IDS WORKING PAPER
Volume **2012** No **390**

Firm Behaviour and the Introduction of New Exports: Evidence from Brazil

Xavier Cirera, Anabel Marin and Ricardo Markwald
March 2012

IDS Institute of Development Studies

Firm Behaviour and the Introduction of New Exports: Evidence from Brazil
Xavier Cirera, Anabel Marin and Ricardo Markwald
IDS Working Paper 390
First published by the Institute of Development Studies in March 2012

ISSN: 2040-0209 ISBN: 978-1-78118-048-8

A catalogue record for this publication is available from the British Library.

Available from:
Central Communications, Institute of Development Studies, Brighton BN1 9RE, UK
Tel: +44 (0) 1273 915637 Fax: +44 (0) 1273 621202
E-mail: bookshop@ids.ac.uk
Web: www.ids.ac.uk/ids/bookshop
IDS is a charitable company limited by guarantee and registered in England (No. 877338)

Firm Behaviour and the Introduction of New exports: Evidence from Brazil

Xavier Cirera,[1] Anabel Marin[2] and Ricardo Markwald[3]

Summary

This paper contributes to understanding the process of export diversification by analysing firm level determinants in Brazil during the period 2000–2009. The first objective of the paper is to establish the set of firm characteristics and processes that are more conductive to new exports; the second, to identify different pathways to diversification regarding relatedness and sophistication and, which firm level behaviours can be associated to the different paths. We answer these questions using a unique dataset that links data on exports, innovation and firms characteristics at the firm level. The paper contributes to the literature on export diversification and on preparation for exporting by identifying firm level behaviours that contribute to the process of diversification. In particular, the findings suggest that firms prepare for diversification by first gaining power in the domestic market and more importantly that they do so by adopting specific innovation and learning efforts.

Keywords: diversification; relatedness; sophistication; trade; innovation; Brazil

JEL: F14; L25

Xavier Cirera is Research Fellow at the Institute of Development Studies. He is a trade economist with particular interest in the impact of trade reform and regional integration, and extensive experience in Southern Africa. His recent research focuses on analysing the impact of different regional and preferential trade agreements, GSP and EPAs. He has also focused his research on how firms introduce new products for exporting, and on trade, poverty and income distribution. His other interests include the analysis of spatial price variations; the functioning of rural markets; the impact of income distribution on food demand and trade and climate change policies.

Anabel Marin is currently based at Centro de Estudios Para Transformación (CENIT), has also worked as a research fellow at the Science Policy Research Unit (SPRU) at the University of Sussex, where she also achieved her PhD in Science and Technology Policy. She has carried out extensive research on innovation including a project investigating 'innovation capability building, learning mechanisms and institutional frameworks in natural resource-related and process intensive industries in emerging economies'.

Ricardo Markwald has worked as a research expert with the Institute of Applied Economic Research (IPEA) since 1980, specialising in the fields of macro-economics, macroeconomic modeling, and the foreign sector of the Brazilian economy. He was appointed Research Coordinator at the Foreign Trade Studies Center Foundation (FUNCEX) in 1995, and is now the General Director of Fundação Centro de Estudos de Comércio Exterior (FUNCEX) in Brazil. Also an assistant lecturer at PUC-Rio de Janeiro, he has published several articles in Perspectivas da Economia Brasileira.

1 Institute of Development Studies, University of Sussex, UK

2 SPRU, University of Sussex; Conicet, Argentina.

3 FUNCEX, Brazil

Contents

Tables

Figures

Acknowledgements

Financial support from the Economic and Social Research Council (ESRC) is gratefully acknowledged. We would like to thanks Alan Winters, Azim Essaji, seminar participants at the University of Sussex, FUNCEX and the Workshop 'Firm, Trade and Development' at Aix-en-Provence, and 33 'Encontro Brasileiro de Econometria' from the Brazilian Econometric Society for comments on earlier versions of the paper.

Acronyms

FDI	Foreign Direct Investment
GDP	Gross Domestic Product
IBGE	Instituto Brasileiro de Geografia e Estatística
MNC	Multinational Corporation
MNE	Multinational Enterprise
PIA	Pesquisa Industrial Anual
PINTEC	Pesquisa de Inovação Tecnológica
PSM	Propensity Score Matching
R&D	Research and Development
SECEX	Secretaria Comercio Exterior
SIC	Standard Industrial Classification
SITC	Standard International Trade Classification

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.

Most previous studies have focused on macro level determinants to the process of diversification. Understanding diversification however, requires understanding the process and capabilities by which firms –the fundamental units of economic organisation – introduce new products for production and export. This paper attempts to explain export diversification by researching the process by which firms introduce new products for export in one particular case study, Brazil.

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). We 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.

More specifically, the paper has two main objectives: (i) to identify and explore empirically firm level determinants of the process of export diversification, and (ii) to identify different pathways followed by Brazilian firms to diversify, taking into account key issues to this process identified by the relevant literature such as survival, relatedness or sophistication and explore which firm level characteristics might affect the pathway to diversification.

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 process of diversification we conduct an extensive review of the existing economic, business, and innovation literature. These strands of literature have all something to say about the process of export diversification.

Nevertheless, they have rarely spoken to each other in the past. We make a contribution to the existing literature by linking these strands of research together. Based on this review, we identify five main types of testable firm level determinants of the process of diversification: (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.

A final contribution of the paper is the development of a unique dataset that links production, trade and innovation data at the firm level in Brazil, covering the period 2000 to 2008. This rich dataset allows us to characterise the process and path of firm diversification simultaneously linking this process with the different types of innovation, production and learning efforts carried out by firms, their main characteristics and their foreign exposure.

Overall we find that: (i) there is a very large number of new products introduced for exporting by new firms, most of which are not sustained beyond one year, (ii) diversification occurs in very closely related activities, where firms have some core competences, (iii) most diversification occurs in new products with lower level of sophistication than existing exports, (iv) all the five groups of firm level determinants identified by the literature review appear to be relevant explaining export diversification, and, however, (v) 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; (vi) none of these elements seem to matter for diversification towards more or less sophisticated products.

The paper is organised as follows. Section 1 surveys the literature analysing what the economic, innovation and business literature has to tell us about export diversification at the firm level. Section 2 describes the dataset and methodology used in the paper to explore the main predictions of these literatures. The next three sections focus on results. Section 3 characterises firm export diversification in Brazil. Section 4 analyses the main determinants of firm export diversification. 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.

1 Explaining firm export diversification. A survey of the literature

Firms' behaviour around the process of diversification have been analysed by different strands of economics, business and innovation literature. Most of the existing studies within these literatures have focused on understanding separately exports (mainly economics), diversification (mainly business) or the introduction of new products (mainly innovation) in general. In this section, we review these studies and the few that have focused specifically on the introduction of new products for exports with the purpose of identifying the main key aspects of the process of export diversification at the firm level and its determinants.

The literature on diversification, whether from economics, innovation or management, emphasises that firms diversify in order to obtain higher profits. The different strands of the literature, however, point in different directions when trying to explain diversification at the firm level. The economic literature emphasises the importance of productivity and preparation for exports, the innovation literature the importance of learning efforts and the business literature the importance of issues such as the international involvement of firms.

We organise the review of the contributions of these different strands of literature in two sections. In the first we review the insights about firm level determinants of the process of diversification, in the second the insights about the different pathways that the process of diversification can follow.

1.1 Factors contributing to export diversification at the firm level

Productivity, size and trade costs

In recent years, a large number of empirical studies have emerged based on Melitz's (2003) model of heterogeneous firms. The key element of this model is the issue of self-selection of firms for exporting. The main elements shaping the decision for exporting are productivity levels and trade costs, which determine what firm will access export markets and to which markets it will be able to export. The model is able to replicate one important stylised fact: firms involved in international trade tend to be larger and more productive (Bernard *et al.* 2007; Mayer and Ottaviano 2008). Nevertheless, its static version treats productivity as exogenous to the firm, and does not address how firms achieve the productivity thresholds that allow them to export.[4]

In recent years, the Melitz's framework has been modified in order to accommodate other important stylised facts. For example, Baldwin and Harrigan (2007) modify the model in order to accommodate quality differentiation and the fact that more productive firms also earn higher (rather than lower) export prices. Ruhl and Willis (2009) modify the model to incorporate entry costs that change with changes in relative prices, productivity and demand shocks. This allows the model to be able to replicate gradual growth in exporters' market share. Finally, Arkolakis and Muendler (2009) modify the model to incorporate multiproduct firms in order to show that higher quality exports are directed to more distant countries.

A more relevant modification of this framework for the question of this paper is provided by Costantini and Melitz (2007). The authors modified the Melitz (2003) model in order to include dynamic elements, where entry, exit and innovation are jointly determined depending on entry and trade costs. In addition, (Aw *et al*. 2009) also develop a dynamic industry model with firm heterogeneity where investment on innovation decisions depends on profitability and entry sunk costs. These models represent a substantial improvement in trying to make productivity dynamics endogenous.

Despite these new advancements to include dynamic issues within the Melitz framework, the focus of this literature remains mainly on entry and trade costs, rather than explaining dynamic productivity changes. Thus, the key prediction for new products is that diversification will occur mainly via reductions on trade costs, tariffs and entry costs.

Learning and other innovation efforts

The innovation literature has focused on explaining the links between firms' decisions about innovation and productivity and, more recently also between firms' decisions about productivity and exports. The general idea of these studies is that the innovative

[4] Several papers have analysed empirically the predictions of the model. A key finding of the literature is the fact that most export growth tends to occur at the intensive margin, on existing products. Breaking into new products and markets tends to be more difficult than increasing existing shares. A collection of papers have analysed export growth at the extensive margin and linked this type of growth to the predictions of the Melitz's (2003) model. The focus of these papers has been how tariffs and trade costs, including the business environment, affect growth at the extensive margin. For example, Amurgo-Pacheco and Pierola (2008), Dennis and Shepherd (2007) find that reducing trade costs via improving the business environment, reduction in transport costs or reducing tariffs increase extensive margin growth (diversification).

performance, as well as the performance in general of individual firms is strongly linked to the conscious efforts that firms make to accumulate technological capabilities via investments in learning and other types of knowledge investments. Increases in productivity as well as in innovation are not an automatic effect of investment decision in technology created outside the firm. Instead it responds to specific efforts of the firm to accumulate capabilities and to capture knowledge advancements that are produced outside the firm.

There are two types of studies that are of special interest in relation to this study, that have used innovation survey data to explore the channels linking (a) learning and innovation efforts and productivity growth, and (b) innovation (efforts and output) and new exports.

The first set of studies exploring the channels linking learning and innovation efforts and productivity growth, most of which are based on the pioneer work by Crespo *et al.* (1998), have typically used a model based on four equations. The first two equations model R&D, the decision to invest in it and its intensity against a group of firm level variables, such as size, market power, demand pull (e.g. links with clients) and technology push factors (e.g. advances in the knowledge base). The third equation, models measures of innovative output, such as product or process innovation, to R&D and firm level determinants. The fourth uses a production function framework to link innovative output to total factor productivity (see Hall and Mairesse (2006) for a summary of these studies). The findings of these studies in general support the idea that there is a link between R&D efforts and innovative output, and between the innovative output of firms and their productivity. It also appears clearly in all the studies that size and market power, as well as technology push and demand pull factors explain improvements in innovation and productivity. In addition, they have identified a set of additional variables which explain why some firms are more innovative and productive than others in emerging countries. In particular in these types of contexts other investments in knowledge, such as skills, and in capital goods, as well as R&D seem to be linked positively to innovation and productivity.

The second group of studies has focused more specifically on the links between innovation and exports. These studies have been less careful than the studies previously discussed in terms of how they deal with the several links between innovation and productivity variables. Nevertheless, they have started to provide some insights about how these two types of variables relate to exports (only a few of them focus specifically on new exports). Within this category of studies, Aw *et al.* (2009) for Taiwan and Cassiman and Martinez-Ros (2007) for Spain, for instance, have found that R&D investments increase the probability of exporting. Becker and Egger (2007) and Cassiman *et al.* (2010) look at the type of innovation that is more likely to encourage exports. Making a link with the economic literature they propose that firms that obtain product and process innovations are more likely to be more productive and self select than to export. Their results suggest that product innovation is more conducive to export than process innovation. In a similar paper using a sample of Belgian firms and , controlling for the endogeneity of innovation,[5] Van Beveren and Vandenbussche (2009) find that it is the mixture of both product and process innovation which increases the propensity to export. The authors also stress the importance of anticipation, which implies the need for controls for past export activity and the fact that innovation efforts occur before exporting. In another related paper, Damijan *et al.* (2010) look at the relationship between innovation and exporting in Slovenia. Using also PSM methods to correct for endogeneity they find no impact of product or process innovation on exports. However, the authors do find that exporting tends to increase innovation efforts.

[5] One problem when looking at the impact of innovation on exports is the endogeneity of innovation expenditure, since exporters tend to invest more on innovation. Most of these studies therefore use Propensity Score Matching (PSM) techniques to control for endogeneity find that product innovation is the key factor for exporting.

As summarised in the previous section, the trade literature has clearly found a strong link between productivity and exports, however, it has not engaged with the factors that explain why some firms are more productive than others. The innovation literature, has engaged with this question, and identified innovative efforts and outputs as key in the explanation of productivity gaps between firms. Also, it has started to explore in some cases how these two types of variables relate to exports. We build on the last category of studies for our empirical approach here, but include some additional variables that have been found to explain both productivity and exports in firms in developing countries. In the next section, we discuss one of them, which has received a lot of attention, foreign direct investment and the international involvement of firms.

