CTEI-2017-15

CTEI WORKING PAPERS

GLOBAL VALUE CHAINS AND PRODUCT SOPHISTICATION – AN EMPIRICAL INVESTIGATION OF INDIAN FIRMS

Karishma Banga[a]

Abstract

This paper analyses the impact of GVC linkages on product upgrading using the Indian firm-level dataset Prowess, and methodologies of System-GMM and Propensity score matching. Defining product upgrading as a movement towards more sophisticated products, we use Haussmann's product sophistication index to calculate a sales-weighted average sophistication level of Indian manufacturing firms in the period 2000/01-2014/15. The first section of the paper empirically investigates whether GVC firms produce more sophisticated goods than non-GVC firms. In the next section, we take the sub-sample of GVC firms to analyse the impact of depth of GVC integration on firm-level sophistication. Lastly, the study draws on Gereffi's (2005)) governance framework and quantifies different types of governance structures to analyse how power asymmetries can impact the sophistication levels of developing country supplier firms. The study finds that the on average, Indian firms participating in GVCs produce more sophisticated goods than non-GVC firms, and increasing GVC integration significantly and positively impacts firm level sophistication. These results lend empirical support to the learning by importing and exporting hypothesis. However, our results also show that firms that link into Captive global value chains produce significantly less sophisticated goods than firms linking into Relational chains.

[a] This paper was drafted whilst I was a Junior Visiting Fellow at the Graduate Institute's Centre for Trade and Economic Integration in Geneva, and a PhD student at the University of Manchester, UK. I would like to thank Prof. Kunal Sen, Dr. Osman Ouattara and Dr. Theresa Carpenter for their support and feedback.

Centre for Trade and Economic Integration (CTEI)

The Centre for Trade and Economic Integration fosters world-class multidisciplinary scholarship aimed at developing solutions to problems facing the international trade system and economic integration more generally. It works in association with public sector and private sector actors, giving special prominence to Geneva-based International Organisations such as the WTO and UNCTAD. The Centre also bridges gaps between the scholarly and policymaking communities through outreach and training activities in Geneva.

www.graduateinstitute.ch/ctei

1. Introduction

A recent stream of economic literature finds that producing more sophisticated goods can serve as an engine of economic growth. Drawing from the theoretical approach in Hausmann and Rodrik (2003), Hausmann et al. (2007) show that the production of highly sophisticated goods shifts out the technological frontier of a country and improves the growth performance, helping the country climb up the export value chain (Hausman et al., 2007). This suggests the importance of taking the discussion beyond the concept of 'comparative advantage' and focusing on understanding what drives 'product upgrading' i.e. the transition from production of low quality goods to high quality goods. While these factors have been previously explored in the economics literature at the macro level, there are very few empirical studies takings firms as the unit of analysis in examining how a move towards more sophisticated products can be made.

One route to product upgrading that has been identified in the existing literature is participation in Global Value Chains (GVCs). It has been well documented that production in today's world is increasingly complex, with intermediate goods and supporting activities being sourced globally in order to increase cost efficiency. Post the middle of the 20th century, there has been rising trade in intermediates, growing liberalization of trade and investments, reduction in transportation and logistics costs and rapid advancements in information and communication technologies (ICTs). These trends have led to a greater interconnection between countries and complex production arrangements, known as GVCs. Trading in these GVCs involves simultaneously importing and exporting, which can generate complementarities that positively impact the sophistication of products produced by developing country suppliers. The import of intermediates can serve as important sources for knowledge, technical know-how and channels of technology transfer, while exporting in GVCs exposes developing country firms to global competitive pressures and increasing incentives for innovating.

However, the impact of GVCs is not uniform across all participating firms and benefits derived are conditional on how production is being governed. The different level of power asymmetries between supplier firms and Lead firms (buyer firms in GVCs), as well as the different learning mechanisms adopted by the supplier firms, can significantly shape opportunities for upgrading. Moreover, while GVC participation is often regarded as a vehicle of growth for developing country firms, several case studies have revealed that gains from linking are not automatic. Internal efforts by the participating firms in their ability to acquire, absorb and utilise the knowledge gained, as well as their investment in building technological capability, are also factors when it comes to product upgrading.

Further, it is important to acknowledge that firms differ in their strategies of profit generation. Once a firm has joined a GVC, increasingly intense competition from global competitors and escalating demand from Northern lead firms and branded manufactures can threaten the survival of the firm. To ensure its survival, firms may look to increase profitability, not by moving to better goods, but by expanding economies of scale. This can actually lead to product 'downgrading' whereby firms are making profit by selling low sophisticated goods in large amounts (Gibbon & Ponte, 2005). A good example of this would be South African Wine industry which achieved a 'better deal' by downgrading from wine to production of low quality grapes (Ponte and Ewert, 2009).

The impact of GVCs on product-upgrading is therefore not straight-forward and presents an interesting empirical challenge. In this study, we take product upgrading as move towards more sophisticated goods and examine the impact of GVC participation on product upgrading of Indian firms. The literature empirically examining firm-level product upgrading has been scarce, particularly for India. In the studies that do exist, product upgrading has mainly been captured as an increase in unit values of products (Hallak , 2006 ;Manova and Zhang , 2012 ; Harding and Javorcik , 2012 ; Bas and Strauss-Kahn , 2015). In this study, we calculate a sales weighted average sophistication level of firms, using Hausmans's product sophistication index (2007). This index captures the income level associated with each product, with more sophisticated products being those that are produced by more developed economies.

In analysing the impact global networks on product upgrading, India in particular presents an interesting case because it has been identified by many studies as an outlier in GVCs (Gupta, 2015, Goldar et al., 2017). Despite being one of the largest and fastest growing markets located in direct proximity to 'Factory Asia' (Baldwin 2006), India is observed to have very low GVC linkages compared to other Asian economies, especially China. This could be due to its stagnant growth in manufacturing sector, low ability to attract FDI in manufacturing, greater focus on domestic markets or low-value exports (Hoda and Rai, 2014). Moreover, analysing Indian linkages, it is found that India mainly participates in the global economy by backward participation i.e. by sourcing intermediates from abroad. Its foreign value added share in gross exports has increased from 12.6% in 1995 to 36.2% in 2011 while its domestic value added share in exports of intermediate goods and services has decreased from 43% to 37% (Banga, 2016). Further, despite having comparative advantages in labour intensive production, India's commodity composition of exports has shifted more towards capital and skill intensive products post-reforms (Veeramani, 2012). Therefore, India has managed to create strong linkages in some key capital intensive sector such as chemical, pharmaceutical and electrical equipment industries.

To our knowledge, while studies have previously examined such GVC integration of Indian industries, there has been no study till date identifying and examining GVC firms in India. The industry level studies that do exist ignore important heterogeneity in firm-level GVC participation in terms of; traders vs non-traders, type of traders; two way/ exporter/importer, SMEs vs large enterprises, domestic vs controlled enterprises, high capability firms vs low capability firm etc. To address this research gap, we conduct firm-level examination based on Indian trade data and follow Baldwin's (2014) methodology of identifying GVC firms as those that are simultaneously importing intermediates and exporting intermediate or final goods. We further examine depth of GVC integration using a firm-level estimate of vertical specialisation which measures the imported intermediates embodied in firm-level exports. To capture the nature of GVC participation, we construct novel measures from supplier level data to identify the different governance structures Indian firms are operating in.

For the purpose of this study, we use a unique firm-level panel database which contains information on annual financial performance of Indian firms and detailed data on the products manufactured by each firm. This panel is constructed for the financial years 2000/2001 to 2014/2015 using the Indian firm-level dataset Prowess dataset. For calculation of firm-level sophistication, we also use information on product-level sales in Prowess and GDP data from WITS (UN COMTRADE). The data indicates that more sophisticated industries such as pharmaceuticals, computer and electronics, rubber and plastic products, chemical and chemical products are also comparatively more integrated into GVCs. We also observe that GVC firms produce more sophisticated goods than non-GVC firms, and amongst GVC firms, captive firms are the least sophisticated.

In establishing a causal relationship between GVC participation and firm sophistication, the main econometric problems are that of reverse causality, omitted variables and simultaneous shocks. To deal with this endogeneity, we employ the IV strategy of System GMM. This also allows us to account for the dynamic nature of firm-level sophistication.

Our GMM results show that GVC firms, on average, have a higher product- sophistication level than non-GVC firms, rendering support to the hypothesis of learning by importing and exporting and generation of complementarities between the two activities. Next, using vertical specialisation as a proxy for GVC integration, we find that as the imported intermediates embodied in firm-level exports increase, firms' product sophistication rises. This suggests that firms which are more integrated into GVCs produce significantly more sophisticated and complex products. However, categorising GVC integration into different governance structures, we find evidence that firms participating in Captive chains produce produce significantly less sophisticated goods than firms in relational chains. These captive chains are characterised by high lead firm control over supplier firm's production and relatively high power asymmetries. This result renders empirical support to the theory that upgrading depends upon governance structures. We also find evidence that innovative capacity of firms positively and significantly impacts product upgrading, implying that efforts made by firms into research and development capability also matter for type of products manufactured.

The rest of the paper is organised as follows. Section 2 reviews the literature and the presents underlying theoretical considerations. Section 3 describes the data sources used and explains the construction of the main variables. Section 4 presents preliminary findings. Section 5 presents the methodology and results for the impact of GVC participation on firm sophistication. Sections 6 and 7 examine the impact magnitude of GVC integration and governance on firm sophistication respectively. Section 8 summarises and concludes the study.

2. Drivers of product upgrading

In analysing drivers of product upgrading, it has been identified in the literature that trade can affect firm-level product upgrading through three channels; exports, imports and two-way trade. Pioneering work on the micro-level analysis of the impact of international trade was undertaken by Bernard and Jensen (1995) and Aw and Hwang (1995) on export market participation. Studies since then have found that exporters are better performers than non-exporters in the same sectors, in terms of productivity, skills, and capital intensity. Under the 'learning by exporting hypothesis', even in the presence of self- selection, exporters can improve productivity because of access to foreign markets which brings in global competitive pressures and opportunities to exploit economies of scale. Firms with access to international markets have a higher probability to acquire new technology (Almeida and Fernandes, 2008) and have a higher incentive to invest in innovating. This can lead to improvements in product quality or a shift towards the production of more sophisticated goods. Using data from a randomised control trial, coupled with survey data, Atkin et al (2017) analyse learning by exporting for rug producers in Egypt. Their results find that exporting leads to large improvements in both product quality and productivity of the producer. They document knowledge flows between foreign buyers, intermediaries and producers and find evidence that learning occurs in part due to information flows.

Similarly, some of the existing studies suggest a 'learning-by-importing' process. For instance, in Grossman and Rossi-Hansberg's (2009) theoretical model, importing intermediate inputs can improve firm productivity by allowing access to foreign inputs and technology that are not domestically available. In studies like Romer (1990), Grossman and Helpman (1991), Kortum (1997), and Eaton and Kortum (1999, 2002), the use of imported intermediate goods also implicitly involves the use of embodied technology and knowledge. Further, intermediate imports grant importers access to better quality and more variety of foreign inputs, allowing them to focus their resources and specialize in activities in which they enjoy a comparative advantage. Empirical studies confirming the positive impact of imports on product scope and quality include Kugler and Verhoogen's study (2009) for Columbia and Goldberg, Khandelwal, Pavcnik, and Topalova's study (2010) for India. Lin and Lin (2010) study the case of Taiwan and find that both imports and FDI have a positive impact on product innovation but find no evidence of reverse causality. Goldberg et al. (2010b) find a positive relationship between access to new foreign intermediate inputs and the introduction of new products by domestic firms. Examining the import effect on product quality, Khandelwal and Amiti (2009) use highly disaggregated product level export data from 56 countries to the US and find that a more liberalised trading environment can cause changes on product patterns of firms, enabling them to produce higher quality goods. The authors find that lower tariffs are associated with quality upgrading for products close to the world quality frontier, whereas lower tariffs discourage quality upgrading for products distant from the frontier. Also analysing the impact of trade liberalisation, Nguyen (2016) uses Hausmann's PRODY index to calculate product sophistication of Vietnamese industries and finds that tariff reductions have a positive impact on the sophistication level on export basket of industries, with a higher impact for non manufacturing sectors.

There are very few studies analysing two-way trade and the studies that do exist advocate that two-way traders may benefit from the complementarities between activities of exporting and importing. This is evidenced in Kasahara and Lapham's (2013) study, on a panel of Chilean manufacturing industries, which found that by engaging in both trading activities a firm can reduce per period fixed and sunk costs. Veugelers et al. (2013) examine GVCs in Europe, defining GVC-involved firms as those that simultaneously import components, maintain production activities located abroad and export their goods. The authors find that while only a few firms are intensively engaged in GVCs, these firms are significantly more likely to introduce new products, new processes solely or in combination. Also defining two- way traders as GVC firms, Baldwin and Yan (2014) show that after controlling for this self-selection, firms that joined a GVC became more productive, and their performance continued to improve. Conversely, firms that stopped participating in a GVC suffered a similar-sized loss in productivity. Lo Turco and Maggioni (2014) find that the joint involvement of a firm in importing and exporting positively impacts product scope and new product introduction, which may contribute towards a more sophisticated product basket. Seker (2012) confirms that two-way traders are the most innovative, in terms of product and process innovation than any other group of firms.

