)	(0.1021)	(0.1447)
number of high skill technical staff	2.6233	3.9368*	3.5247**	6.6779***	4.1034**	5.3677***	5.1454***	7.5850***

	(1.6108)	(2.0929)	(1.7397)	(1.7850)	(1.6005)	(2.0250)	(1.7279)	(1.8272)
group_dep1	0.2270*	0.1508	0.1614	0.6755**	0.2991**	0.2387	0.2349	0.6795**
	(0.1338)	(0.2402)	(0.1719)	(0.2678)	(0.1335)	(0.2402)	(0.1714)	(0.2663)
client_dep1	0.1142	0.0223	-0.0995	-0.6053	0.1655	0.0984	-0.0630	-0.5987
	(0.1672)	(0.2617)	(0.2277)	(0.4108)	(0.1674)	(0.2619)	(0.2279)	(0.4095)
high information from university	-0.1302*	-0.0242	-0.0679	-0.0100	-0.0512	0.1110	-0.0182	-0.0536
	(0.0777)	(0.1186)	(0.1064)	(0.1496)	(0.0814)	(0.1268)	(0.1101)	(0.1590)
foreign_cap1	0.5026***	-0.0053	0.5310***	-0.0050	0.4796***	-0.0295	0.5026***	0.0026
	(0.1048)	(0.1800)	(0.1395)	(0.2219)	(0.1046)	(0.1802)	(0.1391)	(0.2215)
independent or group	-0.0773	-0.2440	-0.2106*	0.0653	-0.0499	-0.2092	-0.1818	0.0770
	(0.0918)	(0.1494)	(0.1259)	(0.1774)	(0.0915)	(0.1489)	(0.1256)	(0.1773)
Observations	9077				9077			
Log-likelihood	-5984.27				-6002.64			
Pseudo R2	0.2233				0.2209			

Exporters that do not diversify are the base category *** significant at 1% confidence level, ** significant at 5% confidence level and * significant at 10% confidence level. Year, region and sector dummy coefficient omitted from the table.

Table A4.8 Multinomial logit relatedness (HS-2 difference) and sophistication (PRODY) exports

	(1)				(2)			
	Related No upgrade	**Related Upgrade**	**Unrelated No upgrade**	**Unrelated Upgrade**	**Related No upgrade**	**Related Upgrade**	**Unrelated No upgrade**	**Unrelated Upgrade**
TFP using Levinsohn and Petrin and value added	0.1257	0.0277	0.2021***	0.0689	0.1257	0.0448	0.2029***	0.0694
	(0.1505)	(0.2637)	(0.0491)	(0.0637)	(0.1505)	(0.2647)	(0.0490)	(0.0637)
log employment	0.2463**	0.2169	0.3934***	0.2370***	0.2672**	0.2298	0.4220***	0.2547***
	(0.1125)	(0.2199)	(0.0369)	(0.0499)	(0.1129)	(0.2202)	(0.0371)	(0.0504)
ratio unit value to product average	0.1022	-0.2834	0.0413	0.0166	0.1045*	-0.3003	0.0435*	0.0181
	(0.0631)	(0.3079)	(0.0257)	(0.0391)	(0.0630)	(0.3124)	(0.0257)	(0.0390)
firm market share by product	1.8758***	-0.0136	0.4587**	0.7065**	1.9169***	-0.0718	0.5129**	0.7395***
	(0.5463)	(1.2860)	(0.2062)	(0.2777)	(0.5438)	(1.3024)	(0.2048)	(0.2767)
Herfindahl concentration normalised of production firm year	-0.8711***	-0.1727	-0.5411***	-0.2689**	-0.8863***	-0.1834	-0.5611***	-0.2849**
	(0.3188)	(0.5643)	(0.0965)	(0.1332)	(0.3188)	(0.5631)	(0.0962)	(0.1332)
distance CNAE 2 digits divisions	0.0291	0.0018	0.0119	0.0083	0.0297	0.0021	0.0124	0.0088
	(0.0249)	(0.0538)	(0.0085)	(0.0119)	(0.0250)	(0.0533)	(0.0084)	(0.0119)
mean) dist	0.0000	-0.0001	0.0000*	0.0000	0.0000	-0.0001	0.0000**	0.0000
	(0.0000)	(0.0001)	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0000)	(0.0000)
dummy for product innovation	0.3969	-0.3037	0.4512***	0.3306***				
	(0.2453)	(0.4651)	(0.0756)	(0.1055)				
dummy for process innovation					0.1206	-0.4030	0.0799	0.0516
					(0.2472)	(0.4627)	(0.0778)	(0.1094)
Other_Innovation dummy	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
significant marketing changes	0.1815	0.2127	0.1870***	0.1160	0.2133	0.2117	0.2277***	0.1475
	(0.2167)	(0.4043)	(0.0685)	(0.0955)	(0.2160)	(0.4027)	(0.0682)	(0.0951)
number of high skill technical staff	-1.8215	3.7743	3.0346**	5.3866***	-0.0612	3.5634	4.5327***	6.4990***