FDI and global engagement

Studies analysing the effect of FDI have proliferated following the large increase in investment projects experienced in developing countries. These studies have not explored specifically the association between FDI and diversification. However, we can derive some relevant insights from this literature. First, is the association between FDI and productivity growth of domestic firms, which is explained by positive technological spillovers emerging when foreign firms set up a plant in host country. This is due to the fact that FDI firms tend to have superior technology than domestic firms. In addition, more generally, firms with global engagement that engage in international trade or multinational enterprises (MNEs) are more likely to innovate because they have better access to technology diffusion (Lederman 2009).

A second relevant insight is the association between FDI and exports. FDI is expected to contribute to exports in two ways. First, because subsidiaries tend in general to export more than domestic firms due to their international involvement (i.e. intensive links with foreign firms both corporative and non corporative) (Marin and Giuliani 2011). Second, via FDI, export spillovers, which are expected to arise because MNCs have better international networks and tend to be much more active in trade than domestic firms, may offer demonstration effects or can help reducing sunk costs of exporting.

Regarding the first type of studies, the literature in general has failed to provide conclusive evidence regarding a positive association between FDI participation (for instance by industry) and productivity growth of domestic firms, particularly when the host country is emerging (see Javorcik (2004) for a discussion of the empirical literature and Crespo and Fontoura (2007) for a survey). However, when some additional variables are taken into account, such as the absorptive capabilities of domestic firms (Konings 2001; Girma 2005) and the type of activity that MNC's subsidiaries carry out in the host country (Marin and Bell 2006; Marin and Costa 2010; Marin and Sashidaran 2010), significant positive effects emerge. In other words, in general FDI do not necessarily have a significant effect on the productivity growth of domestic firms, but positive effects may appear for the more able domestic firms and when MNC's subsidiaries invest heavily in innovation in the host economy. Therefore, in many cases MNCs might benefit domestic firms regarding productivity and other dimensions.

In relation to the effects of FDI on exports, one way of overcoming the uncertainties related to firm export in terms of knowledge, information or capabilities is via spillovers from foreign investors or by participating in foreign value chains. The evidence again shows that these effects might emerge, but they are not automatic.

Swenson (2008), studied the relationship between MNEs and exports in China and found that MNEs enhanced export capabilities of domestic firms mainly via information spillovers. Greenaway *et al.* (2004) found for a panel of UK firms that export propensity was larger in sectors with larger MNE presence. They interpreted this result as evidence of positive spillovers. Bekes *et al.* (2009) analysed the impact of MNEs in Hungary. The authors found

evidence of positive spillovers on more productive firms but not on exporters. Aitken *et al.* (1997) look at the impact of MNE's spillovers using data from Mexico. They find positive spillovers from MNEs on the propensity to export, although the export intensity of MNEs does not affect the propensity to export by domestic firms.

Egan and Mody (1992) studied buyer-seller relationships for the bicycle and garment imports to the US. They found that these relationships help exporters in developing countries to lower entry costs and act as instruments for information, technology diffusion and access to industrial networks. Hobday and Rush (2007) analysed the role of FDI in building export capabilities of subsidiaries in Thailand. They found that some subsidiaries upgraded capabilities while others remained as assembling plants. A key element in determining the outcome seemed to have been the degree of centralisation of technology decisions within MNC. More decentralised networks encouraged subsidiaries to upgrade, while networks where technology decisions and processes were tightly controlled within the parent headquarter, remained as assembly plants. Thus, while in some cases MNEs' share can be an important vehicle for improving subsidiaries export performance, in others they can also become an important constraint in highly centralised organisational structures.

While the studies above suggest that there is evidence of MNEs' spillover effects within industries on export propensity, there is little evidence on the specific role of MNEs, FDI or value chain relationship on firm export diversification. Little is known about the mechanisms through which MNEs, FDI or value chain relationships may facilitate domestic firms' export diversification. In addition, different degrees of value chain governance can play different roles in fostering or constraining firm diversification.

1.2 How the diversification path looks like?

It is a sequential process

The economic literature has characterised the process of export diversification as sequential, one in which firms start introducing new varieties locally, and then export them once they have been proved (prepared) in the local context. Iacovone and Javorcik (2010) using a firm level dataset, document firm market preparation for exporting in Mexico. In their study, firms first introduce varieties domestically and, there is evidence of increases in quality proxied by increases in prices preceding exports. Once domestic varieties have matured, they are then exported. Given the asymmetry and uncertainty of information, multi-product firms start exporting a small number of varieties in small volumes.

López (2009) focuses on a different type of preparation for exporting – firm investments. Analysing the relationship between exporting and productivity, the author shows using Chilean firm-level data significant increases in investment and productivity prior to firms start exporting. This supports the view of self-selection of exporters, in this case using investments to increase productivity.

Albornoz *et al*.(2010) link the preparation to expected profitability. In their model, while there are high sunk costs for exporting, profitability can only be determined once the firm has started to export. This implies a sequential process where firms first decide whether to export based on expected profitability, and then adjust quantities, prices and markets once the real profitability can be estimated. This gradual and sequential export expansion is also affected by distance and trade costs, since expansion to other markets depends on similarity and distance.

The literature, therefore, indicates that diversification may be carried out sequentially, via some domestic preparation efforts via quality and investments first, and increasing the likelihood of product diversification once the firm has sequentially started exporting.

New exports are short-lived

Due to uncertainty regarding trade relationships, most trade flows tend to be small and short-lived. Besedes and Prusa (2006a, 2006b), using trade data for the US at the product level, document the fact that most trade transactions are small and do not survive more than three years. This raises the question of what makes new exports sustainable in time. Some studies have explored this question and found that survival rates are larger for differentiated products since the trade relationships required to sustain this type of trade need to be solid. Martincus and Carballo (2009) find a similar pattern for Peru and suggest that export market diversification is more important than product diversification in increasing survival. Volker (2009) also finds low survival rates for German imports, and the fact that survival depends on how large and close exporters are, the demand elasticity of substitution, and exporters market power. In general, the literature suggests that we should expect very short duration in most diversification cases.

Export 'discoveries' are rare

Following the seminal work of Hausmann and Rodrik (2003) the attention on export diversification at the country level has focused on understanding the processes through which a firm starts exporting a product new to the country. The original model focused on the role of market failures constraining export 'discoveries' when initial entrepreneurs are unable to capture all the positive externalities generated with the product discovery due to rapid imitation by other firms. A large number of empirical work and case studies has focused on describing these discovery episodes.[6]

Freund and Pierola (2009) document new export 'discoveries' for Peru, and compare them with new exports to the firm. The authors find that new exports to the firm have less chance of survival than discoveries, and that only large experienced exporters engage in discoveries, since the latter require larger sunk costs. Klinger and Lederman (2004) find that 'discoveries' are not restricted to 'dynamic' industries but also to traditional sectors such as agriculture. More importantly, the authors find that the determinants of entrepreneurship are not correlated with the frequency of discovery, which implies the existence of market failures.

Klinger (2007) focuses on the different types of uncertainties related to export discoveries. The author analyses eight case studies of different sectors in different countries and suggests two types of uncertainty: uncertainty related to productivity, costs and quality, and; related to demand, market segment characteristics and price. The author finds that when uncertainty is high, two main alternative strategies emerge: discovery in similar products or the use of FDI investors to obtain technology and knowledge.

While the discovery literature focuses on a narrow type of firm diversification, new export products for the country, it stresses relevant elements for firm diversification in general. For example, one may characterise both processes of firm diversification by a similar process where only the magnitude of sunk costs for exporting is different, much larger for discoveries and the capacity for imitating also differs, much lower for discoveries. If this is the case, then the role of productivity and demand uncertainties and the strategy to overcome these

[6] See for example the collection of country case studies commissioned by the IADB for the project 'The Emergence of New Successful Export Activities in Latin America'.

uncertainties, product relatedness and FDI investment, are extremely relevant for firm diversification in general. The size of the export flow increases with distance.

One stylised fact of the trade literature is the Alchian-Allen hypothesis (Alchian and Allen 1964), which establishes that considering per-unit trade costs reduces the relative price of higher quality goods. Therefore, it is more profitable to ship goods with higher quality and higher price to more distant destinations with larger per unit transport costs. As a result, we should expect that the amount exported when introducing new exports depends on the distance to the destination market.

It normally happens in related products

The business literature, in particular the resource-based approach, stresses the importance of related capabilities and resources explaining which new products will be produced and exported in multi product firms. While firms will target demand dynamic sectors for diversification, firms will diversify only to those sectors which are similar and can be handled by existing resources and capabilities (Lien and Klein 2010). This raises the question of relatedness or similarity between new and existing products.

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. They show 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, which are often specialised in the lower end of the value added chain regarding exports. The first question is how they can progress from diversifying into related new exports on to 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.

Multiproduct decisions and product concentration

Traditional models in trade theory have considered firms as producers of one product. The reality, however, is that firms are multi-product and export multiple products to multiple destinations (Bernard *et al.* 2010). This implies that firms look at the decision to produce and export a new product in conjunction with the optimal production and export mix. Moreover, this mix is not random, but is related to firm’s existing capabilities.

Teece *et al.* (1994) emphasise the role of enterprise learning and evolutionary paths of the firm, which together with the extent of competition in each particular sector will dictate a

firms' decision to diversify to more or less related activities. Accordingly, learning and market conditions will dictate the degree of concentration of the production base, which will impact further diversification.

It is important, therefore, when explaining diversification and its path to consider firms' strategies in terms of revenue and product concentration across the production and export mix, and the degree of competition in these sectors. These are likely to impact future decisions to diversify and how related the new products are likely to be.

1.3 A summary of the main determinants of firm export diversification

In section 5 and 6 we test empirically, how these specific issues identified in the review of the literature affect both diversification and its path. In order to implement the empirical section, we summarise these insights of the literature explaining diversification in five main groups that can then be mapped to information available on our existing dataset. These are shown in Table 1.1.

We divide the determinants of export diversification at the firm level into two types: firm level and meso level determinants. Meso level determinants are mainly of two categories: industry level and regional level. These relate to demand and sector factors, and regional export dynamics and policies.

The main focus of this paper is, however, the micro level determinants. These 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 associations 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, made efforts to upgrade their products and processes, have highly skilled personnel and have invested in marketing their products, are also more likely to succeed in gaining the necessary capabilities to diversify exports. Similarly they are more likely to be able to reach further unrelated and sophisticated products.

Table 1.1 Main determinants of export diversification - a summary

Variables	Type	Expected Direction
	Related to:	
Size	**Structural characteristics of the firm**	(+) The hypothesis is that the larger the firm, the greater the likelihood of introducing new exports as Melitz and the innovation literature suggest.
Ownership		(+) In line with the global business literature the hypothesis is that firms with foreign firm share are more likely to introduce new exports, due to their experience in global business.
Global group Integration		(+) In line with the global business literature the hypothesis is that firms engaged in strong links with foreign firms within their group are more likely to introduce new exports, due to the marketing and technological knowledge and support that can access through these links.
Global value chain integration		(?) In line with the global business literature the hypothesis is that firms engaged in strong links with foreign clients are more likely to introduce new exports, due to the marketing and technological knowledge and support that can access through these links. However, at the same time excessive dependency may constrain business strategies to move to other activities.
Quality of firm's products	**Position of the firm in the domestic market**	(+) In line with the literature on preparation for exports we expect that firms with higher quality might be better placed to diversify to new products for exports.
Position in the domestic market		(+) In line with the literature on preparation for exports we expect that the higher the firm's share in the domestic market, the higher the probability that the firm introduces a new product.
Concentration in the value of production	**Characteristics of the product basket of the firm**	(-) Firms with more value concentrated in few products may opt for firm strategies of very little diversification, and focus instead on expanding exports of existing products. As a result, other things constant, we should expect lower diversification for very concentrated firms.
Diversification in production		(+) The variable reflects the evolution of the firm regarding what is able to produce, and we should expect that firms with more diversified sector capabilities should be more likely to diversify at the extensive margin.
Innovative output – Product		(+) In line with the innovation and trade literature we expect that firms that have been successful in introducing some degree of product innovation are more likely to introduce a new product for export, due to the similar capacities that are needed and the chances that this new product is also an innovation for foreign markets.

Innovative output–Process	**Characteristics of the process of production**	(-) In line with the innovation and trade literature we expect that firms that have introduced process innovation are more likely to introduce a new product for export, due to the improvements in the quality and cost of the products that processes innovation allow. However, if these innovations are mainly focused in existing products it may reinforce to focus on existing exports rather than new.
Total Factor Productivity		(+) In line with the heterogeneous firms models we expect that the more productive firms are more capable to introduce new products for export.
Geographical Distance	**Learning efforts of the firm**	(+) In line with the literature on preparation for exports we expect that firms that are able to export to more distant markets are more capable in general to satisfy more demanding markets and having better information and trade networks, and, therefore, more capable to introduce new products for export.
R&D Innovation efforts		(+) In line with the innovation literature the hypothesis is that firm's engagement in R&D will allow acquiring capabilities and increase its chances of introducing new products in general and therefore for exports.
Other innovation efforts different to R&D		(+) The hypothesis is again that firms that carry out other innovation efforts are more likely to introduce new products and for export. These investments are often important for firms in developing countries, which due to sector specialisation (typically in NR linked sectors) innovate typically via other innovation efforts different to R&D.
Efforts in Marketing		(+) In line with the innovation and global business literature we expect that firms that carry out intensive efforts in marketing are more likely to success in the introduction of new products for exports.
Firm's skills		(+) In line with the innovation literature the hypothesis is that firms with a labour force highly skilled are more likely to introduce new products for export.

Source: Author's own elaboration

The next sections analyse empirically firm export diversification in Brazil.

2 Data and methodology

2.1 Data sources

In order to analyse firm export diversification we create a unique dataset that links production, trade and innovation data for Brazilian firms. 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 of 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 the year 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 for the period 2000–2009. 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 for 2000, 2003 and 2005. 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.

2.2 Methodology

2.2.1 Measuring export diversification

The key parameter in this paper is to 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. This implies that in our sample we cannot consider a new export as a product that was exported in 2000, since we do not know whether it was exported in 1999.