More recently, studies have attempted to analyse foreign direct investment as a driver of product sophistication. Eck and Hubert (2016) estimate firm-level sophistication for Indian firms using Hausmann's product-sophistication index (PRODY), and find that vertical backward FDI spill-overs from downstream MNEs to upstream local Indian suppliers have a positive and significant impact on the supplier's firm sophistication. Also focusing on Indian firms, Beyer (2015) finds that foreign owned firms are more active in product innovation and produce on average approximately one product more than had they stayed without foreign ownership. Using data on product-level exports of Turkish firms, Javorcik et al. (2017) measure firms' product complexity using indicators for both diversification and ubiquity, as suggested in Hausmann and Hidalgo (2009). Their analysis suggests that Turkish firms in sectors and regions more likely to supply foreign affiliates tend to introduce more sophisticated products. Evidence therefore seems to suggest that investments made by Multi-national enterprises into local suppliers can lead to transfer of knowledge from the MNE to the foreign affiliates (Arnold and Javorcik, 2009) and also increases the probability of the affiliate itself introducing new products (Brambilla, 2009; Guadalupe et al., 2012). Furthermore, MNEs may provide local producers with better inputs, technical know-how and support, lowering cost of production and enabling product innovation.

Both trade and foreign direct investment are thus found to be important drivers of product upgrading. While, integration through only trade (market integration) and through FDI (vertical integration) form the two extreme ends of the make-or-buy literature, firms integrating into the global economy with a combination of the two are identified as firms operating in 'networks'. The main contribution of Gereffi, Humphrey and Sturgeon (2005) is to differentiate the 'networks' in the economics literature into three types of inter-firm linkages in GVCs; Modular, Relational and Captive or Quasi-hierarchy. According to the authors, the type of chain that arises depends on a combination of three key factors. The first is 'complexity of transactions', or the complexity of the information and knowledge transferred from the lead firm to the supplier firm in order to communicate product and process specifications. The second is 'codifiability of transactions' or the ease with which information can be codified into contracts and standards to reduce the complexity of information. The third is 'supplier capability' –the supplier's ability to deal with the complex and codified information he receives from the lead firm. Power asymmetries between these lead firms and their suppliers, across different forms of governance structures, is seen as an important determinant of the ability of supplier firms to learn and upgrade in the value chain.

While there is no explicit conceptual or theoretical framework of the learning mechanisms for firms in GVCs, Pietrobelli and Rabellotti (2011) have identified learning mechanisms in different type of governance structures, based on an extensive literature review. These learning categories have been derived from direct field observations across different industries such as electronics (Sturgeon, 2002), apparel (Gereffi, 1999), horticulture (Dolan and Humphrey (2000; 2004), and bicycles (Galvin and Morkel) and are summarised in table 1 below.

Table 1: Learning mechanisms for supplier firms across GVC governance

Governance Type	**Key Factors**		**Key Characteristics**	**Learning mechanisms for GVC supplier firm to upgrade products**
MARKET	Complexity Codifiability competence	LOW HIGH HIGH	-Low lead firm control -Buyers and suppliers do not collaborate -Standardized products	-Knowledge spillovers -Imitation
MODULAR	Complexity Codifiability competence	HIGH HIGH HIGH	- Customised products according to buyers' specifications. -Suppliers provide 'turn-key services', going beyond the mere activity of assembly.	-Pressure to accomplish international standards. -Transfer of knowledge embodied in standards, codes, technical definitions.
RELATIONAL	Complexity Codifiability competence	HIGH LOW HIGH	- Mutual dependence between buyers and sellers. - Linkages develop over a long period of time; high switching costs -Complementary competences between buyer and supplier.	- Face-to-face interactions. -The technology and technical expertise transferred *via* the products manufactured for other multinationals and/or in the production for own designed and branded products.
CAPTIVE	Complexity Codifiability competence	HIGH HIGH LOW	- Significant Lead firms control - Risk of suppler failure is high; strict monitoring -Low-skilled suppliers dependent on dominant lead firms; supplier's switching costs are high. - Production of standard consumer goods or labour intensive goods (such as textiles, clothing etc.)	- Knowledge transfer from lead firms but confined to a narrow range of tasks. -Risk of lock-ins because lead firms do not sustain the development of core capabilities.
HIERARCHY	Complexity Codifiability competence	HIGH LOW LOW	-Vertically integrated buyer, who internalises all functions of the seller. - Capable suppliers cannot be found without incurring major transaction costs -Asset specificity is high and management hierarchy oversees subsidiary activities in such linkages.	-Imitation -Turnover of skilled managers and workers -Training by foreign leader/owner -Knowledge spillovers

Source: Adapted from Gereffi et.al (2005) and Pietrobelli and Rabellotti (2011)

Some existing empirical studies in the GVC literature have tried to quantify Gereffi's (2005) governance structures by exploiting survey-based information on inter-firm relationships. For example, Brancati et al. (2016) use the MET survey on Italian firms to collect data on Italian firms' affiliations to corporate groups, presence of strong informal ties with buyers and firm's involvement in the conception of the final product. Similarly, Saloia and Zanfei (2009) use firm-level data for Thailand, collected from the Productivity and Investment Climate Survey (PICS), and define governance based on information regarding; sales made by supplier to suit buyer specification, flow of product information, workers or training from buyer to supplier and imposition of quality standards. More recently, some studies have also used firm-level trade transactional level data to quantify governance. Using such data for 439 of China's largest exporters, Dallas (2015) specifies four indicators that are likely to differ across governance structures; Export Specialisation, Transactional Stability, Industry Stability and Export Entry. Based on different combinations of these indicators, Dallas (2015) classifies large firms under different governance types. Focusing on the automotive industry and using transaction-level data on first and lower-tier supply contracts, Schmitt and Biesebroeck (2017) identify governance modes empirically from the impact on observable market outcomes.

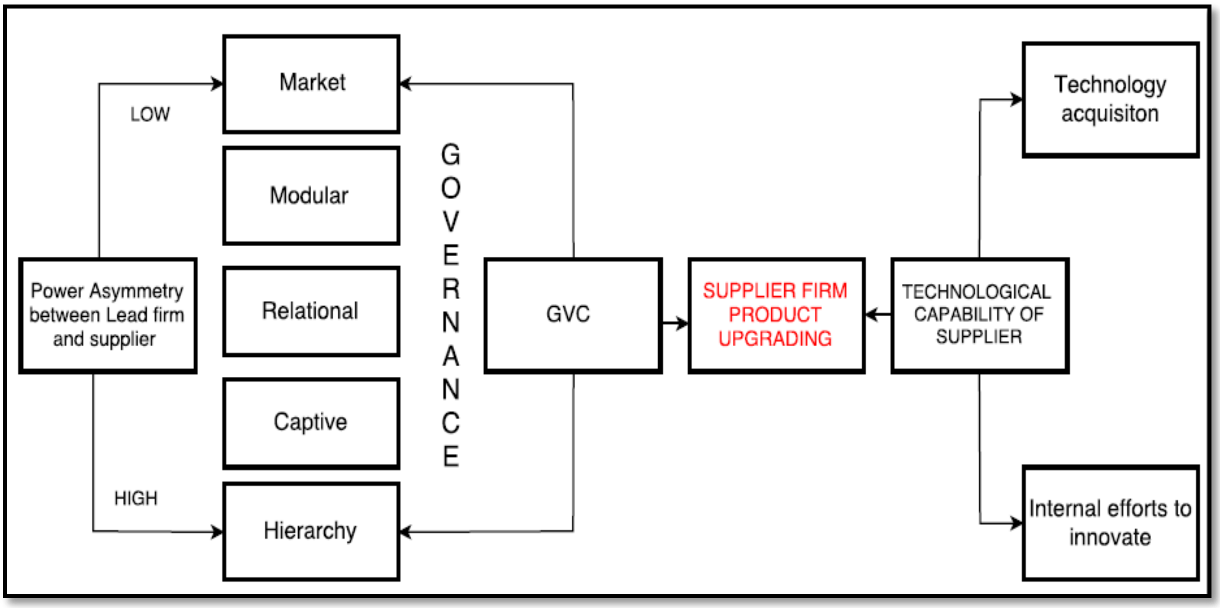
While the theory of GVC governance does provide the most comprehensive framework for analysing different governing relationships within value chains, it cannot account for heterogeneity across firms in the same governance structure. Many studies point out that the understanding of upgrading in the GVC literature is relatively weak in its conceptualisation of how technological progress, learning and capability are acquired by firms (Morrison et al.,2008; Sato & Fujita, 2009; Pietrobelli & Rabellotti, 2011; Nadvi, 2011). While external factors such as trade and foreign direct investment can play an important role in shaping upgrading opportunities for developing country firms, participating in GVCs does not lead to automatic upgrading. It requires efforts in building firm capability.

This drawback of the GVC approach to upgrading is dealt with in the technological capability literature. Research shows that local firms in fact have heterogeneous technological capabilities and very distinctive firm-specific learning strategies (Nelson and Winter, 1982; Bell and Pavitt, 1993; Lall, 1992 and 2001; Giuliani and Bell, 2005), which may prevent them from taking advantage of the foreign knowledge in the same way. Lall (2000) and Bell and Pavitt (1995) present different TC levels such as: basic or simple routine capabilities (experience based); intermediate or adaptive capabilities (search based); advanced or innovative capabilities (research based). Firms can therefore actively engage in the learning process and acquire explicit technological knowledge through learning by doing (producing goods), or learn passively by using capital goods. If knowledge is tacit, passive learning may not be enough to acquire technical knowledge (Bell & Albu, 1999). Some firms may acquire TC from suppliers of machinery or production inputs, while others may need R&D efforts (Bell & Pavitt, 1993). While there exists a rich literature on technological capabilities (Lall, 1992; Bell & Pavitt, 1993; Ernst & Kim, 2001) and on local and national innovation systems

(Nelson, 1993; Cooke et al., 1997; Edquist, 1997; Mytelka, 2000; Lundvall, 2007), there are very few studies examining them in context of GVCs. More recently Morrison, Pietrobelli and Rabellotti (2008) and Pietrobelli and Rabellotti (2011) attempt at examining inter- firm differences in the same governance structure by approaching upgrading from the perspective of 'technological capabilities' and 'innovation systems'.

The above discussion highlights the importance of analysing upgrading through a GVC-TC framework (figure 1) that allows us to account for both external and internal drivers that shape the upgrading opportunities for supplier firms. While external factors include GVC participation and operating in different governance structures, internal factors refer to endogenous upgrading efforts made by the supplier firms.

Figure 1: Supplier firm upgrading in the GVC-TC framework



Source: Author (2017)

3. Data and Construction of main variables

The study primarily uses Prowess, a database of the financial performance of Indian companies, created by Centre for Monitoring Indian Economy Pvt Ltd (CMIE). Information in Prowess is compiled from firm-level Annual reports and balance sheets of listed companies. The total income of all companies in Prowess covers about 80% of India's GDP and as for international trade, Prowess cover about 50% of India's exports and nearly 60% of imports[2]. Prowess classifies firms according to NIC industries at the five-digit level (2008 classification). We restrict ourselves to manufacturing firms[3] and collect firm-level data on identification indicators, total sales,

[2] These statistics are for the year 2013-2014.
[3] The two-digit NIC industries considered in the study are given in the appendix (A.1)

exports, imports, wages and salaries, R&D expenses, royalty, technical know-how and licences fee expenditures etc. We compile a firm-level panel dataset for 8,539 manufacturing firms for the financial years 2000/2001 to 2014/2015[4].

By the 1956 Indian Companies act, Indian firms are also required to report information on product-level sales, capacities and quantities produced[5]. For the purpose of this study, we collect product-level data on 'sales as a share of gross sale of the firm' and we observe that on an avergae, 91% of sales of the firm are covered by product sales. Products considered in this study can be classified into the following sectors: Minerals, Fats, oils and derived products, food products, beverages, tobacco products, textiles, leather and leather products, wood and wood products, pulp and paper products, chemicals and chemical products, plastics and rubbers, non-metallic minerals, base metals, machinery, transport equipment and misc. manufactured articles (optical instruments, furniture, toys., clocks etc.). Since this study focuses only on manufacturing products produced by firms, products classified by Prowess as animal products, agricultural products, services, construction and irrigation are not included in the analysis. As part of data cleaning, we remove firms-year observations with missing or negative value for sales and observations with export-intensity greater than 100. After cleaning the data, we are left with panel of 69,768 firm-year observations. We observe that on an average, a firm is present in the dataset for 11 years out of 15 years[6].