	(6.8533)	(5.5013)	(1.3882)	(1.5161)	(6.5050)	(5.6642)	(1.4076)	(1.5508)
group_dep1	0.3689	0.4863	0.2259*	0.3832**	0.4266	0.4665	0.2956**	0.4395**
	(0.3906)	(0.6947)	(0.1234)	(0.1910)	(0.3897)	(0.6893)	(0.1232)	(0.1906)
client_dep1	0.2399	-33.7109	0.0372	-0.1217	0.2783	-30.8887	0.0849	-0.0825
	(0.4533)	(20500000)	(0.1544)	(0.2241)	(0.4532)	(5145221)	(0.1545)	(0.2242)
high information from university	-0.0698	-0.1164	-0.1159*	-0.0012	-0.0270	-0.0195	-0.0505	0.0514
	(0.2262)	(0.4212)	(0.0703)	(0.0973)	(0.2343)	(0.4523)	(0.0739)	(0.1042)
foreign_cap1	-0.0551	0.9621*	0.5405***	-0.0641	-0.0715	0.9264	0.5162***	-0.0778
	(0.3189)	(0.5599)	(0.0954)	(0.1487)	(0.3184)	(0.5629)	(0.0951)	(0.1487)
independent or group	-0.0457	0.0127	-0.1132	-0.1438	-0.0210	-0.0017	-0.0843	-0.1189
	(0.2600)	(0.5037)	(0.0840)	(0.1207)	(0.2597)	(0.5027)	(0.0837)	(0.1204)
Observations	9103				9103			
Log-likelihood	-5151.21				-5170.66			
Pseudo R2	0.2465				0.2436			

Exporters that do not diversify are the base category *** significant at 1% confidence level, ** significant at 5% confidence level and * significant at 10% confidence level. Year, region and sector dummy coefficient omitted from the table.

Table A4.9 Multinomial logit relatedness (input use) and sophistication (PRODY) core production activity