Second, as suggested above most export flows tend to be short lived (Besedes and Prusa 2006; Martincus and Carballo 2009). The main implication of low survival rates for new exports is that any definition of 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 the issue of firm export diversification and sustainability, we compute an index variable with value 1 for cases of export diversification and 0 otherwise, under three alternative measures that consider different degrees of rigidity regarding the survival of the export flow. The three types of classification are summarised in Table 3.1. Classification 1 is clearly the most rigid methodology since it requires continuously exporting the product once it is introduced, until the end of the period in 2008. All new products identified in Classification 1 are part of the other two methodologies. Classification 2 allows for some degree of intermittence during the period; while the last methodology is the most flexible since it only

requires three years of being exported. The different classification methodologies are applied to the trade dataset ignoring the destination market dimension.[7]

Table 2.1 Methodologies to identify new products

Classification	Description
Classification 1	New product not exported in 2000, and once introduced is exported continuously until 2008; if introduced in 2007, also exported or in 2008 and 2009
Classification 2	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)
Classification 3	New product not exported before 2002 and exported at least three years after

Source: Author's own elaboration

Throughout the empirical part, we conduct sensitivity analysis and analyse the different determinants and diversification paths for the different classification methods. One important element to consider is the fact that firms may introduce more than one new product for exporting in the same year. As a result, for each of the methodologies we use two types of variables: the dichotomous variable with value 1 if the firm introduced one or more new products, and a variable that counts the number of new products for each firm and year.

In order to further explore firm dynamics around diversification, we also differentiate between firms that diversify being existing exporters and firms that export for the first time, new exporters.

Finally, we identify and differentiate those new exports that are a discovery for the country from new products exported for the firm.

2.2.2 Determinants of diversification

The first objective of the paper is to explain diversification using the linked dataset. One important challenge when looking at trading, production and innovation activities is the potential simultaneity or endogeneity of investment and innovation interventions with exporting activities. Firms, as suggested in the literature review, may invest more on innovation activities as a result of trading activities. As a result, in order to minimise simultaneity problems we use lagged variables in t-1 from PIA and PINTEC to explain new exports in t. Since PINTEC has only three years available, we effectively use production, firm and innovation data for each firm *i* (vector X_{kt-1}) in 2000, 2003 and 2005 to explain the probability that firm *i* introduces a new exported product in 2001, 2004 and 2006 (Y_{it}).

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

[7] One could define variety as firm, product and destination. However, survival rates at such level of disaggregation are even lower. Moreover, the main interest of the paper is product diversification, rather than export market diversification.

To explain firm diversification, we proxy the prediction of the literature in Table 2.1 with the proxies summarised in Table 3.2. We use sector dummies *S* at CNAE (Brazilian Industrial Classification) two digits in order to control for sector specific elements such as trade costs, profitability and changes in foreign demand. We also use year *T* and regional dummies *R* to control for year effects and the large correlation between certain regions and exports due to specific policies and firms clustering. These dummies control the main meso factors.

The main proxies for the main micro determinants identified in Table 1.1 are shown in Table 2.2 below.

Table 2.2 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 per cent 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: Authors' own elaboration

2.2.3 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. These types 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 it 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 five measures:

1 **Correlation based on the input use of the input-output matrix in 2005**. We calculate the correlation in terms of 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 in terms of input use 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 (same input use) and, therefore, is highly related to existing exports.

2 **Correlation based on the input use of the input-output Leontieff matrix in 2005**. Same as above, but using the Leontieff input requirements matrix.

3 **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 are exported by countries in the world.[8] 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.

4 **Minimum difference** between the existing and new product at the **HS-4 sector**. All product codes at HS-8 are mapped to their HS-4 sectors, and the different between new and existing product pairs is computed. A difference of zero implies that the new product is in the same HS-4 sector as at least one of the existing products.

5 **Minimum difference** between the existing and new product at the **HS-2 chapter**. As above but looking at the HS-2 chapter rather than the HS-4 sector.

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

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

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 does not diversify; value one if it is an exporter who diversifies to a totally related activity, and; value two if it is an exporter that diversifies to a totally unrelated activity. Table 2.3 shows each measure when diversification is considered related or unrelated.

Table 2.3 Definitions of relatedness in diversification

Measure	Dimension	Definition
Input use	Relatedness in terms of sector input use according the Leontieff 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 CNAE3 sector	Related if difference=0 Unrelated if difference≠0

Source: Authors' 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 try to minimise the risk of endogeneity problems in the decision to diversify. As explanatory variables X_{it-1}, we use the set of variables identified in Table 3.2. 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.

$$D_{it} = \alpha_0 + \sum_k \beta_k X_{kit-1} + \sum_{i=1}^{t} T_{t-1} + \sum_{j=1}^{j} S_j + \sum_{n=1}^{N} R_n + u_{it} \qquad (2)$$

One problem with the formulation of equation (2) 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 3. 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.

2.2.4 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 on 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 (3) 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.

$$PRODY_{kt} = \sum_i \frac{x_{ikt} / X_{kt}}{\sum_i x_{ikt} / X_{kt}} GDPcap_{it} \qquad (3)$$

  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) non 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 categories, 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 as 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 (4) 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.

$$T_{it} = \alpha_0 + \sum_k \beta_k X_{kit-1} + \sum_{i=1}^{t} T_{t-1} + \sum_{j=1}^{j} S_j + \sum_{n=1}^{N} R_n + u_{it} \quad (4)$$

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 Stylised facts on firm export diversification in Brazil

This section characterises the process of firm export diversification in Brazil during the period 2000 to 2009 by focusing on uncovering some of the predictions of the literature review in section 1. Specifically, we focus on describing the export diversification process followed by most firms in Brazil and their path in relation to relatedness and sophistication.

3.1 Diversification

High export diversification activity but low survival rates

As suggested by the survival literature (Besedes and Prusa 2006; Martincus and Carballo 2009) most export flows tend to be short lived. Table 4.1 shows the number of years that each firm's product is exported during the period 2000–2009 for all sectors of the economy. To avoid bias in the results due to changes in classification at highly disaggregated levels, we calculate the number of years exported during the period at three different levels of aggregation in the MERCOSUR nomenclature: HS-8, HS-6 and HS-4 digits.

Only between two and three per cent of flows are observed the entire period. More importantly, only 20 per cent of the flows are exported continuously until the end of the period once they are introduced; and a significant share of these flows correspond to those flows only observed in the last two or three years of the sample period. This suggests very short duration and intermittence of export flows.

Table 3.1 Duration of exports – product by firm– all sectors

	HS-8		HS-6		HS-4	
Number of years exported	Export Flows	% share	Export Flows	% share	Export Flows	% share
1 year	360,300	56.96	314,359	56.14	193,182	53.42
2 year	105,877	16.74	93,723	16.74	61,473	17
3 year	55,351	8.75	48,428	8.65	31,326	8.66
4 year	30,327	4.79	27,470	4.91	19,318	5.34
5 year	21,181	3.35	19,253	3.44	13,499	3.73
6 year	15,968	2.52	14,615	2.61	10,365	2.87
7 year	12,371	1.96	11,556	2.06	8,320	2.3

8 year	9,808	1.55	9,061	1.62	6,452	1.78
9 year	6,959	1.1	6,801	1.21	5,435	1.5
10 year (entire sample period)	14,443	2.28	14,734	2.63	12,262	3.39
Total	632,585	100	560,000		361,632	
Once introduced is exported until the end of the period	122,769	19.41	109,052	19.47	73,454	20.31

Source: Authors' own elaboration from SECEX

Focusing more narrowly on the manufacturing sector also shows very low export survival rates. We repeat the decomposition above only for those firms that may be classified as manufacturing sector according to their main economic activity. The results are very similar. Flows that last the entire sample period range between 3.5 and 5 per cent, depending on the level of aggregation. On the other hand, only between 20 and 21 per cent of the flows are exported continuously after they are introduced.

Higher survival rates are correlated with larger export flows

In order to test the correlation between the duration of the flow and its size, we regress the normalised value of each export flow, on a set of year dummies, product fixed effects and the number of years that the flow was exported in the period 2000-2009. The coefficient on the number of years exported is 0.087 and statistically significant at 99 per cent confidence level. This indicates that export flows that last longer tend to have larger fob values, as suggested by the sequential exporting literature (Albornoz *et al.* 2010).

In addition, we look at the differences in export survival according to destination market. We differentiate between two regions, MERCOSUR and neighbouring countries region (MERCOSUR+), and the rest of the world. Table 3.2 shows the share of export flows by number of years exported. Duration rates are very similar for the two destinations. The main difference is related to the largest prevalence of export flows lasting one year for exports to the rest of the world. This may be explained by the fact that the two most important markets outside the region are the US and the EU. These markets are likely to be more demanding in terms of quality and competitiveness, which makes export failure more likely. However, this difference does not appear to be very significant, reflecting that once the sunk costs and standards are met sustainability may be similar.

Table 3.2 Number of years exported by destination

Number of years exported	Export flows to MERCOSUR+	Export flows to Rest of the World
1	47.84%	51.54%
2	17.29%	16.43%
3	10.21%	9.71%
4	5.73%	5.64%
5	4.44%	4.1%
6	3.45%	3.27%
7	2.93%	2.61%
8	2.34%	2.13%
9	1.75%	1.47%
10	4%	3.09%

Source: Authors' own elaboration from SECEX

Larger export flows are going to more distant markets

As seen in section 1, the Alchian-Allen hypothesis (Alchian and Allen 1964) establishes that per-unit trade costs reduces the relative price of higher quality goods; making more profitable to ship goods with higher quality and higher price to more distant destinations.

While determining the quality level of export flows is a challenging exercise without any information on product attributes, we examine the relationship between the size of the flows and the distance to the destination market. Concretely, 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 per cent 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.

Most exporters concentrate on few products, but there are some very large exporters

Table 3.3 shows basic statistics on the number of products exported by firm and year for the period 2000-2009 by sector of economic activity. The table shows large differences across sectors in terms of average number of products exported by firm. This is explained by a small number of very large exporters, especially in the construction sector.[9] While the average number of products exported by firm in a given year oscillates between two in agriculture and 60 in construction, the median firm exports one or two products. In the industrial sector, the average number of exported products is six, but this is again driven by some large exporters, since the median firm exports two products in a given year.

Table 3.3 Number of exported products by firm – Sector decomposition

Sector	mean	median	99th percentile	Standard deviation	N	max
Agriculture	2.10	1	13	3.27	3,859	107
Trading	9.92	2	143	30.75	50,621	1014
Construction	60.39	2	2105	289.05	623	2578
Industry	6.71	2	71	19.38	124,327	604
Services	7.32	1	106	29.20	6,052	747
Other	5.75	1	63	19.36	1,221	461

Source: Authors' own elaboration from SECEX

Firm export diversification occurs for a relative low number of exported flows, but a significant number of exporters engage in export diversification

The different methodologies described in section 2 for measuring firm diversification are implemented on export flows of firms classified in the manufacturing sector according to their core activity. We also differentiate among two different types of firms:

- New products – these are new products introduced by existing exporters
- New exporters – these are new products introduced by firms that were not exporting previously

[9] This sector is largely dominated by a small number of MNEs with large construction contracts in Latin America and Africa, which export a very large number of products in order to supply their operations abroad.

Table 3.4 shows the number of export flows for each classification method and the percentage of total flows where diversification occurs. Clearly, Classification 3 is the most flexible method since it only requires three years of duration after a flow is introduced to be considered diversification. According to this method, the maximum amount of new flows in a given year, for both existing and new exporters, is around 15.4 per cent in 2003. The share of new flows for this classification is decreasing in time, since the closest to the end of the sample period the less likely it is to have survived the three years. On the other hand, Classification 1, which requires sustainability throughout the period, is the most stringent method with a maximum of 4.6 per cent new flows (for both new and existing exporter) in 2003. The peak in 2007 under Classification 1 and 2 is due to the fact that we only require products to be exported in 2007, 2008 and 2009 to qualify as new export.

Table 3.4 Number of new export flows (firm and product) according to different classifications

	New product existing exporter						New product and new exporter					
Year	Class 1	% of flows	Class 2	% of flows	Class 3	% of flows	Class 1	% of flows	Class 2	% of flows	Class 3	% of flows
2001	1569	2.43%	5435	8.43%	n/a	n/a	546	0.85%	1441	2.23%	n/a	n/a
2002	2247	3.32%	5360	7.92%	n/a	n/a	571	0.84%	1249	1.85%	n/a	n/a
2003	2838	3.53%	5309	6.60%	9562	11.88%	882	1.10%	1522	1.89%	2804	3.48%
2004	3228	3.54%	4483	4.92%	8896	9.77%	816	0.90%	998	1.10%	1965	2.16%
2005	3492	3.65%	3894	4.07%	7189	7.51%	644	0.67%	678	0.71%	1126	1.18%
2006	4109	4.36%	4109	4.36%	5244	5.56%	1083	1.15%	1083	1.15%	1203	1.28%
2007	6173	6.35%	6173	6.35%	6173	6.35%	861	0.89%	861	0.89%	0	0.00%

Source: Authors' own elaboration from SECEX

Table 3.4 also shows important differences between new and existing exporters, and as it should be expected, the contribution of new exporters to diversification is between three or four times lower than the contribution of existing exporters. Growth at the extensive margin, therefore, is mainly carried out by existing firms rather than new firms.

Table 3.5 shows a similar decomposition but focusing on the number of firms and average new products introduced. Comparing the different classification methodologies suggest similar results to the previous table at the export flow level. The main striking difference is that there is a significant share of firms diversifying in a given year. In some years, 25 per cent of firms introduce at least one new product for exporting. This implies that *there is* substantial firm level export diversification activity when one looks at the firm level rather than at the product level. This again is partly the result of concentration of exports on a smaller subset of multi-export firms as compared to a more atomised structure when producing for domestic markets.