3.1 Construction on Firm-level Product sophistication

It is important to note that Indian firms are not required to report product information using any particular classification or governing rule. Therefore, CMIE uses its own internal product classification, loosely based on NIC and HS schedule. In order to calculate sophistication level of each product in Prowess, we need to first match Prowess products to HS products. In order to do this, we assign correspondences between the two by hand, exploiting the fact that both Prowess's classification and HS classification are closely related to ISIC classification[7]. We are able to match around 80% of products in Prowess to four-digit HS products (1996 classification). While matching to six-digit would be ideal, Indian firms do not report detailed product names and therefore are best matched to HS at four-digit [8].

[4] We choose 2000-2001 as the start year since firms were not required to report data on foreign ownership prior to this period.

7. Information on product-level exports is not reported.

[6] Prowess data does not allow one to analyse entry and exit. However, the database has been extensively used in a variety of economics and business literature, and there is no evidence that there is any systematic way in which firm attrition takes place. The following reasons increase our confidence in non-random attrition; (1) data in Prowess mostly comes from medium to large firms, therefore missing data for a firm is most likely due to the fact that the firm has not reported the data rather than it having exited the industry. (2) entry of a firm into Prowess database does not mean that a new firm was formed at the time of the entry. It merely means that Prowess received information for the first time about the firm. (3) Prowess does not drop any firm from its database even if it exits. (4) Finally, we work with an unbalanced panel in which sample size varies from year to year, with only data availability and purging of outliers guiding our sample selection.

[7] This matching process is explained in detail in the appendix (A.2)

8 Matching to six-digit would lead to loss of accuracy, and is therefore avoided.

Once prowess products are matched to HS products, we calculate sophistication level of each HS product using Hausman's Prody index (2007) [9]. It avoids normalisation or any sort of grouping of countries and measures the income per capita associated with each product, weighted by a variant of Balassa's RCA. It is calculated using; $PRODY^k = \sum_c \underbrace{\left(\frac{X_c^k / X_c^\bullet}{\sum_c (X_c^k / X_c^\bullet)}\right)}_{\varphi_c^k} Y_c$, where Y_c is the per capita income of country c, X_c^k denotes country c's export volume of good k, and $X_c^\bullet$ is the sum of exports of country c. The weights φ_c^k add up to one for each good. Hausmann et al. (2007) emphasize that the adjusted weight ensures "that country size does not distort the ranking of goods".

Data on GDP per capita (PPP) in constant US dollars (2011 USD) is collected for 267 countries/ regions from World Development Indicators [10]. Product (at 4-digit level) export data is collected (in terms of 1000 USD) from WITS in UNCOMTRADE. Since the exporting countries and comparative advantages can change over time, calculating product sophistication using different countries in different years can create biases in the indicator (Hausmann et al., 2007). Therefore, it is important to create sophistication indices using data from a consistent sample of countries that report trade data in the period 2001-2015. This means we don't include any country that has trade data missing for even one period in 2001-2015. We get consistent data for 113 countries reporting both export flows and GDP in the period considered[11].

Once the sophistication level of each four-digit product has been calculated, we construct firm level sophistication index PSI, which measures the average extent of product sophistication per firm. It is defined as the sales-weighted average sophistication level of all products, k = 1, ...K, that are produced by a firm i: $PSI_{it} = \sum_k \frac{Sales_{it}^k}{\sum_k^K Sales_{it}^k} \; SOPH^k$, where $SOPH^k$ is the product-specific sophistication of product k, $Sales_{it}^k$ is the sales of product k by firm i at time t. and $\sum_k^K Sales_{it}^k$ is the total sales of firm i across all products at time t. A rise in PSI index thus captures increase in firm sophistication due to 1) introduction of or shift towards new and more sophisticated products, such as moving from production of bicyles (HS 1706) to say automobiles (HS 1704) or 2) diversion of sales towards more sophisticated goods, incase of multi-product firms.

We also calculate seven other firm-sophistication indices, using different methodolgies of calculating product level sophistication. An overview of how each sophistication indictor is calculated, along with its ranking of products and a correlation matrix of the eight indicators is given in appendix A.4.

3.2 Construction of firm level GVC participation

9 This index is calculated using the PRODY command developed by Huber (2016).

10 Using GDP per capita PPP allows us to correct for differences across time (inflation) and across countries (deviations from PPP). This means that the product sophistication indicators can be compared across time and countries

11 The list of countries taken into analysis is given in the appendix (A.3).

In the recent literature, the phrase 'vertical specialisation' has been used to refer to firms that are importing intermediate goods, processing them and then exporting (Hummels, Ishii and Yi, 2001). In this study we follow the concept of vertical specialisation and adopt Baldwin and Yan's (2011) methodology of identifying GVC firms as firms that are simultaneously import intermediate goods and exporting intermediate or finished goods. Using this approach, we capture the sequential integration of production processes across countries. This definition is also consistent with the indicator used by OECD which describes GVC participation as intermediates produced in one country which are included in another country's exports. Therefore, we create a GVC dummy variable which is equal to one for all the firms that are simultaneously importing intermediate goods and exporting intermediate or final goods and services.

Next, to capture the actual magnitude of firm-level GVC integration, we take the sub-sample of GVC firms and measure vertical specialisation, as developed by Hummels, Ishii and Yi (2001). This index estimates the imported intermediates embodied in a firm's exports, thus accounting for both import and export intensities of a firm. Studies in the existing literature have differed on what they account for in imported intermediates. For robustness, we use different measures of imported intermediates and normalise them with total material inputs used in the production process to ensure that these indices are not influenced by firm size. We construct the following indices;

$$\mathrm{VS_1} = \frac{\text{import of; raw materials,stores and spares,capital goods,services}_{it}}{\text{expenditure on; raw material expenses,stores and spares ,gross fixed assets, services.}_{it}} * \left(\frac{\mathrm{Exports_{it}}}{\mathrm{Sales_{it}}}\right)$$

$$\mathrm{VS_2} = \frac{\text{import of; raw materials,stores and spares,capital goods}_{it}}{\text{expenditure on; raw material expenses,stores and spares ,gross fixed assets}_{it}} * \left(\frac{\mathrm{Exports_{it}}}{\mathrm{Sales_{it}}}\right)$$

$$\mathrm{VS_3} = \frac{\text{import of; raw materials,stores and spares}_{it}}{\text{expenditure on; raw material expenses,stores and spares}_{it}} * \left(\frac{\mathrm{Exports_{it}}}{\mathrm{Sales_{it}}}\right)$$

For trade policy purposes, we also measure import content of export (ICE), measured as;

$$\mathrm{ICE} = \left(\frac{\text{Imported intermediates}_{\mathrm{it}}}{\mathrm{exports_{it}}}\right) = \frac{\text{import of; raw materials,stores and spares,capital goods,services}_{it}}{\mathrm{exports_{it}}}$$

3.3 Construction of GVC governance dummies

While transactional trade data can be extremely useful in drawing insights into different types of buyer-supplier relationships, this type of data is not publicly available for Indian firms. Also, the firm-level databases for India, such as PROWESS, do not report destinations from which intermediates are sourced or exported to. With no buyer data available, it becomes impossible to examine inter-firm relations in India by using information on stability of relationships with foreign firms, involvement of foreign firms in the production process, number of buyers etc. [12]. Therefore, instead of characterising buyer-supplier relationships, we take a supplier-firm perspective to analyse what type of chain the supplier is more likely to be a part of. We do this by extracting firm-level data on skilled labour and supplier competence from Prowess.

Lakhani et al.'s (2013) employment systems framework for GVCs argues that skill & knowledge of employees in the supplier firm is strongly related to the nature of task requirements. For example, complex tasks will be associated with highly skilled workers while relatively simple tasks can be performed by less skilled workers. Similarly, if codifiability is low, i.e. it is difficult for the lead firm to codify product specifications into contracts or standards, then the supplier base will need to have a more skilled labour force in order de-codify the transactions and understand the order requirements. Thus, we propose that the share of skilled labour in a firm can be used as an inverse proxy for codifiability of transactions in the firm. By this reasoning, Relational and Hierarchical chains are expected to have a higher share of skilled labour compared to the remaining governance types (see table 1 for reference).

According to Gereffi's typology, the other key factor in determining governance is supplier competence or the ability of the supplier to receive the order and fulfil its requirements. To measure the level of capability in the supply base, we construct a firm-level ICT index that captures how well the supplier is equipped in receiving the order, communicating with the lead firm regarding the order, and its competence in fulfilling all the requirements as per the specifications of the order. Therefore, a firm-level ICT index can be used a proxy for supplier competence.

In line with Gereffi's typology, we propose that interactions of a firm's share of skilled labour and ICT index will determine what type of chain the supplier firm will participate in. These interactions are presented in Table 2.

Table 2: Quantifying governance

Governance	Market	Modular	Relational	Captive	Hierarchy
Average skill level	Low	Low-Mod	High	Low	High
ICT index	High	High	High	low	low

Source: Author (2017)

To capture the average skill level, we construct a 'high skill-level' dummy equal to one if the firm's share of managerial remuneration in total labour compensation is above the median

[12] While some of this information may extracted for a sample of Indian firms from World Bank Enterprise Survey, it is only available for the year 2013-2014.

value in its two-digit industry. For the ICT index, we use Principal Component Analysis on data related to software assets, technology assets and infrastructure assets (see appendix A.5.). We then construct a 'high ICT index' dummy equal to 1 for firms with ICT index greater than median value in two-digit industry. Combinations of these dummies can help us identify governance structures (table 3).

Table3: Construction of governance dummies

Market/ modular[13]	=1 if High skill level dummy=0 and High ICT dummy=1
Market	=1 if Market/ Modular=1 and High Market share=0
Modular	=1 if Market/ Modular=1 and High Market share=1
Relational	=1 if High skill level dummy=1 and High ICT dummy=1
Captive	=1 if High skill level dummy=0 and High ICT dummy=0
Hierarchy	=1 if High skill level dummy=1 and High ICT dummy=0

Source: Author (2017)

3.4 Construction of firm level technological capability

To capture the internal ability of a firm to acquire technology from external sources, we build a technology acquisition index or the operational TC index. This index is constructed using principal component analysis on royalty, technical know-how fee and import of capital goods. The construction of this index is explained in detail in the appendix (A.6). To capture the firm's ability to utilise the acquired technology and build its knowledge base, we calculate firm-level innovative capacity as the share of firm's R&D expenditure in its total sales.

4. Preliminary analysis

4.1 A portrait of Indian firms

Categorising firms according to trading activity (see figure 1), we observe that more than 40% of firms in the period 2000/01-2014/15 did not engage in importing or exporting activities and served only the domestic market. Around 28% of the firms are classified as one-way traders that either import or export and 33% of firms are GVC firms that import intermediate goods, process them and and exporting intermediate or final products. Although the share of GVC firms in our

[13] In order to further identify modular firms from market firms, market share of a firm can be used. GVC firms that are operating under market chains, operate mainly on price and have a large number of competitors. If their price increases, the lead firm can change to another supplier at a low switching cost. Given this, it is expected that the market-share of market based GVC firms will be lower than firms involved in modular chains.

sample is only 33%, the output produced by these firms account for about 74 % of total output produced (figure 2), signifying the importance of GVC firms in manufacturing.

Figure 1: Firms by trading activity

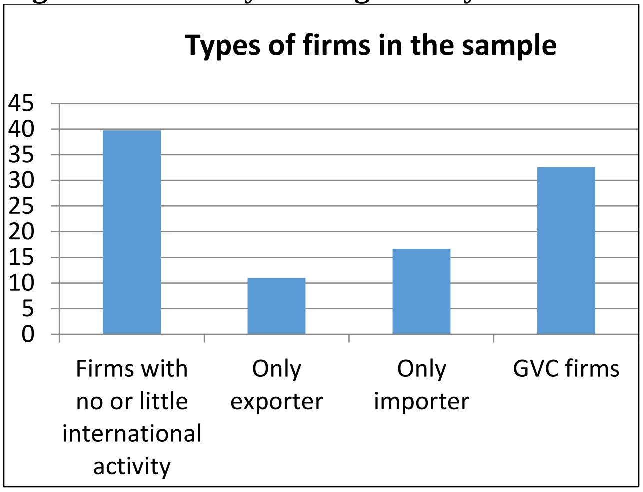
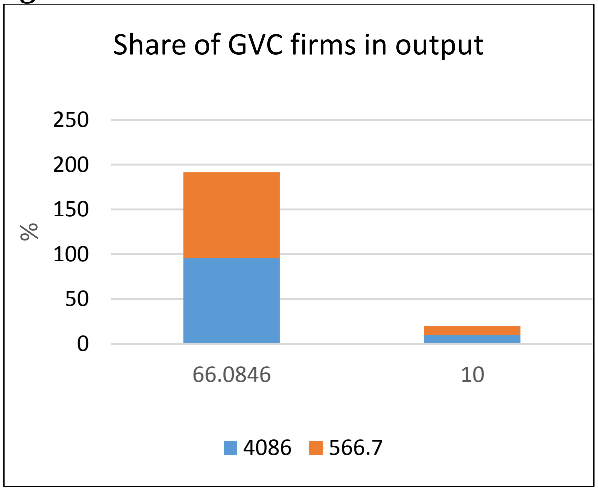


Figure 2: Share of GVC firms



Source: Author(2017)

Examining summary statistics (see A.7), we observe that only 7% of the firms in our sample are classified as foreign firms i.e. firms with more than 10% foreign equity. Using the number of people employed, 45% of firms are classified as big firms with more than 250 employees, 31% as medium sized firms with number of employees between 50 to 249 and 23 % as small sized firms with less than 50 people. Indian firms also operate with heterogeneous technological capability, captured by the varying R&D intensity (innovative TC) and Operational TC index (operational TC) of the firms. The share of R&D expenditure in sales of Indian firms ranges from around 1.5% to 96% but is quite low on average. The sales weighted firm-level sophistication index has been normalised to range between 0 to 100, with average sophistication index being roughly 27 for an Indian firm.