	(1)				(2)			
	Related No upgrade	**Related Upgrade**	**Unrelated No upgrade**	**Unrelated Upgrade**	**Related No upgrade**	**Related Upgrade**	**Unrelated No upgrade**	**Unrelated Upgrade**
TFP using Levinsohn and Petrin and value added	0.2981*	0.0919	0.2084	0.2392***	0.2917*	0.0900	0.2053	0.2404***
	(0.1707)	(0.0649)	(0.1320)	(0.0585)	(0.1712)	(0.0649)	(0.1320)	(0.0583)
log employment	0.1498	0.3127***	0.2637***	0.3807***	0.1221	0.3225***	0.2886***	0.4152***
	(0.1280)	(0.0505)	(0.0933)	(0.0428)	(0.1293)	(0.0509)	(0.0940)	(0.0429)
ratio unit value to product average	-0.3992**	0.0287	0.0593	0.0389	-0.3940**	0.0318	0.0669	0.0432
	(0.1858)	(0.0397)	(0.0511)	(0.0291)	(0.1858)	(0.0396)	(0.0504)	(0.0288)
firm market share by product	0.3155	-0.0154	1.4004***	0.8850***	0.2369	0.0092	1.4721***	0.9507***
	(0.7576)	(0.3000)	(0.4535)	(0.2273)	(0.7576)	(0.2990)	(0.4512)	(0.2253)
Herfindahl concentration normalised of production firm year	-0.5684	-0.4435***	-0.4818*	-0.5045***	-0.5651	-0.4541***	-0.5119**	-0.5298***
	(0.3578)	(0.1305)	(0.2498)	(0.1140)	(0.3575)	(0.1305)	(0.2495)	(0.1135)
distance CNAE 2 digits divisions	-0.0500	-0.0639***	-0.0026	0.0335***	-0.0500	-0.0643***	-0.0024	0.0343***
	(0.0397)	(0.0148)	(0.0247)	(0.0092)	(0.0397)	(0.0148)	(0.0248)	(0.0092)
mean) dist	0.0001	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000**
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
dummy for product innovation	0.0042	0.3049***	0.6055***	0.5864***				
	(0.2792)	(0.1041)	(0.2030)	(0.0889)				
dummy for process innovation					0.2931	0.1408	0.2954	0.1671*
					(0.2832)	(0.1085)	(0.1971)	(0.0901)
Other_Innovation dummy	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000*	0.0000*
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
significant marketing changes	0.2755	0.2400**	0.2853	0.0079	0.2590	0.2642***	0.3205*	0.0544
	(0.2497)	(0.0935)	(0.1736)	(0.0801)	(0.2490)	(0.0931)	(0.1733)	(0.0796)
number of high skill technical staff	6.3379*	3.8750**	4.0192*	3.3959**	6.6247*	4.9231***	5.6155**	5.0495***

	(3.5749)	(1.8548)	(2.3310)	(1.4576)	(3.8275)	(1.8625)	(2.2920)	(1.4503)
group_dep1	0.0812	-0.2636	0.3925	0.1852	0.0369	-0.2251	0.4592*	0.2710**
	(0.5387)	(0.2098)	(0.2470)	(0.1320)	(0.5362)	(0.2096)	(0.2463)	(0.1314)
client_dep1	-0.5189	0.1442	-0.0343	-0.1929	-0.5532	0.1730	0.0042	-0.1408
	(0.7536)	(0.2131)	(0.3430)	(0.1803)	(0.7489)	(0.2129)	(0.3430)	(0.1802)
high information from university	-0.0464	-0.1094	-0.2753	-0.0515	-0.1559	-0.0972	-0.2440	0.0034
	(0.2586)	(0.0967)	(0.1829)	(0.0821)	(0.2737)	(0.1032)	(0.1852)	(0.0851)
foreign_cap1	-0.3720	0.0633	0.4625**	0.5684***	-0.3329	0.0559	0.4318*	0.5354***
	(0.3884)	(0.1455)	(0.2258)	(0.1071)	(0.3863)	(0.1455)	(0.2249)	(0.1066)
independent or group	-0.1007	-0.1507	-0.1599	-0.1362	-0.1006	-0.1290	-0.1440	-0.1075
	(0.3202)	(0.1186)	(0.2037)	(0.0959)	(0.3208)	(0.1182)	(0.2034)	(0.0955)
Observations	8960				8960			
Log-likelihood	-4748.75				-4771.28			
Pseudo R2	0.2734				0.27			

Exporters that do not diversify are the base category *** significant at 1% confidence level, ** significant at 5% confidence level and * significant at 10% confidence level. Year, region and sector dummy coefficient omitted from the table.

Table A4.10 Multinomial logit relatedness (input use) and sophistication (PRODY) – vis-a vis core production activity