Although most firms introduce one new product in a given year, for most years and classifications used, the average number of products exported is around two. When we restrict our sample to firms that have been surveyed by the manufacturing survey (PIA) we obtain that 85 per cent of the observations, firm and year do not engage in diversification activities, and the maximum number of products introduced by a given firm in a year is 44; and 17 new products if the firm is a new exporter.

Table 3.5 Firm that introduce new products for exporting in the manufacturing sector

year		New products Class 1	Class 2	Class 3	New exporters Class 1	Class 2	Class 3	Discovery Class 1	Class 2	Class 3	Number Exporters
2001	Firms	816	1931	0	329	827	0	3	10	0	11,408
	Average new products	1.92	2.81	.	1.66	1.74	.	1	1	.	
	Maximum number	26	74	.	18	25	.	1	1	.	
2002	Firms	1074	2053	0	334	680	0	7	19	0	11,314
	Average new products	2.09	2.61	.	1.71	1.84	.	1	1.16	.	
	Maximum number	44	69	.	22	36	.	1	2	.	
2003	Firms	1335	2096	3212	501	785	1358	4	10	23	12,645
	Average new products	2.13	2.53	2.98	1.76	1.94	2.06	1	1	1	
	Maximum number	37	53	69	45	87	109	1	1	1	
2004	Firms	1474	1857	3112	438	544	1000	4	4	9	13,853
	Average new products	2.19	2.41	2.86	1.86	1.83	1.97	1	1	1	
	Maximum number	170	188	252	81	88	123	1	1	1	
2005	Firms	1649	1777	2757	358	371	593	9	9	13	13,379
	Average new products	2.12	2.19	2.61	1.80	1.83	1.90	1	1	1	
	Maximum number	115	117	148	32	34	43	1	1	1	
2006	Firms	1901	1901	2218	460	460	524	7	7	7	13,110
	Average new products	2.16	2.16	2.36	2.35	2.35	2.30	1	1	1	
	Maximum number	95	95	99	327	327	341	1	1	1	
2007	Firms	2692	2692	2692	405	405	0	22	22	22	13,033
	Average new products	2.29	2.29	2.29	2.13	2.13	.	1.05	1.05	1.05	
	Maximum number	47	47	47	97	97	.	2	2	2	

Source: Authors' own elaboration from SECEX

An interesting element is the sector composition of export diversification. Appendix 1 shows the number of new products and exporters as a share of total export flows per each HS-2 chapter. The percentage, therefore, is based on the number of product lines exported within each HS-2 chapter. The sectors with larger share of new exports are 60 fabrics, 41 skins and leather, 30 pharmaceutical, 31 fertilizers and 81 other base metals.[10]

Discoveries, new exported products for the country, are rare

Table 3.6 shows the number of new products that were not exported before the sample period and are new to the country. In total across classifications, we identify only 75 new discoveries.

[10] We also observe some new exports related to agricultural exports that correspond to firms that also produce agricultural products, even though manufacturing is considered to be their core activity.

Table 3.6 Number of discoveries (firm and product) according to different classifications

Year	Class 1	% of flows	Class 2	% of flows	Class 3	% of flows	number of flows
2001	3	0.00%	10	0.02%	0	0.00%	64509
2002	7	0.01%	22	0.03%	0	0.00%	67686
2003	4	0.00%	10	0.01%	23	0.03%	80461
2004	4	0.00%	4	0.00%	9	0.01%	91058
2005	9	0.01%	9	0.01%	13	0.01%	95773
2006	7	0.01%	7	0.01%	7	0.01%	94243
2007	23	0.02%	23	0.02%	23	0.02%	97259

Source: Authors' own elaboration from SECEX

In line with the findings of the literature, export discoveries in Brazil are rare. The maximum number of discoveries in a single year is 23 new products for 2003 and 2007. In addition, as shown in Table 3.5, discoveries are usually introduced one by one. Only in 2002 under classification method 2, and in 2007, we observe a single firm introducing two discoveries.

Table 3.7 shows the HS-2 sector composition of discoveries. Since some manufacturing firms are exporters in different sectors, we have some discoveries in agricultural products. Regarding manufacturing, most of the discoveries concentrate in the chemicals and pharmaceutical sectors, especially in organic chemicals with around a quarter (23 per cent) of discoveries. These sectors are followed by machinery and telecommunications.

Table 3.7 Product 'discoveries' by HS2 sector

HS2	Description	Number Discoveries	%
01	LIVE ANIMALS	1	1.33%
03	FISH & CRUSTACEANS	2	2.67%
08	ED. FRUITS & NUTS, PEEL OF CITRUS/MELONS	3	4.00%
09	COFFEE, TEA, MATE & SPICES	2	2.67%
10	CEREALS	1	1.33%
20	PREPS OF VEG, FRUIT, NUTS, ETC.	1	1.33%
25	SALT, SULPHUR, EARTH & STONE, LIME & CEMENT	3	4.00%
27	MINERAL FUELS, OILS, WAXES & BITUMINOUS SUB	3	4.00%
28	INORGANIC CHEMICALS	5	6.67%
29	ORGANIC CHEMICALS	17	22.67%
30	PHARMACEUTICAL PRODUCTS	6	8.00%
32	TANNING OR DYEING EXTRACTS; TANNINS	1	1.33%
38	MISCELLANEOUS CHEMICAL PRODUCTS	3	4.00%
39	PLASTICS AND ARTICLES THEREOF	2	2.67%
44	WOOD & ARTICLES OF WOOD, WOOD CHARCOAL....	2	2.67%
46	MATERIALS,	1	1.33%
48	PAPER AND PAPERBOARD; ARTICLES OF PAPER	1	1.33%
50	SILK	1	1.33%
51	WOOL, FINE OR COARSE ANIMAL HAIR; HORSE	1	1.33%
52	COTTON, INC. YARNS & WOVEN FABRICS THEREOF	1	1.33%

55	MAN-MADE STAPLE FIBRES,	2	2.67%
63	MADE-UP TEXTILE ARTICLES ...	1	1.33%
71	PEARLS, STONES, PREC. METALS, IMITATION JEWELLERY	1	1.33%
84	NUCLEAR REACTORS, BOILERS, MACHINERY,	4	5.33%
85	ELECTRICAL MACHINERY & EQUIP. & PARTS,.,..	4	5.33%
86	RAILWAY OR TRAMWAY LOCOMOTIVES, ROLLING STOCK,..	1	1.33%
90	OPTICAL, PHOTOGRAPHIC, CINEMATOGRAPHIC, , ..	2	2.67%
91	CLOCKS AND WATCHES AND PARTS THEREOF	2	2.67%
95	TOYS, GAMES & SPORTS EQUIP, PARTS & ACCES	1	1.33%

Source: Authors' own elaboration from SECEX

New exports travel to closer destinations and tend to have larger values

New exports tend to go, on average, to closer destinations. The average distance for new exported products is 5,287 kilometres, while the average for all export flows is 5,589 kilometres. This is consistent with the sequential exporting literature, suggesting that firms introduce products in geographically closed markets with less uncertainty, and then sequentially move to more distant markets. However, the difference in average distance does not appear to be large.

The sequential exporting literature also predicts that the size of the flow is sequentially increasing. In our case, the average size of new export flows is more than 2.5 times the size of an average export flow. Nevertheless, this is largely explained by the fact that all exports include a very large number of short lived flows, and as seen above the duration of the flows increase its size.

No clear increase in domestic unit values prior to export diversification

Iacovone and Javorcik (2010) find for a sample of Mexican firms preparation activities in the domestic market prior to exporting in general. The main idea is that firms upgrade product quality domestically before engaging in exporting that specific product. In line with this argument we look for new products dynamics in the domestic market prior to diversification in international markets.

There are three main challenges when looking at previous dynamics for the domestic market. First, the conversion from trade (HS-8) to production codes (prodlist) is not one to one. Some trade codes are mapped to more than one prodlist code. Second, our production data (PIA produto) is based on a survey. This means that not all firms and not all products are included in the survey. Third, the length of our sample is limited, and this constrains the number of periods where we can document domestic production dynamics.

In order to overcome these problems, we focus only on those new products that have one to one mapping from production to trade classifications. We analyse the evolution of unit values and quantities prior to exporting based on data on domestic production. We define unit values as the export value divided by the quantity deflated by the retail price index and normalised by the mean and standard deviation of each product. Unit values and quantities are regressed on a set of year dummies, product fixed effects and a time trend for the period before the product is introduced to capture any changes in growth of unit values and quantities. The results of the regressions suggest a coefficient on the time trend that is not statistically significant, and no evidence of any prior increases in unit values of domestic production for products being introduced for exporting is found.

It is possible that this result is explained by the fact that our sample only starts in 2000, and, therefore, there is little time dimension to oberve the dynamics prior to the introduction of new products. On average we only observe 2.7 years of new exports prior to being exported. However, for those new products with more than two years prior information, it was only in a very small number of cases that we observed a sustained increase in real unit values every year.

3.2 Relatedness

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

So far we have observed that although growth at the extensive margin of trade is small, the number of firms involved in diversification activities in a given year is quite substantial. As the resource based approach to firm diversification (Lien and Klein (2009)) suggests, firms diversify mainly to those activities where there are existing capabilities within the firm. If concentration of export activities is large, this should be translated into most firm diversification activities being similar or related to existing ones.

In the context of multi-product firms, measuring relatedness is problematic, since diversified production and export baskets imply a larger array of potentially similar diversification paths. In other words, for firms producing or exporting different products, it is easy to find a larger number of additional products that can be introduced by the firm and that are similar or related to the existing ones.

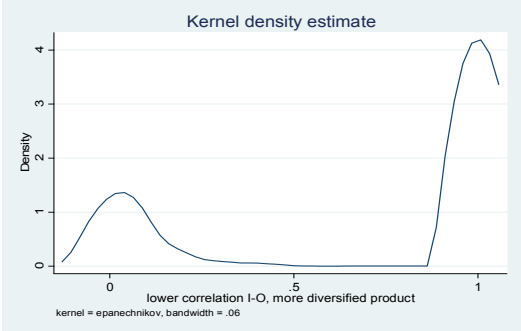
In order to identify how related the diversification path is, we implement the methodologies described in section 2 to our dataset and determine how new products relate to existing products. We use two different sets of products as reference. First, we compare new products exported in t with the closest product exported in t-1. Second, we use as a reference product the core activity for domestic production in t-1.

Figure 3.1 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.

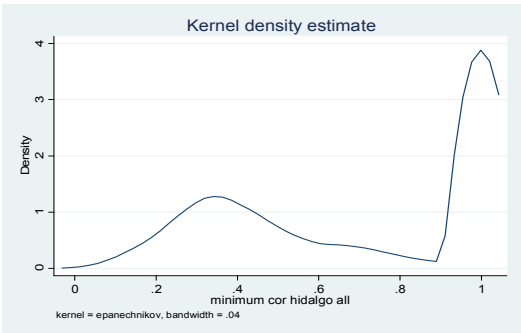
We define *related diversification* as the introduction of new products that are in the same product classification (zero distance) as the reference product or that have correlation one under the input use or Hidalgo's product space measures. Focusing on diversification vis-a-vis the export basket in 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.

Figure 3.1 Relatedness in diversification

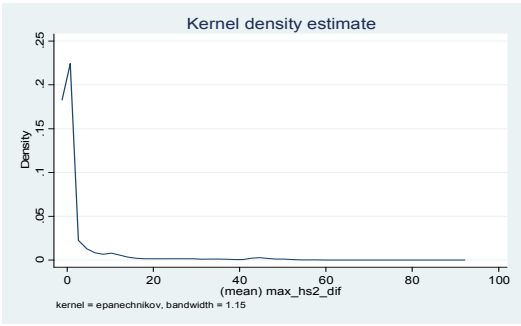
Relatedness vis-a-vis exports



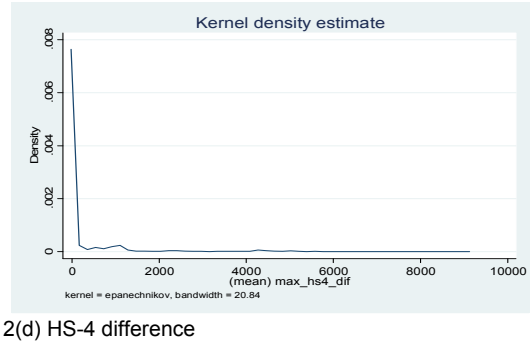
2(a) Input use similarity



2(b) Hidalgo *et al*. similarity

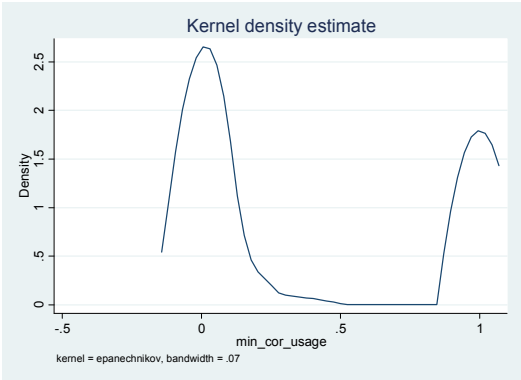


2(c) HS-2 difference

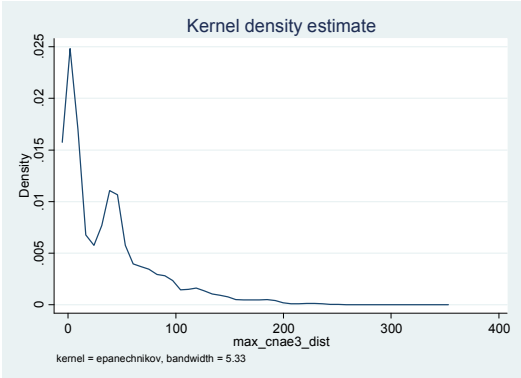


2(d) HS-4 difference

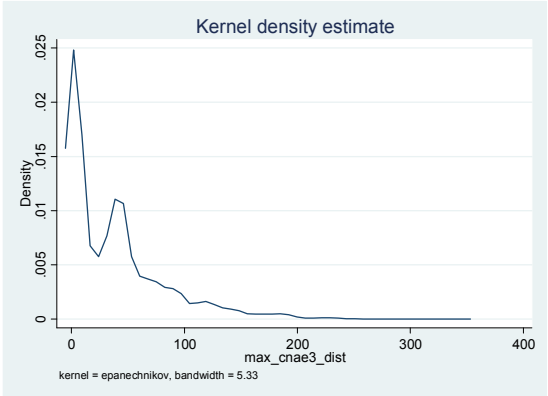
Relatedness vis-a-vis core product



2(e) Input use similarity



2(f) CNAE-2 difference



2(f) CNAE-3 difference

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 3.1 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 the 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.[11] The probability distribution functions plotted in the second column of Figure 3.1 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.