4.2 Portrait of GVC integration and firm sophistication over time

Between 2003/04 to 2007/08, India witnessed a period of significant growth acceleration. During this period, there was a sharp increase in the percentage of GVC firms and a steady increase in the average magnitude of firm-level GVC integration (See figures 3 and 4). The average level of firm- sophistication also increased (figure 5). During the period 2007/08-2009/10, when the financial crisis hit, the percentage of GVC firms in India fell by almost 5% and average GVC integration also declined. Both the level of firm-sophistication and its growth (figure 6) showed a similar trend, as observed by the sharp dips in the two series during the period of the crisis. Post the crisis, firm-level sophistication recovered and has been steadily increasing, with a particularly sharp rise observed in the last few years. The share of GVC firms in the sample also steadily increased after the crisis, implying that a larger number of firms started engaging in both importing and exporting activities.

However, post 2010, there seems to be a downward trend in the depth of GVC participation by Indian firms, mirroring the global deceleration of international trade since 2011. Possible explanations for declining participation rates in GVCs include; changes in the composition of final demand after the crisis, with shift in spending from more trade-intensive durable investments and consumer good to services which are mainly domestically produced, an increase in trade costs due to increased protection (Evenett and Fritz ,2015), rise in local production capabilities and substitution of imports by domestic goods (Kee and Tang, 2016), or simply realignment after overshooting of fragmentation, as suggested in Harms et al. (2012) and Baldwin and Venables (2013).

Figure 3: Share of GVC firms in total firms

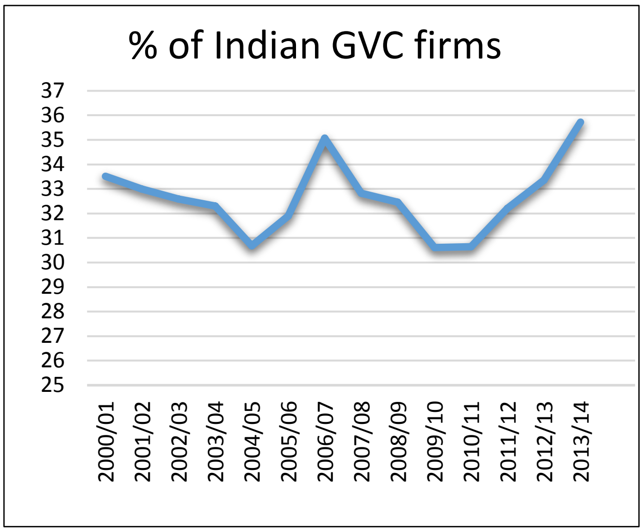
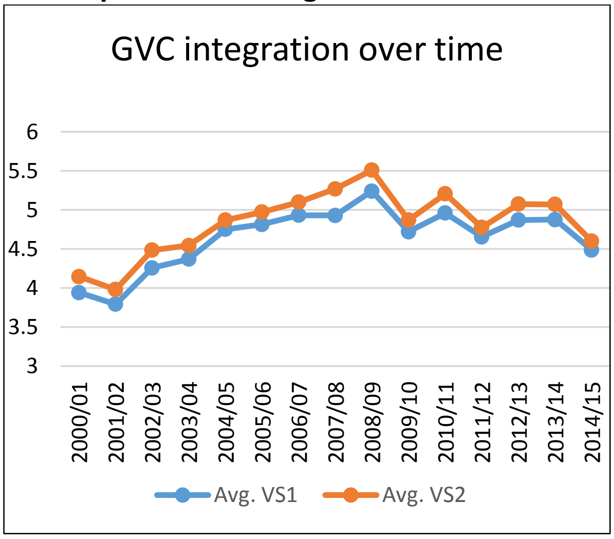


Figure 4: Depth of GVC integration



Source: Author (2017)

Figure 5: Changes in Firm sophistication

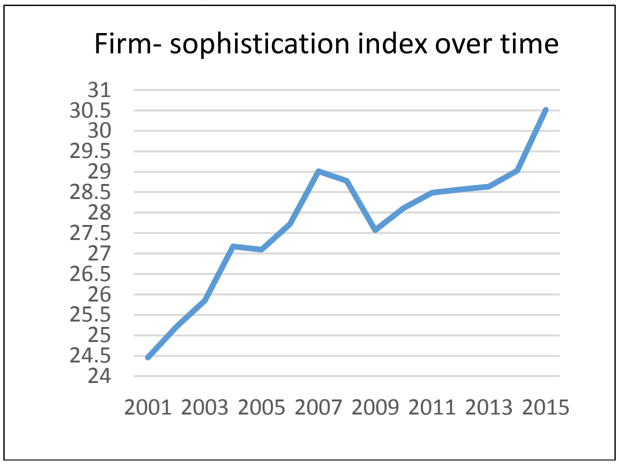
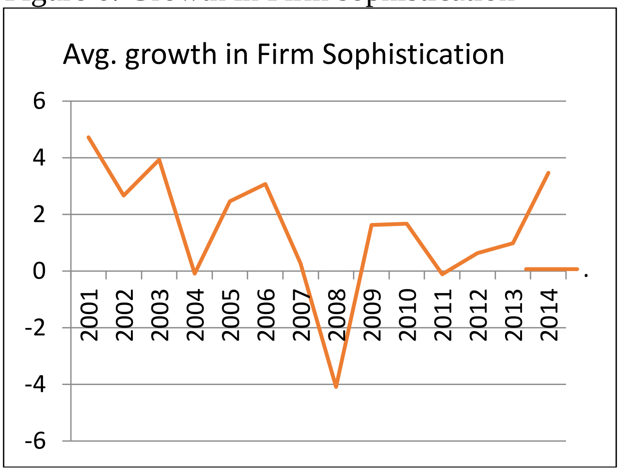


Figure 6: Growth in Firm sophistication

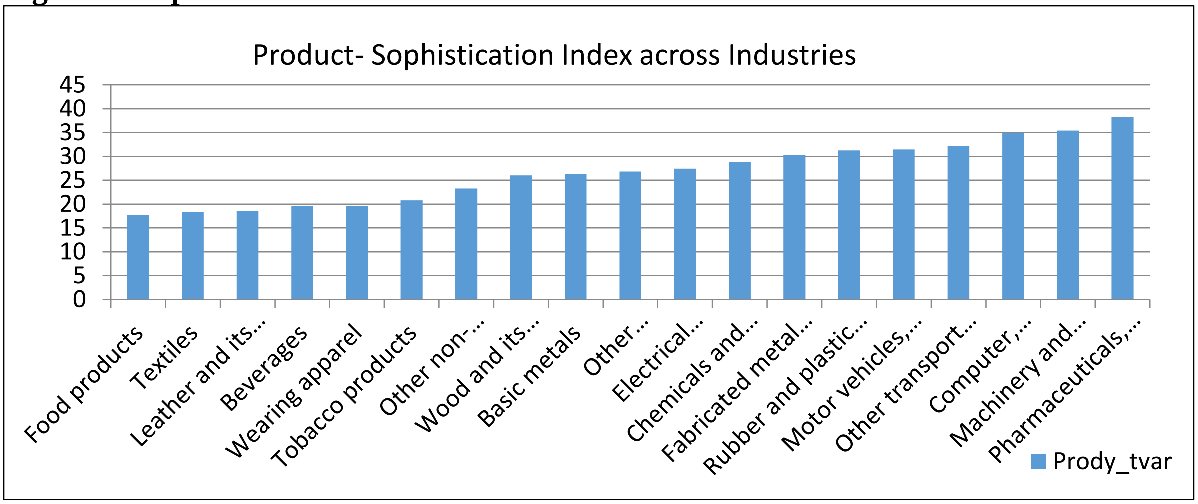


Source: Author(2017)

4.3 Portrait of GVC integration and firm sophistication across sectors

Analysing sophistication across industries, we find that on average, firms in pharmaceuticals, machinery and equipment, computer and electronics, motor vehicles, transport equipment, chemicals etc. are more sophisticated than industries of food products, beverages, leather, textiles etc. (figure 7)[14]. These capital-intensive industries are also observed to have a higher percentage (roughly above 40%) of GVC firms operating in them (figure 8). Using VS1 as a measure for depth of GVC integration, we find that firms classified under NIC 32- other manufacturing- have the highest average vertical specialisation (figure 9). These are firms engaged in manufacture of; jewellery, sports goods, games and toys, medical and dental instruments and musical instruments. We also observe that that the more sophisticated industries of pharmaceuticals, computer and electronics, rubber and plastic products, chemical and chemical products are also comparatively more integrated into GVCs. The only exceptions are industries of machinery and equipment, motor vehicles and trailers and other transport equipment. These Indian industries produce sophisticated goods, but are not found to have relatively high linkages in GVCs. While the automotive sector is very prominent in the GVC literature and globally exhibits a high import content of exports, India continues to impose relatively high tariffs in this sector. This may contribute to relatively lower imports of auto parts by Indian firms and consequent lower embodiment of intermediates in firm-level exports.

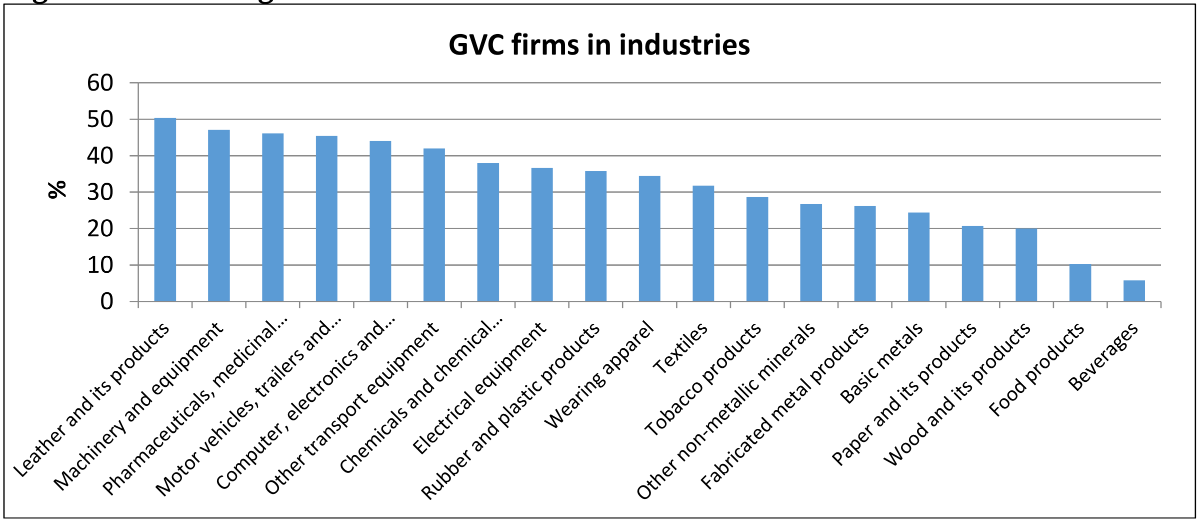
Figure 7: Sophistication level of Indian industries



Author(2017)

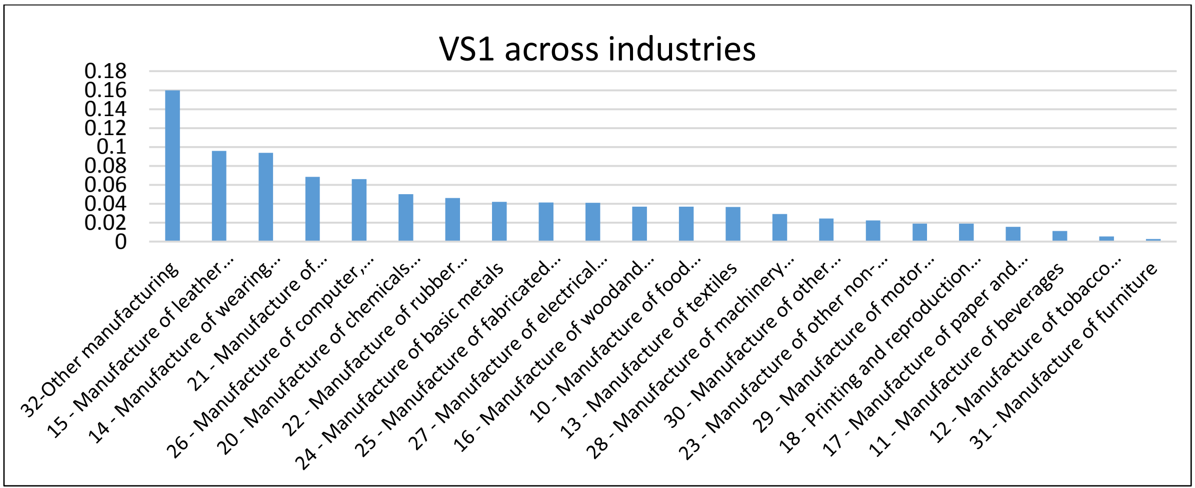
[14] These results are similar to Eck and Hubert(2014).