	(1)				(2)			
	Related No upgrade	**Related Upgrade**	**Unrelated No upgrade**	**Unrelated Upgrade**	**Related No upgrade**	**Related Upgrade**	**Unrelated No upgrade**	**Unrelated Upgrade**
TFP using Levinsohn and Petrin and value added	0.1361	0.1619	0.2339**	0.1665***	0.1366	0.1610	0.2314**	0.1667***
	(0.4483)	(0.1353)	(0.1090)	(0.0477)	(0.4527)	(0.1354)	(0.1092)	(0.0476)
log employment	0.1388	0.3448***	0.2375***	0.3551***	0.1865	0.4046***	0.2372***	0.3764***
	(0.3018)	(0.0970)	(0.0791)	(0.0363)	(0.3108)	(0.0975)	(0.0797)	(0.0365)
ratio unit value to product average	-1.7031**	-0.0703	0.0276	0.0434*	-1.6968**	-0.0715	0.0348	0.0480*
	(0.7039)	(0.0896)	(0.0512)	(0.0257)	(0.7049)	(0.0898)	(0.0510)	(0.0256)
firm market share by product	2.5365	1.2869**	0.9593**	0.4995**	2.6453	1.3478***	0.9854**	0.5482***
	(1.6676)	(0.5181)	(0.4035)	(0.1998)	(1.6742)	(0.5155)	(0.4019)	(0.1985)
Herfindahl concentration normalised of production firm year	-1.2001	-0.4482*	-0.4564**	-0.4732***	-1.3108	-0.4763*	-0.4701**	-0.4911***
	(0.8155)	(0.2558)	(0.2140)	(0.0952)	(0.8105)	(0.2567)	(0.2140)	(0.0949)
distance CNAE 2 digits divisions	0.0625	0.0270	-0.0273	0.0008	0.0539	0.0286	-0.0270	0.0012
	(0.0723)	(0.0222)	(0.0220)	(0.0084)	(0.0729)	(0.0222)	(0.0221)	(0.0084)
mean) dist	-0.0001	0.0000	0.0000	0.0000	-0.0001	-0.0001	0.0000	0.0000*
	(0.0001)	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0000)	(0.0000)	(0.0000)
dummy for product innovation	0.7892	0.5989***	0.3412**	0.4619***				
	(0.6337)	(0.2044)	(0.1696)	(0.0748)				
dummy for process innovation					0.1158	-0.0491	0.2840*	0.1709**
					(0.6321)	(0.2039)	(0.1689)	(0.0770)
Other_Innovation dummy	-0.0001	0.0000	0.0000*	0.0000*	-0.0001	0.0000	0.0000*	0.0000*
	(0.0001)	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0000)	(0.0000)	(0.0000)
significant marketing changes	0.6131	0.1097	0.2912*	0.0959	0.6645	0.1736	0.3064**	0.1327**
	(0.5467)	(0.1778)	(0.1496)	(0.0678)	(0.5461)	(0.1772)	(0.1493)	(0.0674)
number of high skill technical staff	-19.7959	4.4042**	5.4000**	3.3709**	-13.7071	5.7003***	6.5145***	4.7996***

	(29.3750)	(2.0437)	(2.0912)	(1.3854)	(26.4547)	(2.0335)	(2.1114)	(1.3914)
group_dep1	0.9601	0.1227	0.3606	0.1162	1.0814	0.2132	0.3848*	0.1783
	(0.8314)	(0.3326)	(0.2337)	(0.1229)	(0.8270)	(0.3327)	(0.2328)	(0.1224)
client_dep1	0.3183	0.1225	-0.1253	-0.0773	0.3911	0.1948	-0.1068	-0.0347
	(1.1373)	(0.4015)	(0.3230)	(0.1554)	(1.1377)	(0.4021)	(0.3225)	(0.1551)
high information from university	-0.3528	-0.0284	-0.1954	-0.0853	-0.2436	0.1241	-0.2103	-0.0526
	(0.5732)	(0.1869)	(0.1566)	(0.0693)	(0.5988)	(0.1934)	(0.1605)	(0.0729)
foreign_cap1	0.3867	0.0400	0.1502	0.4161***	0.3836	0.0273	0.1472	0.3979***
	(0.7570)	(0.2690)	(0.2005)	(0.0953)	(0.7537)	(0.2687)	(0.2001)	(0.0950)
independent or group	-22.5548***	-0.1930	-0.0645	-0.1422*	-22.5939***	-0.1626	-0.0502	-0.1144
	(4.3315)	(0.2285)	(0.1766)	(0.0827)	(4.3440)	(0.2281)	(0.1767)	(0.0823)
Observations	8960				8960			
Log-likelihood	-4265.88				-4284.81			
Pseudo R2	0.2713				0.2681			

Exporters that do not diversify are the base category *** significant at 1% confidence level, ** significant at 5% confidence level and * significant at 10% confidence level. Year, region and sector dummy coefficient omitted from the table