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

We further look at the issue of the unrelated diversification path described above by taking advantage of the fact that in our sample some firms introduce more than one new product simultaneously. 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.

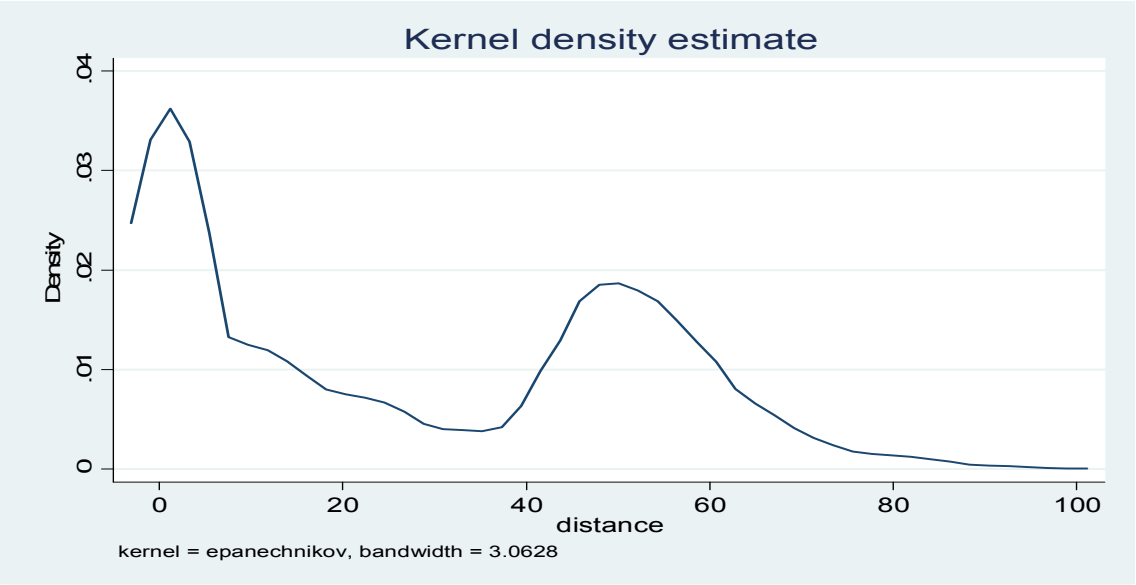
Figure 3.2 shows the probability distribution function of the calculated distances. Interestingly, there are two modes in the distribution. The first main mode occurs for observations with zero distance, where the new products introduced by the firm are in the same HS-2 chapter. There is, however, a 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 number of cases where diversification occurs in products of different sectors, these products appear to be part of different stages of the same value chain.

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

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

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.

Figure 3.2 Probability distribution function HS-2 distances – New products basket



Source: Author's own elaboration from SECEX

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 sector 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 significant frequency. Therefore, an interesting question in addition to what are the processes through which diversification occurs, 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?

3.3 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 2.2). 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 3.8 below shows the results. In the case of the PRODY index, in 64 per cent of cases new exports are less sophisticated, in 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 per cent indicate 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 one core production product 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 3.8 Sophistication/technological content changes in Diversification

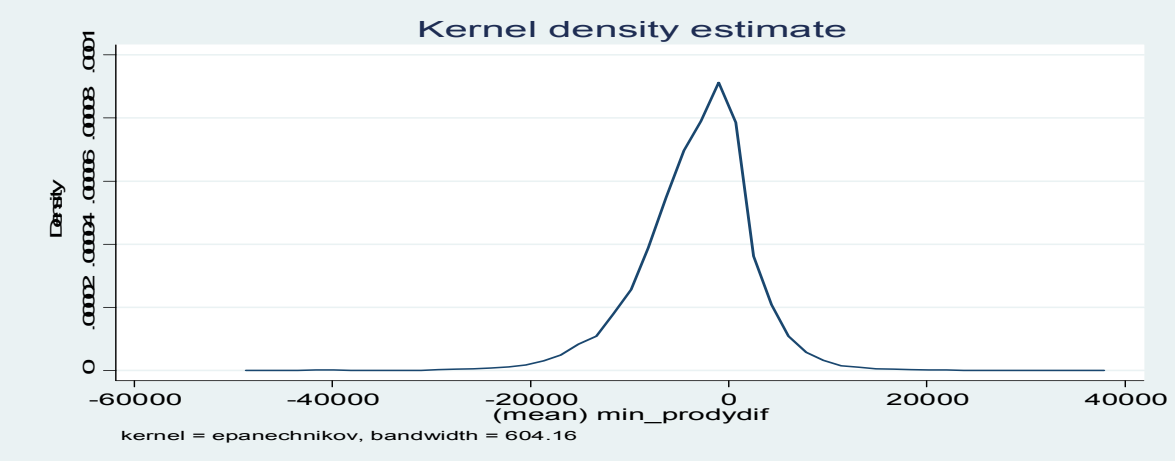
	Exports		Core production	
change in sophistication/technology content	PRODY	OECD	PRODY	OECD
lower level	63.70%	33.47%	6.67%	2.70%
same level	4.28%	60.13%	1.88%	12.20%
higher level	32.03%	6.40%	91.45%	85.10%

Source: Authors' own elaboration

In order to look at the size of upgrading/downgrading we plot the probability distribution function of the PRODY index with regard to differences between existing and new exported products (Figure 3.3). The figure shows the larger prevalence of new products have been of lower sophistication. It also shows a larger tail on the left indication that positive changes in PRODY, when they 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 3.3 Distribution of change in sophistication in product diversification (PRODY)



Source: Authors' own elaboration

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.

4 Determinants of firm export diversification in Brazil

In this section we analyse what factors are more important when explaining firm export diversification.

4.1. Baseline results

In order to determine the main factors explaining firm export diversification we estimate equation (1) using two alternative definitions of the dependent variable: (i) a dummy variable that assumes value 1 when the firm has introduced a new export under any of the possibilities included in Table 2.2 (Classifications 1, 2 and 3), and (ii) a variable that counts the number of new exports introduced by each firm.

One potential problem when estimating equation (1) with the dummy index for diversification is the risk of sample selection bias if we omit from the sample non-exporter firms. It is possible that unexplained factors that determine whether a firm is an exporter are also correlated with the probability of firm diversification. In order to correct this potential problem, we use a Heckman selection model and use the Heckprobit estimator. We first calculate the inverse mills ratio (IMR) from a first selection model for the probability of exporting using regional dummies for identification in the first stage given the large regional concentration of exporters. Then we use the IMR as a regressor of the level equation -the probability of a firm diversifying.

In a second stage, we focus on explaining the number of products introduced by firms that diversify. In this case, since the dependent variable is a count of the number of products, zero if the exporter does not diversify, we use a random effects Tobit estimator.

Table 4.1 shows the Heckprobit estimates for all firms that diversify. Even specifications show the selection equation for exporting, while odd specifications show the estimates for the probability of diversifying. As identifying variable for exporting, we use the share of national inputs used by the firm, Exporters, tend to use better inputs and engage more on trading activities. The assumption, therefore, is that the share of national inputs affects the selection into exporting but not the decision to diversify.

The results of the selection equation into exporting (even columns) show that larger and more productive firms tend to self-select for exporting. Also, as expected, the largest the share on national inputs the less likely to be an exporter.

Column (1) shows the result specification that maximises the size of the available dataset since it does not include any innovation variables, which are less frequent in the dataset. The remaining columns show the results for observations with more information on micro determinants, specifically a key determinant, product related innovations, and other sources of learning and innovation.

The Pseudo R^2 ranges from 0.10 and 0.28, and the Chi test for the overall non significance of the model is rejected. More importantly, the coefficient on the IMR is not statistically significant for most of the specifications, which implies that given our identification strategy we cannot accept the hypothesis of sample selection bias.

Comparison across specifications reveals the following patterns regarding the determinants of export diversification at the firm level. First, all five groups of determinants - the structural

characteristics of the firm; its position in the domestic market; the characteristics of the firm's production basket and processes; and the learning efforts carried out by the firm – seem to be relevant in explaining export diversification.

Second, when we look inside every group of determinants, the analysis of the impact of specific variables reveal that:

- Size definitely matters. Across all specifications the coefficient for natural log of employment is positive and significant. Larger firms are more prone to diversify.
- The multinationality of the firm, captured by the share of foreign capital in the firm ownership (ownership) also has a clear and significant positive effect on export diversification.[12]
- Integration into global business captured by the extent to which firms interact with foreign firms within their groups has a positive effect on the introduction of new exports. However, in this case the effect is less clear, since the coefficients are only significant at 10 per cent confidence level. This might be reflecting the fact that foreign share already captures some of the links and support from foreign firms within the same group to export.
- Global value integration proxied by client or buyer dependency appears to have a marginal negative impact on export diversification, but the coefficients are not statistically significant. Client dependency, however, positively impacts the selection of firms into export markets.
- Firms with larger market power find it easier to diversify and enter new product markets. This result suggests that in preparation for exporting, the position in the domestic market, as reflected by domestic shares, makes a positive contribution to the possibility to introduce and sustain new exports. This is because firms with larger shares are likely to have more financial leverage to introduce new exports.
- The effect of increasing quality, proxied by the unit value, does not appear statistically significant in all specifications. It is possible that firms focus on improving quality on existing products rather than on introducing new products.
- Firms with a basket of production highly concentrated in value have less probability to introduce new exports. This is explained by the fact that these firms might opt to expand exports of existing products, their core business, rather than expanding the type of products exported.
- Sector diversification of the production base of the firm has a positive effect on firm selection to exporting and export diversification. This reflects the importance of diverse production and technological capabilities for new exports. Firms with capabilities that span along different types of products are more likely to introduce new products. However, the effect on diversification is not statistically significant when considering innovation activities. Not surprisingly, and in line with the findings of previous studies, firms that engage in product related innovations are more likely to diversify into new exports.
- The characteristics of the process of production regarding both efficiency and innovativeness, are also highly relevant in explaining the capacity of firms to introduce new products for export. More productive firms are more likely to diversify. In addition, process innovation also increases the likelihood of diversifying.
- Engagement in R&D activities is a significant learning effort for firms aiming to diversify, the variable enters positive and significant in most estimations in which we include it. We find the same with skills and marketing efforts. They appear both highly significant in the two estimations included.
- The coefficients on other investments in innovation are not statistically significant. Some of the effects of these other investments may be captured by the sector dummies, since these types of efforts are only relevant for certain sectors, such as the natural resources related sector.
- Finally, the effect of geographical distance of firms' exports to proxy for exporting experience is not statistically significant in most specifications.

[12] We do not have an indicator of multinationality of domestic firms, so we cannot see if this variable also operates in the case of this type of firm.

Table 4.1 Heckprobit estimates determinants of export diversification

	(1) Diversification	(1) Selection	(2) Diversification	(2) Selection	(3) Diversification	(3) Selection	(4) Diversification	(4) Selection
TFP using Levinsohn and Petrin and value added	0.1117***	0.3168***	0.1046***	0.2809***	0.1161***	0.2367***	0.1143***	0.2356***
	(0.014)	(0.005)	(0.019)	(0.009)	(0.025)	(0.015)	(0.025)	(0.015)
log employment	0.1617***	0.3767***	0.1403***	0.3580***	0.1652***	0.3507***	0.1731***	0.3496***
	(0.011)	(0.005)	(0.017)	(0.009)	(0.023)	(0.013)	(0.023)	(0.013)
ratio unit value to product average	0.0130**	0.0142***	0.0055	0.0166**	0.0015	0.0106	0.0030	0.0119
	(0.004)	(0.003)	(0.009)	(0.005)	(0.012)	(0.009)	(0.012)	(0.009)
firm market share by product	0.2857***	1.2750***	0.1910*	1.0825***	0.3101**	0.9642***	0.3278***	0.9776***
	(0.051)	(0.037)	(0.075)	(0.061)	(0.095)	(0.088)	(0.095)	(0.088)
herfindahl concentration normalised of production	-0.2521***	-0.0553***	-0.2478***	-0.0624**	-0.2673***	-0.0400	-0.2790***	-0.0468
	(0.021)	(0.012)	(0.033)	(0.021)	(0.043)	(0.033)	(0.043)	(0.033)
distance CNAE 2 digits divisions	0.0071***	0.0172***	0.0070*	0.0143***	0.0056	0.0144***	0.0056	0.0145***
	(0.002)	(0.001)	(0.003)	(0.002)	(0.004)	(0.004)	(0.004)	(0.004)
(mean) distance	-0.0000		-0.0000		-0.0000*		-0.0000**	
	(0.000)		(0.000)		(0.000)		(0.000)	
dummy for product innovation					0.2338***	0.1323***		
					(0.034)	(0.028)		
dummy for process innovation							0.0820*	0.0942**
							(0.035)	(0.030)
RDdummy			0.1620***	0.1697***				
			(0.030)	(0.027)				
Group dependency			0.1415*	0.1087	0.0842	0.1095	0.1174*	0.1284
			(0.056)	(0.089)	(0.057)	(0.088)	(0.057)	(0.088)
Client dependency			-0.0754	0.8964***	-0.0443	0.8867***	-0.0273	0.8954***
			(0.069)	(0.119)	(0.072)	(0.119)	(0.072)	(0.120)
Foreign capital			0.1293**	0.6465***	0.2184***	0.7612***	0.2024***	0.7595***
			(0.045)	(0.052)	(0.054)	(0.054)	(0.054)	(0.054)
Other Innovation expenditure					0.0000	0.0000	0.0000*	0.0000
					(0.000)	(0.000)	(0.000)	(0.000)
marketing					0.0969**	0.1295***	0.1137***	0.1377***
					(0.031)	(0.025)	(0.031)	(0.025)
number of high skill technical staff					1.7245**	-0.0525	2.3988***	-0.0467