Figure 8: Percentage of GVC firms in different industries



Author(2017)

Figure 9: Depth of GVC participation in different industries



Author(2017)

4.4 Cross-firm analysis of GVC integration and firm sophistication

The industry level in section 4.3 analysis suggests that industries which produce more sophisticated goods are mostly those with a high percentage of firms operating in GVC. Moving on to firm-level analysis, we conduct difference-in-means test (table 4) to examine whether sophistication levels are significantly different across types of firms. The larger coefficient on GVC firm dummy and the significant p value of the t- test statsitic (0.000), indicates that GVC firms have signficantly higher sophistication than non-GVC firms. This results holds true across all eight indicators of product sophistication that we have calculated (see appendix A.8). On average, sophsitication levels of foreign firms are higher than domestic firms and those of multi-product firms are higher than single-product firms. Similarly, high TC firms are found to be more sophsiticated than low TC firms. These differences are also statistically significant.

Table 4: Average Level of product sophistication across types of firm categories

Categories	Firm-level product-sophistication	No. of observations	P-value for t-test.
GVC	29.91	19974	0.000
Non-GVC	26.52	36155	
High innovative TC	31.23	7505	0.000
Low innovative TC	29.15	11983	
High operational TC	31.25	11152	0.000
Low operational TC	28.98	8336	
Domestic-Owned	27.37	50989	0.000
Foreign-owned	31.52	4569	
Single-product firms	27.99	25056	0.000
Multi-product firms	26.96	29273	

Note: Null hypothesis of t-test is that the average product sophisticaition does not differ signficantly across the two categories considered. Sophistication is calculated using prody_tvar.

From table 4, it is clear that GVC firms are producing significantly more sophisticated goods than non-GVC firms. Further disaggregating non-GVC firms, we observe that on average, sophistication levels are highest in GVC firms, followed by only importing firms, only exporting firms and lastly, domestic firms (figure 10). Moving from levels of sophistication to growth, we again find on average, growth in product sophistication in the period 2000-2014 has been much higher for GVC firms than non-GVC (figure 11). Also, categorising firms according to sophistication quintiles, we find that around 45% of the firms producing in the top-most quintile are GVC firms (figure 12). Thus, majority of the highly sophisticated firms in our sample are two-way traders.

Fig 10: Average firm sophistication

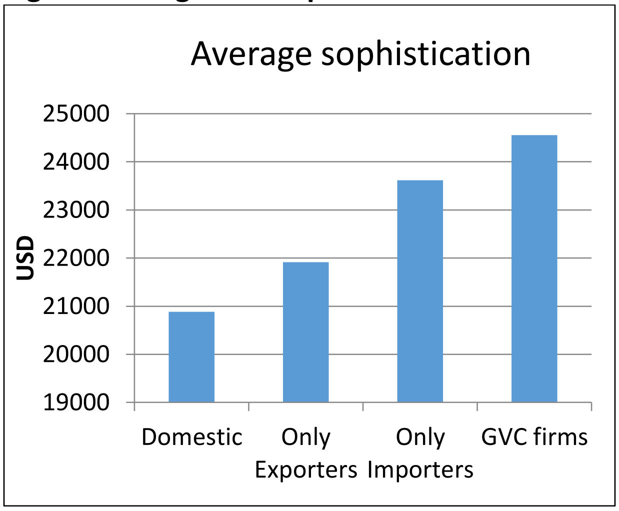
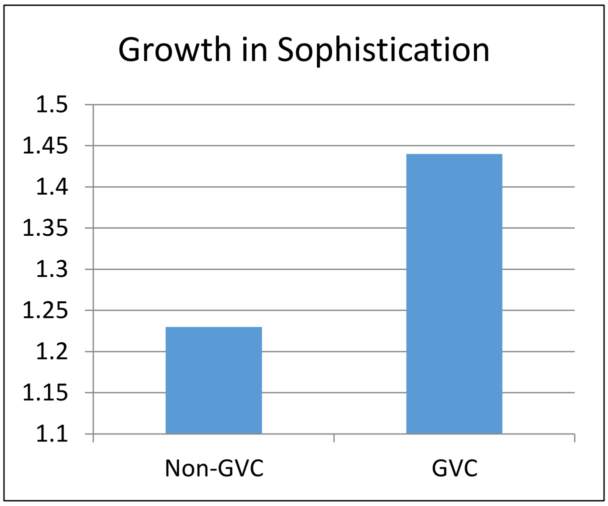
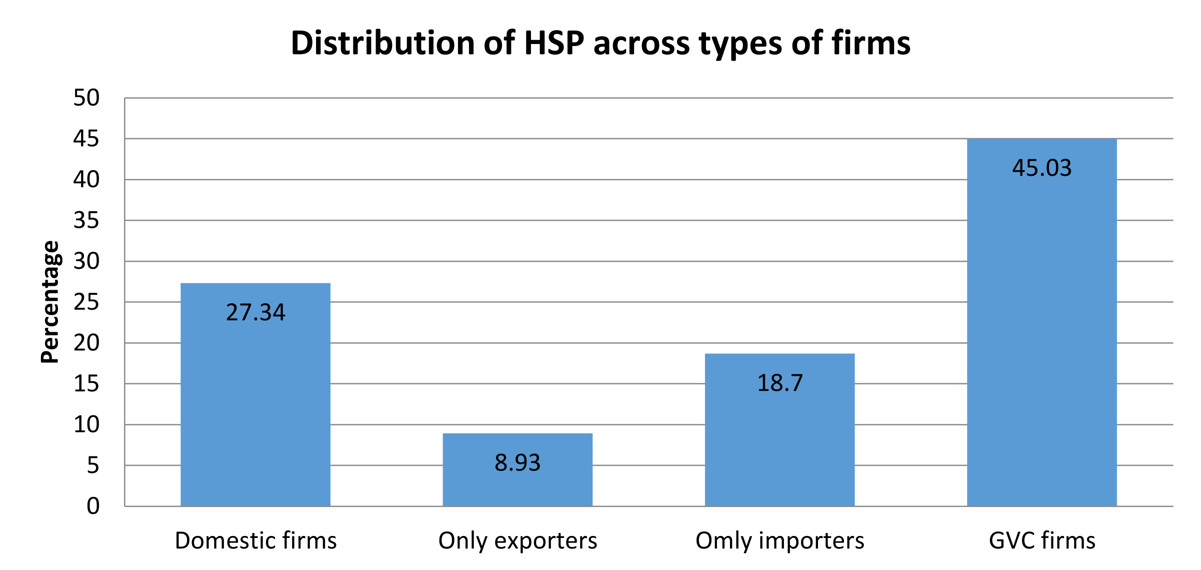


Fig 11: Growth in sophistication-index



Note: product sophistication is calculated using prody_tvar. We have presented actual sophistication levels, instead of normalised values. Using ttest, it is found that that there is a statistically significant difference in means of growth in sophistication across GVC and non-GVC firms at 5%.

Figure 12: Highly-sophisticated producers

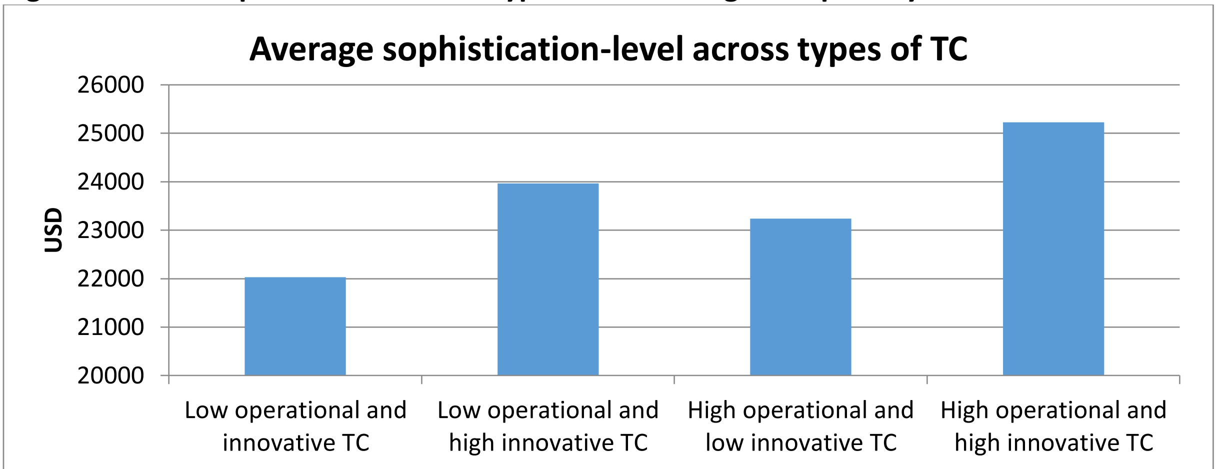


source: Author (2017)

4.4.1 Sophistication across firm capability and ownership

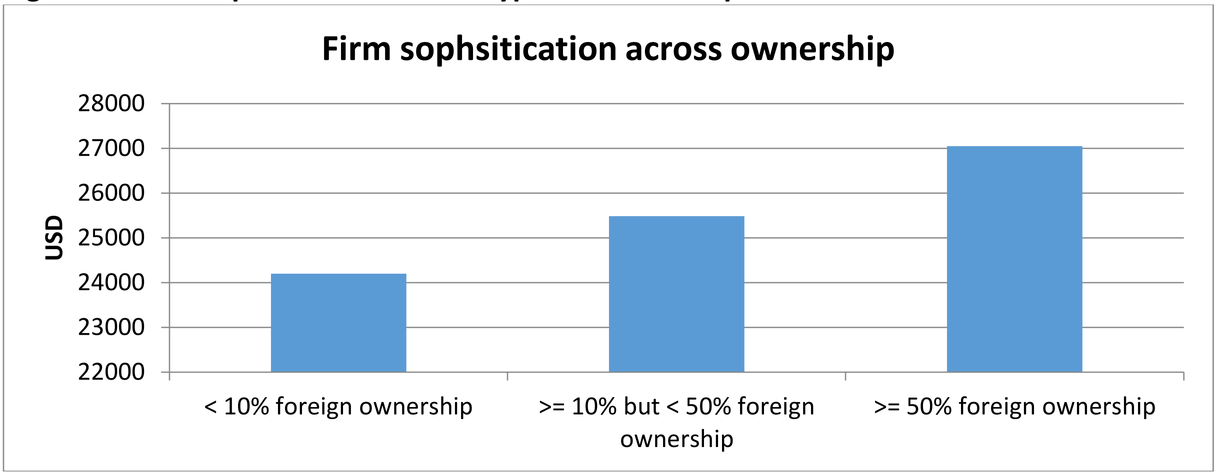
From figure 13, we observe that on average, firms with high operational TC and high innovative TC, produce more sophisticated goods. We also find that firms with low operational TC but high R&D intensity have on average a higher sophistication level that firms with high operational TC but low R&D intensity. This could suggest that, in determining product sophistication level of the firms, firm's own efforts into building its innovative technological capability matter more than building the capacity to absorb external information/knowledge, which is captured by the technology acquisition or operational TC index. Sophistication of a firm can also be influenced by foreign investments. Focusing on the ownership structure of GVC firms, we find that on average, foreign subsidiaries have higher sophistication levels, followed by foreign-owned firms with foreign ownership between 10 to 50% (see figure 14).