					(0.648)	(0.077)	(0.636)	(0.072)
university					-0.0054	0.0232	0.0117	0.0110
					(0.031)	(0.025)	(0.033)	(0.029)
independent or group					-0.0380	-0.0167	-0.0231	-0.0078
					(0.037)	(0.034)	(0.037)	(0.034)
IMR	-0.1761**		-0.1433		-0.0181		-0.0412	
	(0.054)		(0.077)		(0.119)		(0.119)	
National input		-0.0085***		-0.0080***		-0.0046***		-0.0048***
		(0.000)		(0.000)		(0.001)		(0.001)
Constant	-3.1118***	-6.3201***	-2.9797***	-6.0423***	-3.3875***	-5.7455***	-3.4787***	-5.7676***
	(0.563)	(0.205)	(0.804)	(0.353)	(0.857)	(0.399)	(0.865)	(0.400)
Observations	37.686	107.199	16.065	36.923	9.101	15.716	9.101	15.716
log-likelihood	-19882	-53331	-8038	-18330	-4780	-7610	-4801	-7617
Pseudo R2	0.104	0.269	0.124	0.275	0.140	0.288	0.137	0.288

Dependent variable new products under all classifications for existing exporters and new exporters

*** 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 4.2 Random effects Tobit estimates of number of new products for exporting

	(1)	(2)	(3)	(4)	(5)
TFP using Levinsohn and Petrin and value added	-0.109*	-0.17*	-0.184**	-0.175*	-0.198**
	(0.0534)	(0.0702)	(0.0702)	(0.0703)	(0.0707)
log employment	1.072**	1.238**	1.188**	1.237**	1.276**
	(0.0772)	(0.1056)	(0.1052)	(0.1069)	(0.1072)
ratio unit value to product average	0.143**	0.149*	0.151*	0.156*	0.16*
	(0.0514)	(0.0741)	(0.0737)	(0.0739)	(0.0737)
firm market share by product	3.379**	2.77**	2.786**	2.845**	3.09**
	(0.4076)	(0.5379)	(0.5347)	(0.5398)	(0.5374)
Herfindahl concentration normalised of production	-1.755**	-1.76**	-1.756**	-1.844**	-1.89**
	-0.1992	-0.2733	-0.2722	-0.2740	-0.2739
distance CNAE 2 digits divisions	0.002	-0.015	-0.011	-0.015	-0.008
	-0.0200	0.0254	-0.0244	0.0250	-0.0258
(mean) distance	0.000**	0.000**	0.000**	0.000**	0.000**
	0.0000	0.0000	0.0000	0.0000	0.0000
dummy for product innovation		1.669**	1.847**	0.334	0.564**
		0.2344	0.2066	0.2012	0.2104
RDdummy	1.03**	0.427		1.249**	
	0.1746	(0.2454)		(0.2180)	
Other_Innov_Exp	0.000*		0.000**		0.000**
	(0.0000)		(0.0000)		(0.0000)
foreign_cap1	1.427**	2.297**	2.288**	2.293**	2.309**
	(0.2452)	(0.2771)	(0.2760)	(0.2779)	(0.2782)
group_dep1	0.764*	0.587	0.581	0.831*	0.863**
	(0.3056)	(0.3298)	(0.3282)	(0.3298)	(0.3294)
client_dep1	-0.341	-0.469	-0.398	-0.377	-0.219
	(0.3789)	(0.4043)	(0.4061)	(0.4054)	(0.4056)
number of high skill technical staff		12.322**	12.393**	13.265**	17.832**
		(3.8267)	(3.7328)	(3.8338)	(3.6767)
independent or group		-0.156	-0.193	-0.118	-0.094
		(0.2229)	(0.0000)	(0.0000)	(0.0000)
significant marketing changes			0.476**		0.621**
			(0.0000)		(0.0000)
high information from university			-0.252		-0.089
			(0.0000)		(0.0000)
Constant	-10.568**	-12.102**	-11.73**	-11.656**	-11.716**
	(0.7145)	(0.9713)	(0.0000)	(0.0000)	(0.0000)
Observations	16069	9103	9103	9103	9103
Number of group	9014	5406	5406	5406	5406

*** 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.

In the second stage we focus on explaining the number of products, given the fact that some firms introduce in a given year, more than one new product for exporting. The sample is restricted to firms of exporters, with a zero value to those that do not diversify. We first try to estimate equation (1) with a Poisson fixed effects estimator. However, two main problems

arose when trying to implement this estimator. Firstly, estimates do not converge when using sector dummies. As a result it is possible that some of the coefficients are capturing sector wide effects. Secondly, and more importantly, the FE Poisson estimator eliminates groups (firms) with only one observation or where there are only zero outcomes. This implies that is mainly estimating the sample of diversification firms.

As a result of these problems, we implement a Tobit random effects estimator. Although the estimator does not converge with sector dummies, it can handle the zero truncation in the sample at the same time as allowing us to estimate the model for the full exporter sample.

Table 4.2 shows the results of the Tobit RE estimator. The Chi test for the overall non significance of the model is rejected, and the random effects model explains 35 per cent of the variance and, therefore, should be preferred to the pooled estimator. The results for the number of new products introduced are very similar to the ones in Table 4.1. Most of the channels for acquiring capabilities discussed above, explain the number of new products. However, there are now three main differences.

The first difference is the role of productivity. More productive firms appear to introduce less new products. In addition, the role of the sector diversification of the product base is now not statistically significant, while quality levels of the firm appear to marginally impact the increase in the number of products.

On the other hand, strong determinants of the number of new products introduced are size, foreign ownership, lack of business concentration, average distance of exports, introduction of product improvements and number of high skilled staff.

4.2 Robustness checks

4.2.1 Other measures of diversification

The results shown above are based on diversification identified using the more flexible methodologies regarding sustainability of exports. As a robustness check, we re-estimate the same baseline models using the most conservative methodology, which requires that once the new product is introduced, it is exported until 2008.

Table A4.1 in Appendix 4 shows the results of the estimates for the Heckprobit second stage, the probability of diversification, with very similar results to the ones in Table 4.1 above, both in size and sign of the coefficients, suggesting that the results above are not sensitive to the diversification methodology used.

4.2.2 Survivor sample

One consequence of having to use some criteria of sustainability when identifying diversification is the possibility of distorting the sample that we are using. Concretely, one concern is the fact that if firm turnover is large we may be 'contaminating' our estimates with factors that are not related to exporting but to firm closure. With the methodologies used in the paper, any firm that diversifies in t but goes out of business in t+1 or t+2 would not qualify as a firm that diversifies, but it would as an exporter in t. If the prevalence of these firms is large, we would be comparing our diversification firms with exporters, a large share of which are likely to be out of business in the coming periods. As a result, some of the factors affecting firm closure would contaminate the firm dynamics and strategies of exporters.

In order to test whether closure factors may be affecting our results, we restrict our sample to only those firms that are in the dataset for the entire period of the sample (our survivor sample) and re-estimate the baseline model. The tables A4.2 and A4.3 in Appendix 4 show the results. The loss of number of firms and observations is moderate, indicating that closure was not an important part of the sample. More importantly, the results are very similar to the ones in Table 4.1 and Table 4.2, suggesting that main estimates are robust to non-survivor biases.

4.3 Summing up

The results of the estimations indicate that all five groups of determinants identified in the literature play important roles explaining firm export diversification. Firm characteristics regarding size and ownership appear to be important in explaining the decision to introduce a new product for exporting. The position in the domestic market especially regarding market share and the business concentration of firm activities are also important when explaining the decision to diversify. Also, another important finding is the importance of the learning efforts of the firm. These are particularly important in relation to improvements in products, process, other R&D activities, marketing and using high skilled staff. These activities are clearly conducive to acquiring the required capabilities for diversification. The results, however, are less clear in relation to global engagement. In some specifications, belonging to a group increases the likelihood of diversification, but client or buyer dependency reduces the likelihood that the firm will introduce a new product for exporting. However, the results are not statistically significant in most specifications.

In general, what the results uncover is that specific firm characteristics are more conducive than others to firm export diversification. Firms prepare for diversification by using their position and business strategies in the domestic market, and more importantly adopting specific innovation efforts and learning.

5 Explaining the export diversification paths of Brazilian firms

In previous sections we identified two features to distinguish diversification paths that 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. These two key elements are the degree of relatedness and sophistication of the diversification path. The first helps to understand how capable firms are of expanding in new activities, potentially in new sectors away from traditional exports. The second is important in determining how capable firms are of moving to higher value added activities. Countries with firms more capable of moving 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 matter more or less in order to achieve diversification paths that are more unrelated and sophisticated.

5.1 Relatedness

In section 3 we decomposed diversification cases according to whether they occurred in highly related activities, *related diversification*, or in somehow unrelated activities; based on a different set of classification methodologies. The descriptive analysis suggested higher prevalence of diversification towards very similar or highly related activities. Most diversification occurs in a related diversification path. However, as Figure 3.1 shows, there are some cases where diversification to non-totally related activities occurs. Section 3 shows that the extent of unrelatedness in diversification largely depends on the measure used. It tends to be larger for measures based on input use and Hidalgo's product space correlation, and lower for trade and industry classification based measures.

In this section, we use our existing dataset to go beyond diversification and try to explain related and unrelated diversification paths. The main assumption is that different firm characteristics, capabilities and business strategies, are required to achieve more or less related diversification paths that are more or less similar to existing capabilities and competencies. The objective, therefore, is to determine which of the five groups of determinants identified in the literature on diversification is more conducive to unrelated diversification.

5.1.1 Results

5.1.1.1 Relatedness vis-a-vis exports

Based on the methodology described in section 2, we estimate equation (2) using a multinomial Logit model. The results of the estimations are summarised in Table 5.2. 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 Leontieff 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. As we saw in the previous section, productivity and size increase the probability of a firm diversifying. The relative contributions of these two variables to both categories, related and unrelated, 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 has a larger increase in 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 Leontieff 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 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.

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, as we saw before firms that have more concentrated business structures and whose sales depend on less 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 within 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 also have 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 confirm the findings of the previous section on the determinants of diversification. When looking more specifically to related and unrelated diversification, only the quality level of the firm, proxied by unit values and, in particular, the degree of diversification of the production structure, has 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 in related diversification for input use and classification based methodologies, and also for unrelated diversification under Hidalgo's product space methodology. This suggests some degree of similarity between the technology dimension of relatedness measures adopting input use and classification based measures.

Table 5.2 Multinomial Logit estimates related diversification vis-a-vis exports

	Leontieff		Leontieff		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.

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 continuous 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 5.3 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, in higher sector unrelatedness.

The parameter rho is the panel-level variance component. It is statistically significant in most cases, but not in the case of HS-4 differences, which indicated for this case that 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 business strategy that focuses on expanding a few sets of existing products. The results suggest that the former is more important than the latter. Finally, product improvements in t-1 are likely to result in a greater likelihood of the firm introducing new, unrelated products.

With regard to unrelated diversification measured by changes in HS classification, an interesting result is that the higher the degree of dependence from clients and buyers, the more related 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 with similar goods, or the lack of incentives for firms to develop different export products and find 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.

Table 5.3 Tobit RE estimates related diversification vis-a-vis exports

	leontieff 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(empresa_)	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.

5.1.1.2 Relatedness vis-a-vis core production activity

Relatedness can also be expressed in relation to the core production activity of the firm. 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 on how firms build the specific capabilities required for exporting, such as information or marketing.

As suggested by the left hand side panel in Figure 3.1, 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 5.4 shows the results of the multinomial logit estimations for relatedness with 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 Leontieff 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 increase the odds of exporters to diversify. In this case, however, the coefficient suggests that productivity and size increase the probability of diversifying more 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 instead continue to export 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 it is that exporters will diversify to related activities and the more likely they are to diversify to unrelated sectors. 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, largely statistically insignificant. 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 greater 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.

Table 5.4 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.

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.

The main similarity between both types of relatedness, in relation to exports and to core activities, 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 5.5). Positive signs increase relatedness diversification of the Leontieff 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.

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 5.3. 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.

5.1.1.3 Robustness

In addition to using different measures of relatedness and different variables specifications, we also re-estimate the different specifications using the survivor sample. As suggested above, the fact that we have used a criterion of sustainability of exports in the definition of diversification implies that we could be introducing some bias in the coefficients by comparing firms that survive with firms that disappear.

Table 5.5 Tobit RE estimates related diversification vis-a-vis core production activity

	leontieff 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 survivor sample, however, contains a large proportion of the normal sample, which implies that it is very unlikely that survivor factors may be important in the estimations. This is confirmed by the estimations that produce very similar results to the ones described above.[13] The results, therefore, are not affected by factors that may determine firms going out of business

5.1.2 Summary of relatedness results

The results in this section confirm the main determinants of diversification found in section 5. In general most of the determinants of diversification also explain both related and unrelated diversification, and there are only a few sets of determinants that appear to have more impact in one type of diversification or the other.

When looking more specifically at the difference between both types of diversification we found that the most important variable explaining 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.

Section 3 showed that part of this 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 chain 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. One implication of this is that measures based on input use and classification differences are not well suited for measuring this type of relatedness. Furthermore, measures based on the product space may also fail to capture this type of relatedness when, for example, size or ownership may constrain the number of firms able to acquire all the capabilities.

Other findings suggest that firms with larger market power are more likely to diversify to unrelated activities where there is more room to grow, especially in relation to core activities; and that this was also the case with firms that introduced product innovations. Finally, 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.

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 3 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 greater 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 3 showed that there is larger prevalence of introducing new products with less sophistication and the same or less 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.

[13] The results are available upon request.

The objective of this section is to identify which of the group of determinants reviewed above are more conducive to one type of diversification or the other. In order to do so we implement the methodology described in section 2.2.4, and estimate equation (4) for both, the PRODY measures and the OECD technology index.

5.2.1 Results

5.2.1.1 Sophistication vis-a-vis exports

The results of the estimates regarding the sophistication measure in relation to exports are summarised in Table 5.6. 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 it 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 is it likely that two different products have the same PRODY value. The Pseudo-R^2 is still low, around 0.09.

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 which are 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 in explaining relatedness, such as market power, quality and, more importantly, the diversification of the production structure, are now not statistically significant.

Perhaps the most important result is the importance of the average distance of exports. 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 likely to be biased towards exporting primary commodities and processed natural resources to developed and Asian markets, while exporting manufacturing products to the region; therefore with very low scope for quality differentiation.

A puzzling result 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 5.7 shows the estimates when using the technological content index. Now the control group 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.

Table 5.6 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 5.7 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

5.2.1.2 Sophistication vis-a-vis core activity

Table A5.1 in Appendix 5 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 A5.2, 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.

5.2.2 Relatedness and sophistication as a joint decision

One way of addressing this potential sample 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 A5.3 to A5.10 in Appendix 5 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 Leontieff and the OECD index, and in relation to exports, but the results are 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 Conclusion and policy implications

This paper analysed the determinants of firm export diversification in Brazil. We used a unique dataset that links data on trade, innovation, production and firm characteristics in order to understand the main firm determinants of export diversification during the period 2000 to 2008.

A first decomposition of the export diversification process in Brazil shows an interesting number of stylised facts. Some of these stylised facts support some of the existing findings in the trade literature, while others are new contributions in the area of export diversification at the firm level. Notably, the results of the decomposition suggest the following stylised facts:

- There is high export firm activity but a low survival rate of export flows, suggesting very short duration and intermittence of export flows at the firm level.
- Higher survival rates are correlated with larger export flows. Export flows that last longer tend to have larger fob values.
- Larger export flows are going to more distant markets.
- Most exporters appear to concentrate on few products for export, but a few firms export a very large number of products.
- Firm export diversification occurs for a relative low number of exported flows, but a significant number of firms that are exporters engage in export diversification. This implies that growth at the extensive margin is mainly carried out by existing firms rather than new entrants.
- Discoveries, new exported products for the country, are rare. In total across our sample period we identify only 75 new discoveries.
- New exports travel to closer destinations and tend to have larger values.
- 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;
- However, 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 with less sophisticated or technology intensive products than existing ones, and with more sophisticated or technological intensive products than the core competences of the firm.

The paper also sheds some light on the main firm determinants of diversifications as identified in the literature review. We divide these determinants into five main groups: the structural characteristics of the firm; its position in the domestic market; the characteristics of the firm's production basket; the characteristics of the production processes; and the learning efforts carried out by the firm.

The econometric analysis of export diversification reveals that all of these five groups are relevant determinants of firms introducing new products for exporting. We find that:

- Larger and foreign owned firms are more prone to diversify, as well as firms that interact with foreign firms within their groups.
- Firms with larger market power find it easier to diversify and enter new product markets.
- Firms with a basket of production highly concentrated in value have less probability to introduce new exports.
- The innovativeness of the production basket is highly significant in explaining the introduction of new exports.
- The characteristics of the process of production regarding both efficiency and innovativeness, are also highly relevant in explaining the capacity of firms to introduce new products for export.
- Engagement in R&D activities is a significant learning effort for firms aiming to diversify, and the same happens with increasing skilled labour and marketing efforts.

The paper also looked at two dimensions of the diversification path: relatedness and sophistication. Regarding the degree of relatedness in diversification, we found that the most important variable explaining 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. It also potentially reflects firm specialisation in different stages of a value chain. We also found that firms with more market power were more likely to diversify to unrelated activities, especially in relation to core activities, and that firms that introduced product innovations were also more likely to diversify to unrelated activities.

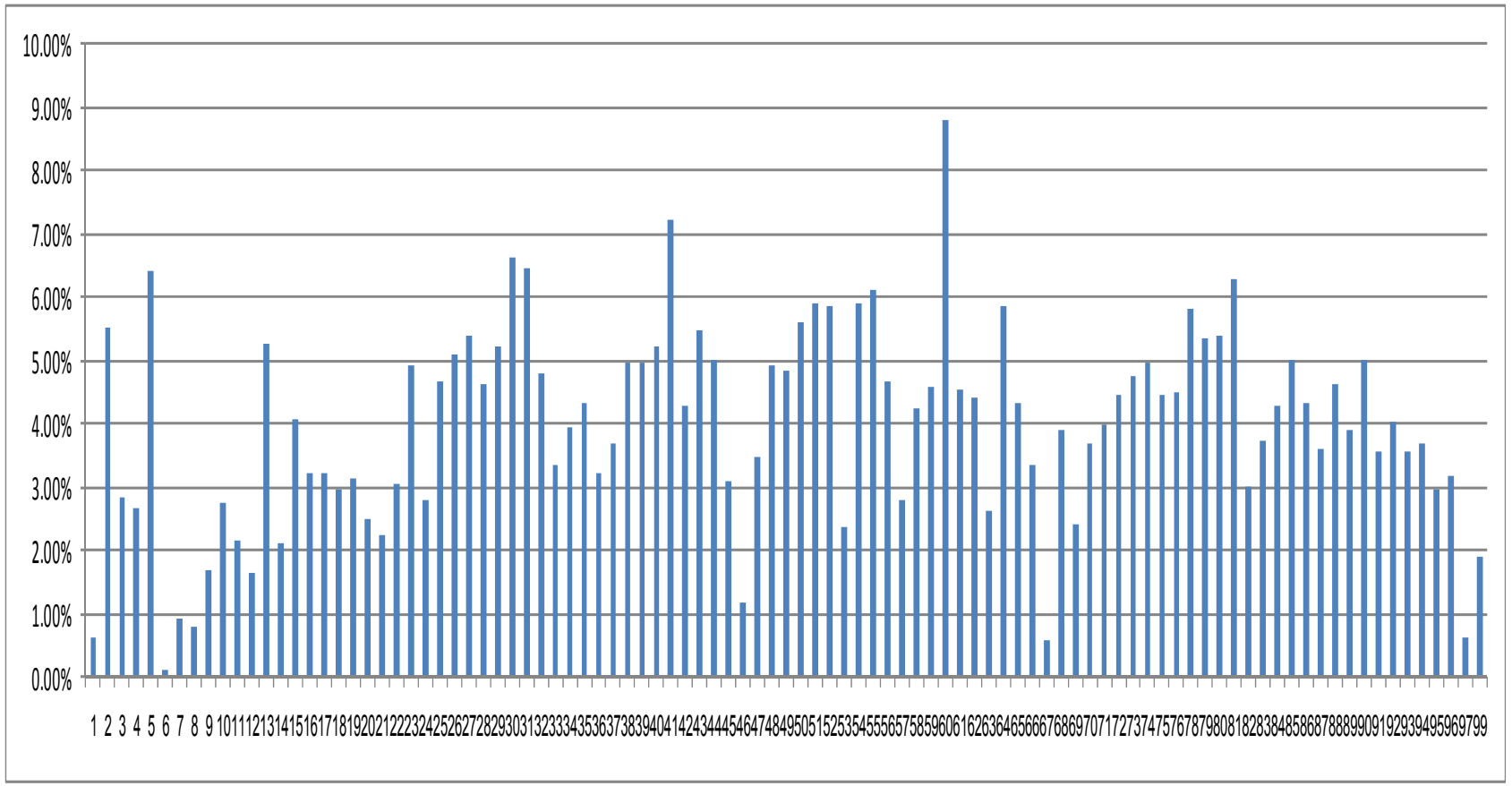
Finally, we attempted to explain the determinants of upgrading/downgrading in sophistication and technological content. The results, however, do not point out any of the five groups of firm determinants robustly explaining upgrading/downgrading. The level of sophistication of the diversification path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

In summary, this paper contributes to the literature on export diversification and on preparation for exporting. We have documented the specific set of firm characteristics that appear to be more conducive to firm export diversification. More importantly, the results uncover that firms prepare for diversification by using their position and business strategies in the domestic market, and more importantly adopting specific innovation efforts and learning activities.

The findings suggest the following policy implications. First, innovation activities of the firm are important in acquiring the required capabilities to diversify. Therefore, policies that support innovation also support diversification. Second, domestic production dynamics such as market power or sector diversification, are important determinants of diversification. This suggests that export support policies should also consider the extent to which firms first improve their production base domestically prior to exporting. Finally, foreign exposure also appears to increase diversification, and, therefore, policies that facilitate FDI and links with international markets also support diversification.

This research has raised some important questions for further research. First, we need a better understanding on the relationship between diversification for the domestic market and diversification for exporting. This will lead to a better understanding of the specific capabilities required for exporting. Second, 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. Third, we need a better understanding of the specific processes through which firms are able to diversify to more sophisticated products.

Appendix 1 Distribution of new products and exporters by HS2 chapter – percentage of total product lines by sector



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 MILK; PASTRYCOOKS' PRODUCTS	77	15	19.48%	0.37%	6	7.79%	0.12%	20	25.97%	0.29%
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;	12	5	41.67%	0.12%		0.00%	0.00%	5	41.67%	0.07%

	HORSEHAIR YARN AND WOVEN FABRIC										
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	total	PRODY upgrade firms	PRODY upgrade % upgrade	PRODY upgrade % total	OECD upgrade firms	OECD upgrade % upgrade	OECD 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 Diversification estimates

Table A4.1 Heckprobit estimates conservative classification

	(1)	(2)	(3)	(4)
TFP using Levinsohn and Petrin and value added	0.0991***	0.1157***	0.0989***	0.0976***
	(0.015)	(0.022)	(0.027)	(0.027)
log employment	0.1540***	0.1336***	0.1629***	0.1733***
	(0.012)	(0.018)	(0.024)	(0.024)
ratio unit value to product average	0.0117*	0.0034	-0.0090	-0.0065
	(0.005)	(0.009)	(0.014)	(0.014)
firm market share by product	0.2750***	0.2129*	0.3225**	0.3461***
	(0.055)	(0.083)	(0.103)	(0.103)
herfindahl concentration normalised of production	-0.2219***	-0.2254***	-0.2507***	-0.2640***
	(0.023)	(0.036)	(0.048)	(0.048)
distance CNAE 2 digits divisions	0.0058**	0.0101**	0.0077	0.0079
	(0.002)	(0.003)	(0.004)	(0.004)
(mean) distance	-0.0000	-0.0000	-0.0000**	-0.0000***
	(0.000)	(0.000)	(0.000)	(0.000)
dummy for product innovation			0.2621***	
			(0.038)	
dummy for process innovation				0.0743
				(0.039)
RDdummy		0.1648***		
		(0.033)		
Group dependency		0.1244*	0.0825	0.1186*
		(0.059)	(0.060)	(0.060)
Client dependency		-0.0513	-0.0201	0.0042
		(0.076)	(0.079)	(0.079)
Foreign capital		0.0933	0.1866**	0.1672**
		(0.048)	(0.058)	(0.058)
Other Innovation expenditure			0.0000	0.0000
			(0.000)	(0.000)
marketing			0.0900**	0.1099**
			(0.034)	(0.034)
number of high skill technical staff			0.5725	1.3441
			(0.724)	(0.694)
university			-0.0460	-0.0201
			(0.035)	(0.036)
independent or group			-0.1020*	-0.0847*
			(0.041)	(0.041)
IMR	-0.1931***	-0.1139	-0.0272	-0.0558
	(0.058)	(0.084)	(0.130)	(0.130)
Constant	-8.8910***	-8.3322	-8.1709	-8.3159
	(0.401)	(0.000)	(0.000)	(0.000)
Observations	37,686	16,019	9,071	9,071
log-likelihood	-16083	-6269	-3788	-3811
Pseudo R2	0.101	0.117	0.135	0.130

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

Table A4.2 Heckprobit estimates determinant of export diversification – Survivor sample