Figure 13: Firm sophistication across types of technological capability



source: Author (2017)

Figure14: Firm sophistication across types of ownership

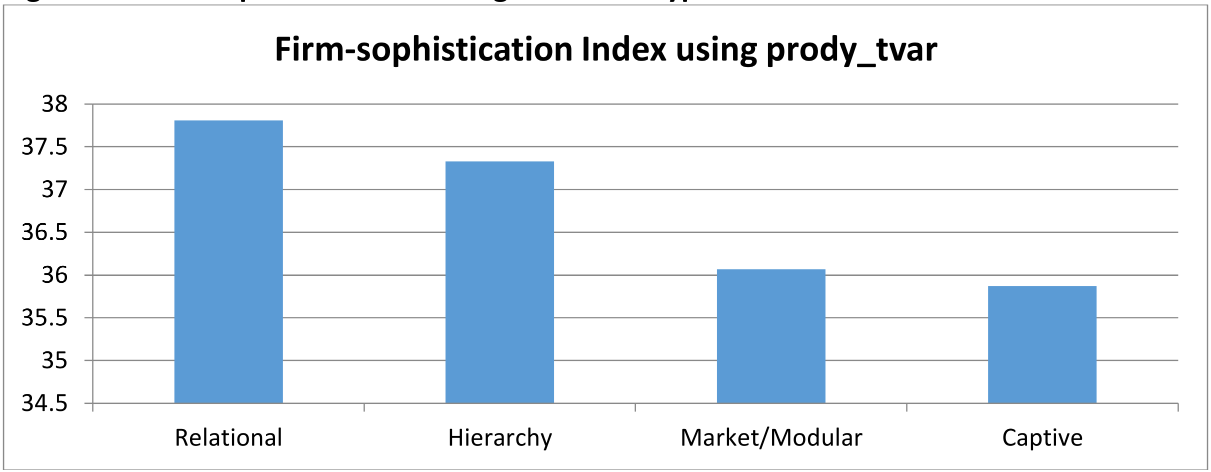


source: Author (2017)

4.4.2 Sophistication across governance structures

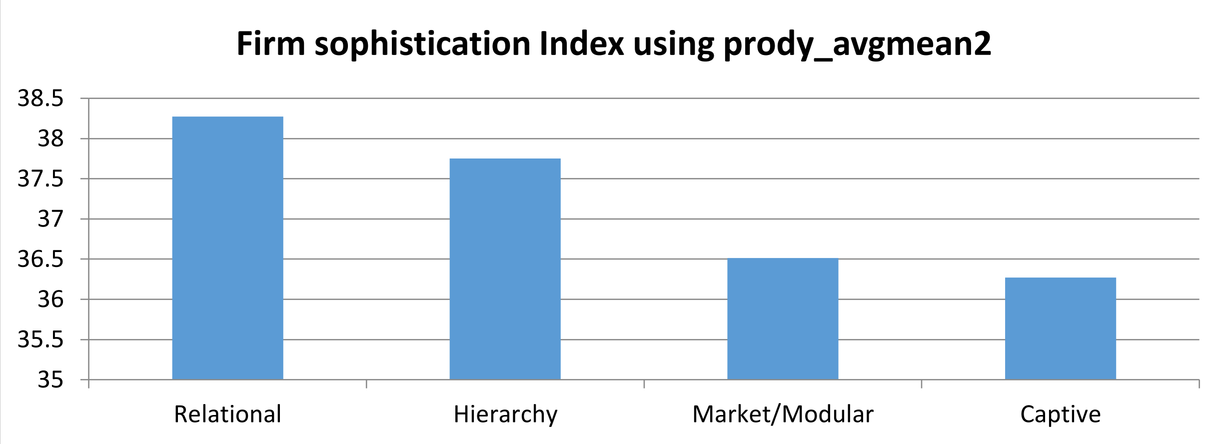
Figure 15 shows that average firm sophistication is highest for firms in relational chains, followed by firms in hierarchy, market/modular and lastly captive. We cross check our results using a different indicator of product sophistication and confirm that firms in captive chains have on average the lowest sophistication-level, while relational firms exhibit the highest (see figure 16). This does not seem to be surprising since a lot of GVC case studies have documented that relational chains provide ideal opportunities for product upgrading. They involve learning through face-to-face interactions and complementary competences between the suppliers and the lead firms. There is also a balance of power between the two, with relationships based on mutual trust and dependence under such governance.

Figure 16: Firm sophistication across governance types



source: Author (2017)

Figure 17: Firm sophistication across governance types (using alternative index for soph.)



Note: Source(Author). See appendix A.4 on how prody_avgmean2 is calculated.

From the above analysis, it is clear that there is a positive relationship between product-sophistication and; two-way trade, firm capability and foreign ownership. To establish a causal relation between sophistication and firm-level characteristics, we move on to empirical investigation in the next section.

5. GVC status and firm sophistication; Empirical analysis

5.1 Econometric model

To analyse the causal impact of GVC integration on firm-level product sophistication, we first use a dummy variable $GVCparti_{it}$ to capture the GVC status of a firm. The baseline specification for this model is given below;

$$\log(PSI_{it}) = \alpha_0 + \log(PSI_{it})_{t-1} + \beta_1 GVCparti_{it} + \beta_2 \log(innovTC_{it}) + \beta_3 OperationalTCindex_{it} + \beta_4 \log(LabourProductivity_{it}) + \beta_5 HHI_{jt} + \beta_6 X_{it} + a_i + a_t + a_j + e_{ijt} \ldots\ldots \text{Model 1}$$

The dependent variable, $\mathrm{PSI}_{\mathrm{it}}$ is firm i's sales weighted average product-sophistication index at time t, where sophistication of a product is calculated used Hausmann's (2007) PRODY index.

$\mathrm{PSI}_{\mathrm{it}_{t-1}}$ is the firm's sophistication level in the previous period and is expected to show a positive sign since there is likely to be persistency in the products produced by a firm. $GVCparti_{it}$ dummy is equal to 1 for all GVC firms at time t. These firms are two-way traders, which can learn by importing as well as exporting. Therefore, we expect GVC firms to have higher product sophistication than one-way traders or domestic firms, resulting in positive sign on the GVC dummy.

$\mathrm{InnovTC_{ijt}}$ is the innovative technological capability of a firm which is measured by R&D intensity while Operational TC index captures the technology acquisition capability of the firm. The coefficients of both these variable are expected to be positive since firms with higher technological capability can better acquire, absorb, assimilate and implement the technology learnt.

$\mathrm{LabourProductivity_{it}}$ is measured by value added per unit of labour and it is expected that more efficient labour can produce more complex goods.

HHI is the Herfindalh-index, defined as the sum of squared market shares of all firms operating in a particular industry. It controls for concentration of the five-digit industry a firm is operating in. If the sign on HHI is positive, it implies that the more concentrated a firm's industry is, the higher are the profits for the firms in that industry. These profits are being re-invested into more sophisticated lines. A negative sign implies that firms in highly concentrated industries faces lower competition and thus lower incentives to innovate.

$\mathrm{X_{ijt}}$ refers to set of variables controlling for firm characteristics such as age, size, foreign shares and product scope (single product firms or multiple product firms). A positive sign on the variable age suggests that older firms have had more experience in 'learning by doing' and may have already established marker power, allowing them to focus on innovation. However, younger firms have more incentive to innovate and remain competitive in the market, in which case the sign on age will be negative. The coefficient on foreign shares is expected to be positive since spill-overs from foreign firms can lead to transfer of knowledge, technology and skill flow. The expected sign on product scope remains ambiguous. Introduction of new products or product innovation can increase the sophistication of products on one hand, but on the other firms may choose the economies of scale route to rent generation and maybe producing a new but low-sophisticated good. Firm, industry and time fixed effects are captured by a(i), a(j) and a(t) terms.

5.2 Empirical strategy and methodology

As our analysis aims to investigate the impact of participating in a GVC on average sophistication level of a firm, the main problem that we need to deal with is endogeneity. Unobserved firm characteristics, such as unobserved productivity, may affect both firm-sophistication and the GVC status of the firm, leading to spurious correlation between the two. Endogeneity and biased results may also arise when unobservable time-invariant firm effects are correlated with regressors in the empirical model. Moreover, reverse causality can bias the results if an increase in firm-level sophistication inturn impacts the explanatory variables. Mainly, we are concerned with the reverse causality between firm sophistication and the GVC dummy. One way traders that start producing more sophisticated goods may find it easier to bear the sunk costs associated with international trade and may therefore be encouraged to become GVC firms. Similarly, firms that have higher sophistication are more likely to attract foreign investment, invest in improving labour efficiency and import more capital goods, making foreign shares, labour productivity and operational TC index potentially endogenous.

To deal with such econometric problems of reverse causality and endogeneity, the study employs the System-GMM methodology. The System GMM estimator simultaneously runs the econometric model in levels and differences, using lagged values of levels as instruments for first differences and lagged values of first differences as instruments for levels. It allows us to examine the cross sectional relationship between participating in a GVC and firm sophistication, since the firm-specific effect is not eliminated but rather controlled by the lagged differences of the independent variables as instruments. Here, the assumption is that differences are not correlated to firm-specific effects compared to levels. It also allows us to include the lagged value of firm sophistication as an explanatory variable, enabling us to deal with problems of: (i) autocorrelation of disturbances in the panel estimation, (ii) time-invariant firm characteristics correlated with explanatory variables, and (iii) some regressors that may be predetermined variables and not strictly exogenous. This estimator uses extra moment conditions that rely on certain stationarity conditions. If these conditions are satisfied, the Sys-GMM estimator has much better finite sample properties, in terms of bias (Blundell and Bond,1998 and Blundell et al, 2000).

To employ system GMM we use the xtabond2 command by Roodman (2009). We use two-step GMM estimations with robust standard errors, clustered at the firm-level to ensure that the standard errors are robust to heteroskedacity and autocorrelation. We also use the 'orthog' option (Arellano and Bover, 1995) which preserves the sample size in unbalanced panels by subtracting the average of all available future observations, rather than subtracting the previous observation from the current one. We include time dummies in all our models since the assumption of no correlation across individuals in the idiosyncratic disturbances, as made by the AR test and robust estimates, is more likely to hold if time effects are included. In system GMM, we can include time-invariant regressors, such as industry dummies, which cannot be included in difference-GMM. Asymptotically excluding them does not affect the coefficient estimates of other regressors because all instruments for the levels equations are assumed to be orthogonal to fixed effects. Moreover, Roodman (2006) mentions that it would be a mistake to include explicit fixed-effects dummies as they would still effectively cause within- groups transformation to be applied. Therefore, we don't include industry fixed effects in our main models, but instead use them as part of robustness checks.

Following the rule of thumb, we keep the instrument count below the number of groups and collapse the instrument set when there are too many instruments. To check against misspecification of instruments, we carry out the Arellano-Bond's autocorrelation test to ensure that there is no autocorrelation in the first differences residuals. A p-value greater then 0.05

indicates that there is no autocorrelation at the second order lag at 5%. The validity of results is also checked using the standard Sargan/Hansen's J -test of over-identifying restrictions. A p-value of Hansen's test lower than 0.05 casts suspicion on the exogeneity of the instrument set. We also report p values for Difference-in-Hansen tests (Roodman, 2009) for the full set of instruments for the levels equation as well as for the subset based on the dependent variable.

5.3 RESULTS

The System-GMM estimates of model 1 are shown in table 5. Seven alternate models are shown. The dependent variable is log(firm-sophistication), where sophistication is sales weighted average sophistication of the products the firm manufactures. The key explanatory variables are lagged values of firm sophistication, GVC participation dummy, log R&D intensity, firm size, firm age, foreign ownership, industry concentration, labour productivity and product scope. The lagged dependent variable, GVC firm dummy, foreign shares, labour productivity, operational TC and firm size are treated as endogenous variables due to potential reverse causality between these variables and the firm-level product sophistication.

The lagged values of firm sophistication are included as explanatory variables to capture persistency in the type of products the firms produce. While studies commonly use one lagged term, in this study two lagged terms have been used. Results in table 5 show that there is persistency in the sophistication level of firms, indicated by the positive and statistically significant coefficient of the two lagged terms of firm sophistication across the seven models. This persistency in firm sophistication confirms the need of using a dynamic GMM approach.

The coefficient of our main variable, the GVC firm dummy, is positive and significant across all models. Controlling for age, size, ownership and product scope model 3 reports that ceteris paribus, the sophistication level of GVC firms is on average around 2% higher than the sophistication level of non-GVC firms. Models 4 and 5 confirm this result using different lag specifications and model 6 and 7 do so by using different specifications explanatory variables. The positive sign suggests that GVC firms are learning by both importing and exporting and the complementarities that arise allow them to produce significantly more sophisticated goods than non-trading firms and one-way traders.

The positive and significant coefficient of HHI indicates that as the concentration of the industry to which the firm belongs increases, the firm's sophistication level rises. Firms in more concentrated industries may be having higher market power and earning higher profits. This could result in these firms re-investing in more sophisticated product lines (Eck and Huber, 2014).

As expected, firm's innovative technological capability, captured by its R&D intensity, is positively related to production of more sophisticated goods. As a firm invests more into building its own knowledge base and innovation, it is more likely to shift towards production of better and more sophisticated products.

Controlling for firm size in model 3, we find that the coefficient on log(age) is negative and significant. This result is in line with the finding of Eck and Huber (2014) that younger Indian firms have a higher level of sophistication. While older firms are more likely to hold greater market power and the ability to innovate, it is the younger firms who have more incentive to innovate and remain competitive. The survival and growth of younger firms may depend on product innovation and manufacturing of more sophisticated goods.

Table 5: Results; Dependent variable: Log(Firm sophistication).

VARIABLES	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
L.log(Firm soph.)	0.606***	0.630***	0.721***	0.699***	0.701***	0.676***	0.633***
	(0.0656)	(0.0811)	(0.0474)	(0.0599)	(0.053)	(0.0837)	(0.0867)
HHI	0.0537***	0.0645***	0.0548***	0.0564***	0.0624***	0.0587***	0.0640***
	(0.0118)	(0.0143)	(0.0120)	(0.0125)	(0.0142)	(0.0152)	(0.0145)
Log(R&D)	0.00417***	0.00412***	0.00279***	0.00359***	0.00274***	0.00338**	0.00432***
	(0.000978)	(0.00126)	(0.00102)	(0.00125)	(0.000976)	(0.00166)	(0.00162)
GVC firm	0.0173*	0.0179*	0.0231**	0.0236**	0.0244*	0.0217**	0.0191**
	(0.00921)	(0.0100)	(0.00903)	(0.00921)	(0.0143)	(0.00929)	(0.00931)
Operational TC	-0.000362	-0.000227	-5.89e-05	-2.57e-05	-0.00025	-0.000110	0.00664*
	(0.000309)	(0.000316)	(0.000192)	(0.000211)	(0.000314)	(0.000225)	(0.00397)
Log(labour prod.)	0.0246*	0.0181	-0.00459	0.00147	-0.00152	0.00537	0.0115
	(0.0139)	(0.0171)	(0.00844)	(0.00959)	(0.00974)	(0.0143)	(0.0134)
Log(age)		-0.00588	-0.0109**	-0.0105**	-0.0121**	-0.00827	-0.00801
		(0.00581)	(0.00486)	(0.00505)	(0.00519)	(0.00596)	(0.00576)
Foreign shares		-0.000350	-0.000207	-0.000261	-0.00010		-0.000300
		(0.000300)	(0.000274)	(0.000264)	(0.00028)		(0.000282)
Multi product firm		-0.0113*	-0.00529	-0.00233	-0.00813	-0.00877	-0.00770
		(0.00637)	(0.00493)	(0.00509)	(0.00519)	(0.00544)	(0.00553)
Log(sale)			-0.000132	-0.00477	0.0030		
			(0.00496)	(0.00578)	0.004		
Foreign firm						-0.00674	
						(0.0220)	
Log(employment)						-0.000224	-0.00278
						(0.00515)	(0.00523)
Lags	2 to 5	2 and 3	2 and 3	2	2 to 4, 3 and 4	2 and 3	2
Time fixed effects	yes	yes	yes	yes	yes	yes	yes
Standard errors	Two-step robust clustered on firms	Two-step robust clustered on firms	Two-step robust clustered on firms	Two-step robust clustered on firms	Two-step robust clustered on firms	Two-step robust clustered on firms	Two-step robust clustered on firms
AR1 AR2 Hansen	0.000 0.15 0.168	0.000 0.147 0.120	0.000 0.149 0.059	0.000 0.146 0.134	0.000 0.149 0.367	0.000 0.142 0.08	0.000 0.135 0.340
DIFF IN Hansen GMM instruments	0.273	0.28	0.120		0.09 for L.dep. 0.63	0.318	
Diff in Hansen IV instruments	0.178	0.08	0.043	0.285	0.65	0.266	0.543
Observations	45,599	43,554	43,554	43,554	43,554	43,554	43,554
Number of firms	6,896	6,526	6,526	6,526	6,526	6,526	6,526
Number of instruments	36	33	36	30	36	36	30

Note: Model 7 calculates operational TC as the share of overall expenditure on royalty and technical know-how fees in firm sales. Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.

5.4 Robustness checks of GMM results

To ensure that the results obtained above are robust, we implement a number of robustness checks in table 5.1. First, we estimate the models using one-step GMM estimation (model1) and also add interactions to two-step results (model2). The coefficient on the GVC dummy remains positive and significant, with similar magnitude as two-step estimations and the interactions are not found to be significant. Next, we check robustness to measurement and lag specifications by running model 1 with labour productivity calculated differently (model 3) and use different number of lags for endogenous variables (model 4). We find that the coefficients remain comparable in terms of magnitude and significance and have the same signs. Then, to check for attrition bias in our results, we also run the empirical model on a smaller sub-sample of firms that consistently report data for the period 2001-2015 (model 5). In this balanced panel also, we find that GVC firms are producing more sophisticated goods than non-GVC firms. Finally, we also check robustness by using a different dependent variable to capture firm sophistication. In model 6, we calculate firm sophistication as sales weighted average of firm's sophistication where product sophistication is calculated as the average of PRODY over time. Results remain comparable and similar to those obtained earlier.

Table 5.1 Robusntess checks

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	**One step**	**With Interaction**	**With LP calculated differently**	**Different Lag specification**	**Balanced panel**	**Different dependent variable**
L.log(Firm soph.)	0.563***	0.725***	0.718***	0.701***	0.560***	0.674***
	(0.0965)	(0.0469)	(0.0480)	(0.0541)	(0.0616)	(0.113)
HHI	0.0688***	0.0544***	0.0527***	0.0624***	0.0613***	0.0550**
	(0.0136)	(0.0119)	(0.0131)	(0.0144)	(0.0172)	(0.0217)
Log(R&D)	0.00245*	0.00271***	0.00251**	0.00274***	0.00297***	0.00374**
	(0.00140)	(0.00100)	(0.00102)	(0.000976)		(0.00176)
Foreign shares	-0.000168	-0.000263				-0.000146
	(0.000277)	(0.000482)				(0.000214)
Operational TC index	0.000100	-5.77e-05	-0.000117	-0.000262	-0.000140	-9.83e-05
	(0.000240)	(0.000188)	(0.000200)	(0.000314)	(0.000154)	(0.000202)
GVC	0.0214**	0.0226**	0.0232**	0.0244*	0.0377**	0.0140**
	(0.0103)	(0.00920)	(0.00900)	(0.0143)	(0.0168)	(0.00693)
GVC*foreign shares		0.000145				
		(0.000487)				
Log(age)	-0.00266	-0.0112**	-0.0117**	-0.0121**	-0.0184*	-0.00395
	(0.00627)	(0.00482)	(0.00503)	(0.00519)	(0.0102)	(0.00567)
Log(Labour prod.)	0.0148	-0.00547	-0.00546	-0.00813	0.00914	0.0107
	(0.0134)	(0.00828)	(0.00607)	(0.00519)	(0.00934)	(0.00840)
Multi-product firm	-0.00829	-0.00511	-0.00679	-0.00679	-0.0228**	-0.00840
	(0.00587)	(0.00492)	(0.00504)	(0.00504)	(0.00931)	(0.00631)
Foreign firm			0.00396	0.00396	-0.0189	
			(0.0195)	(0.0195)	(0.0218)	
Standard errors	One step robust	Two step robust	Two-step robust, clustered on firms	Two-step robust clustered on firms	Two-step robust clustered on firms	Two-step robust clustered on firms
AR1	0.000	0.000	0.000	0.000	0.000	0.03
AR2	0.11	0.148	0.136	0.149	0.91	0.462
Hansen	0.11	0.10	0.126	0.359	0.19	0.356
Diff in Hansen for iv Diff in Hansen for GMM	0.48	0.15 0.09	0.37 0.56	0.4 .606 0.09 for dep.	0.8 0.11	0.497 0.134
Control for firm size	yes	yes	yes	yes	yes	yes
lags	2	2 and 3	2 and 3	2 to and 3to4	2 and 3	2 and 3
Time Fixed effects	yes	yes	yes	yes	yes	yes
Observations	43,554	43,554	43,620	43,554	13,977	
Number of firms instruments	6,526 31	6,526 39	6,533 36	6,526 36	1,047 39	6526 36

Note: Firm sophistication calculated using prody_tvar. Constants and year dummies not reported.

5.5 Robustness using PSM

If GVC and non-GVC firms have very different characteristics, the econometric issue of selection into GVCs can arise. To check robustness against this issue, we need to first make GVC and non-GVC comparable and then analyse the impact of linking into GVC on firm sophistication. To do this as a robustness check, we take a subsample of firms in the period 2011/2012-2013/2014 and conduct propensity score matching. After matching GVC firms with non-GVC firms, that are similar in terms of firm characteristics such as R&D intensity, size, productivity, ownership etc., we employ fixed-effects estimation to check if the impact of participating in a GVC on firm-sophistication changes after accounting for selection.

5.5.1 Methodology and empirical strategy for PSM-FE

The first step in PSM-FE is to choose variables to match treatment and control groups on. When choosing matching variables to match GVC and non-GVC firms, we follow Caliendo and Kopeinig (2005) and choose only those variables that simultaneously affect the participation decision (decision to participate in GVCs in our case) and the outcome variable (firm level product sophistication). We also ensure that the variables used for matching are exogenous to the treatment, by measuring them before participation. We avoid over-parameterised models since including extraneous variables in the propensity score model exacerbates the support problem and inclusion of insignificant variables in the propensity score specification can increase their variance (Bryson et al., 2002).

Next we model the probability of a firm being a GVC firm, conditional on observable pre-treatment variables. We use a Logit model for each year – 2011/2012, 2012/2013 and 2013/2014- separately and model the propensity score as a function of firm-specific attributes at time t-1 and industry dummies. We include industry dummies to avoid firms in very different industries being matched to each other.

Our final propensity score specification for a firm to be a GVC firms is as follows:

$$\Pr(GVC_t = 1) = \emptyset[industry\ controls_t, \begin{Bmatrix} \{\text{Log(R\&Dintensity)}, \text{Smallfirm}, \text{Bigfirm}, \text{Foreignfirm} \\ , \text{Log(Labour prod.)}, \text{Log(age)}\} \end{Bmatrix}_{t-1}]$$

where, industry controls are industry dummies, Smallfirm and Bigfirm are dummy variables capturing size of the firm and Foreign firm is a dummy variable equal to one for all firms that have more than 10% foreign ownership.

The results for estimation of the above specified propensity are given in the appendix, section A.9. We ensure that for each year the balancing property holds in all blocks for all covariates. This ensures that there is no significant difference in means across treatment and control groups in each block.

After obtaining the probability that a firm participates in a GVC, we match GVC firms with non-GVC. Matching is done separately for each year and using Stata's Psmatch2 command. We use Kernel Matching[15], a non-parametric matching estimator in which each treated individual is assigned a weight of 1 and the counterfactual outcome is constructed using the weighted average of all individuals in the control group. Thus, one major advantage of using this estimator is that it uses all possible information which results in this estimator having lower variance compared to other matching estimators (Caliendo and Kopeinig, 2005).

After Kernal matching, we check the quality of matching using 'pstest' command (developed by Leuven and Sianesi (2012)) in Stata. For each variable included in the model, the pstest command reports the means of treatment and control groups and a standardised percentage bias, and performs t-test for equality of means across the treatment and control groups[16]. It also provides us with overall measures of covariate balance between treatment and control groups, including pseudo-R-squared score, the value of likelihood-ratio test of the joint significance of all regressors, the mean and median bias, and the Rubin's B and R score.

Table 5.2 compares the unmatched and matched sample and finds that there has been a significant reduction in bias in each variable. The pseudo-R-squared score indicates how well the regressors in model explain the probability of selection into the treatment group (Sianesi, 2004). After matching, there should not any systematic differences in the distribution of covariates between the GVC and non-GVC firms. Therefore, pseudo-R-squared should be fairly low (Caliendo and Kopenig, 2008). We find that in our case (see table 4.3), the regressors explained around 15% of the selection in the unmatched sample, while in the matched sample this has gone down to 0.002, which is quite low. Table 5.3 also reports that the likelihood-ratio test of joint insignificance of all regressors in the logistic regression model has a p-value of 0.358, implying that the null hypothesis of joint insignificance cannot be rejected. For treatment and control groups to be sufficiently balanced, Rubin's B statistic should be less than 25 and Rubin's R statistic should lie between 0.5 and 2 (Rubin, 2001). Our Rubin's B statistic of 10.1 and Rubin's R statistic of 1.12 in table 5.3 clearly satisfy these conditions. Thus, we can conclude that our matching is successful.