	(1) Diversification	(1) Selection	(2) Diversification	(2) Selection	(1) Diversification	(3) Selection	(2) Diversification	(4) Selection
lnTFPLP	0.1307***	0.3286***	0.1071***	0.3002***	0.1111***	0.2368***	0.1075***	0.2353***
	(0.015)	(0.006)	(0.022)	(0.011)	(0.026)	(0.017)	(0.026)	(0.017)
lnl	0.1802***	0.3860***	0.1455***	0.3634***	0.1767***	0.3627***	0.1828***	0.3629***
	(0.012)	(0.006)	(0.018)	(0.011)	(0.024)	(0.016)	(0.024)	(0.016)
uv_ratio	0.0165***	0.0153***	0.0098	0.0215***	0.0039	0.0037	0.0065	0.0049
	(0.005)	(0.003)	(0.009)	(0.006)	(0.013)	(0.011)	(0.013)	(0.011)
prod_share	0.4138***	1.2973***	0.2924***	1.0960***	0.3957***	1.0061***	0.4141***	1.0174***
	(0.055)	(0.041)	(0.081)	(0.068)	(0.101)	(0.098)	(0.101)	(0.098)
herf2	-0.2643***	0.0767***	-0.2775***	0.0855***	-0.2971***	-0.0344	-0.3092***	-0.0404
	(0.022)	(0.014)	(0.035)	(0.024)	(0.045)	(0.037)	(0.045)	(0.037)
dist_cnae2	0.0076***	0.0160***	0.0073*	0.0126***	0.0061	0.0122**	0.0062	0.0123**
	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
prod_inno					0.2362***	0.1123***		
					(0.036)	(0.031)		
process_inno							0.0904*	0.0767*
							(0.037)	(0.033)
Other_Innov_Exp					0.0000	-0.0000	0.0000*	-0.0000
					(0.000)	(0.000)	(0.000)	(0.000)
marketing					0.1103***	0.1232***	0.1262***	0.1301***
					(0.033)	(0.027)	(0.032)	(0.027)
skills_rd					2.4006**	4.3734***	3.1289***	4.8976***
					(0.783)	(0.979)	(0.781)	(0.984)
RDdummy			0.1650***	0.1883***				
			(0.032)	(0.030)				
group_dep1			0.1490*	0.0917	0.0834	0.0845	0.1166	0.1029
			(0.060)	(0.098)	(0.061)	(0.098)	(0.061)	(0.098)
client_dep1			-0.0611	0.8758***	-0.0226	0.8401***	-0.0053	0.8496***
			(0.073)	(0.131)	(0.076)	(0.131)	(0.076)	(0.131)
foreign_cap1			0.1790***	0.6727***	0.2879***	0.7959***	0.2696***	0.7946***
			(0.048)	(0.058)	(0.058)	(0.060)	(0.058)	(0.060)
dist_allexp	-0.0000*		-0.0000		-0.0000*		-0.0000**	
	(0.000)		(0.000)		(0.000)		(0.000)	

v3a					-0.0056	0.0044	0.0090	0.0126
					(0.039)	(0.038)	(0.039)	(0.038)
university					-0.0002	0.0357	0.0120	0.0257
					(0.033)	(0.028)	(0.035)	(0.032)
imr2	-0.0523		-0.0806		0.0766		0.0402	
	(0.058)		(0.084)		(0.127)		(0.127)	
		-		-		-		-
Constant	-3.5933***	6.5407***	-3.1506***	6.3390***	-3.6017***	5.9185***	-3.6495***	5.9362***
	(0.573)	(0.209)	(0.817)	(0.364)	(0.870)	(0.416)	(0.877)	(0.417)
Observations	32,230	86,802	13,590	28,536	7,994	13,115	7,994	13,115
log-likelihood	-17772	-43622	-7134	-14312	-4324	-6247	-4342	-6250
Pseudo R2	0.103	0.270	0.125	0.275	0.145	0.288	0.141	0.287

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

Table A4.3 Random effects Tobit estimates number of new products – Survivor sample

	(1)	(2)	(3)	(4)	(5)
TFP using Levinsohn and Petrin and value added	-0.159**	-0.244**	-0.226**	-0.219**	-0.243**
	(0.0556)	(0.0711)	(0.0711)	(0.0713)	(0.0715)
log employment	1.069**	1.322**	1.194**	1.244**	1.276**
	(0.0818)	(0.1053)	(0.1076)	(0.1093)	(0.1094)
ratio unit value to product average	0.167**	0.185*	0.184*	0.19*	0.2**
	(0.0534)	(0.0768)	(0.0763)	(0.0769)	(0.0766)
firm market share by product	3.649**	3.091**	2.739**	2.826**	3.053**
	(0.4185)	(0.5423)	(0,5402)	(0.5445)	(0.5423)
Herfindahl concentration normalised of production firm year	-1.767**		-1.745**	-1.839**	-1.87**
	(0.2036)		(0.2731)	(0.2749)	(0.2746)
distance CNAE 2 digits divisions	0.002	0.028	-0.007	-0.009	-0.003
	(0.0200)	(0.0239)	(0.0241)	(0.0237)	(0.0273)
(mean) dist	0**	0**	0**	0**	0**
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
dummy for product innovation		1.36**	1.736**	0.326	0.548**
		(0.2317)	(0.2050)	(0.1988)	(0.2084)
RDdummy	0.937**	0.759**		1.088	
	(0.1748)	(0.2364)		(0.2172)	
Other_Innov_Exp	0				
	(0.0000)				
number of high skill technical staff			13.227**	14.229**	18.902**
			(4.0450)	(4.1850)	(3.9962)
foreign_cap1	1.41**	2.369**	2.341**	2.367**	2.367**
	(0.2474)	(0.2807)	(0.2780)	(0.2801)	(0.2805)
group_dep1	0.994**	0.918**	0.795*	1.024**	1.06**
	(0.3058)	(0.3302)	(0.3272)	(0.3272)	(0.3272)
client_dep1	-0.309	-0.419	-0.366	-0.331	-0.183
	(0.3768)	(0.4068)	(0.4022)	(0.4037)	(0.4067)
independent or group			-0.111	-0.029	-0.023
			(0.2265)	(0.2231)	(0.2300)

Constant	-9.452**	-12.472**	-10.487**	-10.439**	-10.425**
	(0.7448)	(0.9488)	(1.0045)	(0.9857)	(1.0121)
Observations	13594	7996	7996	7996	7996
Number of group	7127	4530	4530	4530	4530

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

Appendix 5 Sophistication estimates

Table A5.1 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 A5.2 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 A5.3 Multinomial Logit estimates relatedness (Leontief input use) and technological content (OECD) – exports

	(1) Related No upgrade	Related Upgrade	Unrelated No upgrade	Unrelated Upgrade	(2) 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 A5.4 Multinomial Logit estimates relatedness (HS-2 difference) and technological content (OECD) – exports

	(1) Related No upgrade	Related Upgrade	Unrelated No upgrade	Unrelated Upgrade	(2) 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 A5.5 Multinomial Logit estimates relatedness (Leontief input use) and technological content (OECD) – core production

	(1) Related No upgrade	Related Upgrade	Unrelated No upgrade	Unrelated Upgrade	(2) 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 A5.6 Multinomial Logit estimates relatedness (HS-2 differences) and technological content (OECD) – core production

	(1) Related No upgrade	Related Upgrade	Unrelated No upgrade	Unrelated Upgrade	(2) 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 A5.7 Multinomial Logit relatedness (input use) and sophistication (PRODY) exports

	(1) Related No upgrade	Related Upgrade	Unrelated No upgrade	Unrelated Upgrade	(2) 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 A5.8 Multinomial Logit relatedness (HS-2 difference) and sophistication (PRODY) exports

	(1) Related No upgrade	Related Upgrade	Unrelated No upgrade	Unrelated Upgrade	(2) 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 A5.9 Multinomial Logit relatedness (input use) and sophistication (PRODY) core production activity

	(1) Related No upgrade	Related Upgrade	Unrelated No upgrade	Unrelated Upgrade	(2) 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 A5.10 Multinomial Logit relatedness (input use) and sophistication (PRODY) – vis-a vis core production activity

	(1) Related No upgrade	Related Upgrade	Unrelated No upgrade	Unrelated Upgrade	(2) 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.

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

Albornoz, F.; Calvo-Pardo, H.; Corcos, G. and Ornelas, E. (2010) *Sequential Exporting*, Centre for Economic Performance Discussion Paper 974, London: London School of Economics (LSE)

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

Amurgo-Pacheco, A. and Pierola, M.D. (2008) *Patterns of Export Diversification in Developing Countries : Intensive and Extensive Margins,* Policy Research Working Paper 4473, International Trade Department, Washington: The World Bank

Arkolakis, C. and Muendler, M. (2009) 'The Extensive Margin of Exporting Goods: A Firm-Level Analysis', Yale University mimeo, 44

Aw, B.Y.; Roberts, M.J. and Yi Xu, D. (2009) *R&D Investment, Exporting, and Productivity Dynamics*, NBER Working Paper 14670, Cambridge MA: National Bureau of Economic Research, Inc.

Baldwin, R. and Harrigan, J. (2007) *Zeros, Quality and Space: Trade Theory and Trade Evidence*, NBER Working Paper 13214, Cambridge MA: National Bureau of Economic Research, Inc.

Becker, S.O. and Egger, P. (2007) *Endogenous Product Versus Process Innovation and a Firm's Propensity to Export*, CESifo Working Paper Series 1906, Munich: University of Warwick and Institute for Economic Research (IFO)

Bekes, G., Kleinert, J. and Toubal, F. (2009) 'Spillovers from Multinationals to Heterogeneous Domestic Firms: Evidence from Hungary', *World Economy* 32.10: 1408–33

Bernard, A.B.; Jensen, J.B.; Redding, S.J. and Schott, P.K. (2007) 'Firms in International Trade', *Journal of Economic Perspectives* 21.3: 105–30

Bernard, A.B.; Redding, S.J. and Schott, P.K. (2010) 'Multiple-Product Firms and Product Switching', *American Economic Review* 100.1: 70–97

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

—— (2006b) 'Product Differentiation and Duration of Us Import Trade', *Journal of International Economics* 70.2: 339–58

Van Beveren, I. and Vandenbussche, H. (2009) *Product and Process Innovation and the Decision to Export: Firm-Level Evidence for Belgium,* LICOS Discussion Paper 24709*,* Leuven: LICOS Centre for Institutions and Economic Performance

Broda, C. and Weinstein, D.E. (2010) 'Product Creation and Destruction: Evidence and Price Implications', *American Economic Review* 100: 691–723

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

Cassiman, B. and Martinez-Ros, E. (2007) 'Product Innovation and Exports. Evidence from Spanish Manufacturing', mimeo*,* Madrid: IESE Business School

Costantini, J. and Melitz, M. (2007) 'The Dynamics of Firm-Level Adjustment to Trade Liberalization', mimeo, Harvard: Harvard University Department of Economics

Damijan, J.P.; Kostevc, C. and Polanec, S. (2010) 'From Innovation to Exporting or Vice Versa?', *The World Economy* 33.3: 374–98

Dennis, A. and Shepherd, B. (2007) *Trade Costs, Barriers to Entry, and Export Diversification in Developing Countries*, Policy Research Working Paper Series 4368, Washington: The World Bank

Egan, M.L. and Mody, A. (1992) 'Buyer-Seller Links in Export Development', *World Development* 20.3: 321–34

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

Freund, C. and Pierola, M.D. (2009) *Export Entrepreneurs: Evidence from Peru*, Policy Research Working Paper 5407, Development Research Group, Washington: The World Bank

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

Greenaway, D.; Sousa, N. and Wakelin, K. (2004) 'Do Domestic Firms Learn to Export from Multinationals?', *European Journal of Political Economy* 20.4: 1027–43

Haddad, M.E.; Lim, J.J. and Saborowski, C. (2010) *Trade Openness Reduces Growth Volatility When Countries are Well Diversified*, WPS 522, World Bank Policy research Working Paper, Washington: The World Bank

Hall, B.H. and Mairesse, J. (2006) 'Empirical Studies of Innovation in the Knowledge-driven Economy', *Economics of Innovation and New Technology, Taylor and Francis Journals*, 15.4–5: 289–99

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

Hausmann, R.; Hwang, J. and Rodrik, D. (2007) 'What you Export Matters', *Journal of Economic Growth* 12.1: 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 R.S. Newfarmer, W. Shaw and P. Walkenhorst (eds), *Breaking into New Markets: Emerging Lessons for Export Diversification,* Washington, DC: The World Bank

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

Hidalgo, C.A.; Klinger, B.; Barabasi, A.L. and Hausmann, R. (2007) 'The Product Space Conditions the Development of Nations', *Science* 317.5837: 482–87

Hobday, M. and Rush, H. (2007) 'Upgrading the Technological Capabilities of Foreign Transnational Subsidiaries in Developing Countries: The Case of Electronics in Thailand', *Research Policy* 36.9: 1335–56

Hummels and Klenow (2005) 'The Variety and Quality of a Nation's Exports', *American Economic Review* 95.3: 704–23

Imbs, J. and Wacziarg, R. (2003) 'Stages of Diversification', *American Economic Review* 93.1: 63–86

Lacovone, L. and Javorcik, B.S. (2010) 'Multi-Product Exporters: Product Churning, Uncertainty and Export Discoveries', *The Economic Journal,* 120.544: 481–99

Javorcik, B.S. and Spatareanu, M. (2009) 'Tough Love: Do Czech Suppliers Learn from Their Relationships with Multinationals?', *Scandinavian Journal of Economics* 111.4: 811–33

Klinger, B. (2007) *Uncertainty in the Search for New Exports*, CID Working Paper 16, Cambridge MA: Centre for International Development at Harvard University

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

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

Lederman, D. (2009) *The Business of Product Innovation : International Empirical Evidence*, Policy Research Working Paper 4840, Development Research Group, Trade Team, Latin American and the Caribbean Region Office of the Chief Economist, Washington: The World Bank

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

López, R.A. (2009) 'Do Firms Increase Productivity in Order to Become Exporters?', *Oxford Bulletin of Economics and Statistics* 71.5: 621–42

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

Mayer, T. and Ottaviano, G. (2008) 'The Happy Few: The Internationalisation of European Firms', *Intereconomics* 43.3: 135–48

Melitz, M.J. (2003) 'The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity', *Econometrica* 71.6: 1695–725

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

Ruhl, K.J. and Willis, J.L. (2009) 'New Exporter Dynamics', mimeo, New York: NYU Stern School of Business

Swenson, D.L. (2008) 'Multinationals and the Creation of Chinese Trade Linkages', *Canadian Journal of Economics/Revue Canadienne D'économique* 41.2: 596–618

Teece, D.J.; Rumelt, R.; Dosi, G. and Winter, S. (1994) 'Understanding Corporate Coherence: Theory and Evidence', *Journal of Economic Behavior and 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