[15] Bandwith used for matching - 0.06

[16] T-test makes controversial assumptions- including normal distribution of covariates in treatment and control groups- and is sensitive to sample size. Therefore, many studies advise against performing t-test based comparisons after PSM (Austin, 2009)

Table 5.2: Results from PSTEST command

Variable	Unmatched/ Matched	Mean		Bias		t-test	
		Treated	control	% bias	% reduction in bias	t	p>|t|
L.Log(R&Dintensity)	U	-8.4363	-12.193	69.5		37.89	0
	M	-8.0215	-8.4617	8.1	88.3	1.76	0.079
L.Small firm	U	0.0414	0.24803	-61.4		-28.98	0
	M	0.04012	0.04274	-0.8	98.7	-0.33	0.739
L. Big firm	U	0.75572	0.41483	73.7		37.14	0
	M	0.76775	0.75558	2.6	96.4	0.73	0.468
L. Foreign firm	U	0.13833	0.04286	33.7		18.65	0
	M	0.13889	0.11295	9.2	72.8	1.99	0.047
L.Log(Age)	U	3.2227	3.0151	29.9		15.2	0
	M	3.2547	3.2344	2.9	90.2	0.76	0.449
L.Log(Labour prod.)	U	-4.7589	-4.7211	-3.6		-1.81	0.07
	M	-4.7054	-4.7085	0.3	91.7	0.09	0.931

Table 5.3: Results from the PSTEST command

Sample	Ps R2	LR chi2	p>chi2	MeanBias	Med Bias	B	R	%Var.
Unmatched	0.153	2294.41	0	45.3	47.6	99.7*	0.94	100
Matched	0.002	6.62	0.358	4	2.8	10.1	1.13	33

Having successfully matched GVC firms with similar non-GVC firms, we find that 12.32 % of our sample lies outside the common support (see table 5.4). For each year, we drop all these observations that are not matched. Dropping these observations, we are left with 11,456 observations in the common support region. In this sample, 35.84 percent of the observations are classified as Treated (table 5.5). Other summary statistics of this truncated sample are given in table A.13.

Table 5.4: Identifying common support after matching

COMMON SUPPORT	FREQUENCY	PERCENT
0	1,610	12.32
1	11,456	87.68

Table 5.5: Treated and Control group in the truncated sample

Treatment	FREQUENCY	PERCENT
0	7,350	64.16
1	4,106	35.84

Next, we use the Hausmann's (1978) specification test to choose between Fixed effects or Random effects. A significant p-value of the Hausman's test statistic implies that the FE estimation is appropriate17. Following the literature on treatment effects after PSM, we present results with bootstrapped standard errors (150 reps.)

5.5.2 Results for PSM-FE

Results are shown in table 5.6. In model 1, we include only those controls which we have used in the propensity score specification for being a GVC firm. We find that even after accounting for selection, the coefficient on GVC firm is positive and significant. Similar to results in GMM estimations, we find that bigger firms produce less sophisticated goods compared to medium-sized firms. While age is positive and significant in this model, we find that it loses significance when we control for time-fixed effects in model 2. Model 3 adds a control for industry concentration to model 2 and finds that as the firm's industry becomes more concentrated, the firm's sophistication level rises. This is similar to what we obtained in the GMM estimations. Model 4 adds a multi-product firm dummy to control for effects of product scope. Model 5 controls for operational technological capability by adding firm's expenditure on royalty and technical know-how, as a share of total sales. We find that across all the five models, the coefficient on GVC firm remains positive, significant and comparable to our GMM results. The results hold across different specifications in the balanced panel as well (Models 6 and 7).

[17] Results are shown in appendix A.14.

Table 5.6: Results for PSM-FE. Dependent variable- Log (Firm sophistication)

VARIABLES	Model1 (unbalanced)	Model2 (unbalanced)	Model3 (unbalanced)	Model4 (unbalanced)	Model5 (unbalanced)	Model 6 (balanced)	Model 7 (balanced)
GVC firm	0.00756*	0.00789*	0.00824*	0.00892**	0.00891*	0.00931**	0.0102**
	(0.00430)	(0.00471)	(0.00436)	(0.00447)	(0.00498)	(0.00468)	(0.00427)
L.Log(R&D intensity)	0.000316	0.000211	0.000231	0.000264	0.000264	-0.000147	-7.90e-05
	(0.000708)	(0.000636)	(0.000704)	(0.000707)	(0.000757)	(0.000733)	(0.000739)
L. Small firm	0.00466	0.00396	0.00298	0.00288	0.00289	0.00123	0.000286
	(0.00835)	(0.00857)	(0.00953)	(0.0105)	(0.00961)	(0.00778)	(0.00826)
L. Big firm	-0.0135*	-0.0133*	-0.0140**	-0.0141**	-0.0141*	-0.0119*	-0.0128
	(0.00713)	(0.00729)	(0.00650)	(0.00658)	(0.00731)	(0.00674)	(0.00884)
L. Foreign Firm	0.0222	0.0245	0.0245	0.0256*	0.0255	0.0263	0.0276
	(0.0153)	(0.0185)	(0.0165)	(0.0149)	(0.0185)	(0.0165)	(0.0208)
L. Log(Age)	0.0860***	0.0208	0.0222	0.0234	0.0234	0.0150	0.0172
	(0.0140)	(0.0147)	(0.0155)	(0.0146)	(0.0165)	(0.0181)	(0.0182)
L. Log(Labour prod.)	-0.000535	-0.00116	-0.00133	-0.00138	-0.00138	-0.00125	-0.00149
	(0.00239)	(0.00273)	(0.00249)	(0.00271)	(0.00301)	(0.00269)	(0.00283)
HHI			0.101***	0.0981***	0.0981***		0.102***
			(0.0266)	(0.0248)	(0.0252)		(0.0291)
Multi-product firm				-0.00500	-0.00499		-0.00331
				(0.00776)	(0.00748)		(0.00707)
Royalty/Sales					0.000330		0.000393
					(0.00396)		(0.00352)
2013.year		0.00562**	0.00478*	0.00517**	0.00516**	0.00749***	0.00718**
		(0.00257)	(0.00267)	(0.00250)	(0.00254)	(0.00273)	(0.00327)
2014.year		0.0147***	0.0127***	0.0128***	0.0128***	0.0155***	0.0137***
		(0.00309)	(0.00323)	(0.00344)	(0.00312)	(0.00308)	(0.00370)
Constant	3.020***	3.213***	3.190***	3.177***	3.177***	3.235***	3.199***
	(0.0481)	(0.0488)	(0.0495)	(0.0496)	(0.0528)	(0.0603)	(0.0591)
Observations	9,075	9,075	9,075	8,750	8,750	7,404	7,145
R-squared	0.006	0.009	0.011	0.011	0.011	0.009	0.011
Number of Firms	3,700	3,700	3,700	3,574	3,574	2,536	2,455

Note: Standard errors are bootstrapped (150 reps).

6. GVC integration and firm- level sophistication; Empirical analysis

To examine the impact of increasing GVC participation on firm-level sophistication, we take the sub-sample of GVC firms and estimate the following baseline model;

$$\log(PSI_{it}) = \alpha_0 + \log(PSI_{it})_{t-1} + \beta_1 \log(GVCintegration)_{it} + \beta_2 \log(innovTC_{it}) + \beta_4 \log(LabourProductivity_{it}) + \beta_5 HHI_{jt} + \beta_6 X_{it} + a_i + a_t + a_j + e_{ijt} \ldots\ldots \text{ Model 2.}$$

Here, GVC integration is proxied by different measures of vertical specialisation. A positive sign on this variable will imply that as the intermediates embodied in firm level exports increases (it becomes more integrated into GVCs), its product-sophistication level rises.

To estimate this model, we again employ system GMM to account for endogeneity caused due to omitted variable bias and reverse causality. Table 6 presents GMM results with GVC integration proxied by VS1. VS1 measures the import content embodied in exports and includes raw materials, capital goods, stores and spares and services in intermediate inputs. In model 1, we regress firm sophistication on lagged values of firm sophistication, logVS1, age of the firm and size. In models 2 and 3, we add controls for foreign ownership of the firm. In model 4, we control for innovative technological capability of a firm. In model 5, HHI, labour productivity and product scope are introduced as controls. In model 6, we include a high technology sector dummy which is equal to 1 for firms operating in high technology sectors.

Across the models, we find that, two period lagged sophistication positively and significantly affects current period's sophistication level. From this we can infer that there is persistency in firm sophistication. Similar to results obtained in section 5, older firms are producing less sophisticated goods and firm size, innovative capacity and industry concentration are positively related to firm sophistication. The coefficient on foreign firm dummy and foreign shares variables is positive but not significant; foreign ownership does not seem to be affecting sophistication levels of Indian firms.

Turning to our main variable of interest, log vs1, we find that the coefficient on vertical specialization is positive and significant. This implies that other things constant, as the imported intermediate embodied in firm level exports increases, the average product sophistication level of firm also increases. Firms are able to gain access to more variety and better quality foreign inputs and also learn from the technology embedded in foreign inputs. This results empirically supports the learning by importing and exporting hypothesis.

As part of robustness checks (table 6.1), we control for industry characteristics by adding dummy for firms operating in medium to high technology sectors[18]. This classification is based on technology intensity of industries and classifies the following industries as high-tech; aircraft and spacecraft, pharmaceuticals, office, accounting and computing machinery, radio and TV communication equipment, motor vehicles, machinery and equipment, chemicals etc. The positive and significant coefficient on the high-tech sector dummy in model 1 implies that medium to high technology industries produce more sophisticated goods. Next, we check sensitivity of results to the measure of vertical specialization used. In model 2 and 3, we use VS2 measure to capture GVC integration. This measure excludes services from intermediates inputs. We find that the coefficient on logvs2 is also positive and significant and comparable to the coefficient on logvs1. In models 4 and 5, we use logVS3 to capture magnitude of GVC participation. This measure excludes both services and capital goods from intermediate goods and account for only imported raw materials and stores and spares embodied in a firm's exports. Again, we find that results remain similar and comparable.

[18] High technology sectors are classified as NIC 20 (chemicals and chemical products), 21 (pharmaceuticals), 26 (computer, electronics and opticals), 27 (electrical equipment), 28 (machinery and equipment), 29 (motor vehicles), 30 (other transport).

Table 6: Dependent variable log (firm sophistication)

VARIABLES	Model1	Model2	Model3	Model4	Model5
L.Log(firm soph.)	0.808***	0.802***	0.798***	0.790***	0.793***
	(0.0429)	(0.0430)	(0.0432)	(0.0471)	(0.0457)
L2.Log(firm soph.)	0.0777*	0.0795*	0.0830**	0.0851**	0.0817*
	(0.0404)	(0.0406)	(0.0400)	(0.0421)	(0.0440)
Log(VS1)	0.00767**	0.00613*	0.00695*	0.00657*	0.00621*
	(0.00366)	(0.00374)	(0.00375)	(0.00360)	(0.00377)
Log(age)	-0.0104*	-0.0115**	-0.0113*	-0.0144**	-0.00899*
	(0.00581)	(0.00584)	(0.00582)	(0.00564)	(0.00544)
Log(employment)	0.00965*	0.00896*	0.00993**	0.00842	0.00597
	(0.00498)	(0.00507)	(0.00501)	(0.00513)	(0.00518)
Foreign firm		0.0209			
		(0.0168)			
Foreign shares			0.000202	0.000177	0.000113
			(0.00019)	(0.000191)	(0.00018)
Log(R&D)				0.00131*	0.00144*
				(0.000694)	(0.00075)
HHI					0.00718
					(0.00647)
Multi product firm					-0.000647
					(0.00421)
Log(Labour prod.)					0.00597
					(0.00872)
Constant	0.396***	0.410***	0.409***	0	0
	(0.136)	(0.144)	(0.145)	(0)	(0)
Observations	11,670	11,611	11,611	11,611	11,259
Number of firms	1,980	1,961	1,961	1,961	1,899
No. of instruments	35	38	38	42	44
AR(1)	0.000	0.000	0.000	0.000	0.000
AR(2)	0.436	0.470	0.526	0.593	0.558
Hansen p val	0.166	0.08	0.24	0.05	0.07
Sargan p val	0.179	0.206	0.302	0.10	0.124
Time fixed effects	Yes	Yes	Yes	Yes	Yes

Note- VS1 = (import of raw materials + import of stores and spares+ import of capital goods+ import of services)/ (raw material expenses+ SS consumed+ GFA+ total expenses on services). Standard errors are two-step robust, clustered on firms.

Table 6.1: Robustness checks with different measures of vertical specialisation

VARIABLES	Model1	Model2	Model3	Model4	Model5
L.Log(firm soph.)	0.803***	0.788***	0.795***	0.802***	0.795***
	(0.0478)	(0.0479)	(0.0477)	(0.0458)	(0.0474)
L2.Log(firm soph.)	0.046	0.0807*	0.050	0.0891*	0.0582
	(0.0431)	(0.0436)	(0.0441)	(0.0464)	(0.0458)
Log(vs1)	0.0064*				
	(0.0035)				
Log(vs2)		0.00763**	0.0063*		
		(0.00377)	(0.0037)		
Log(vs3)				0.00699*	0.00695*
				(0.00412)	(0.00404)
Log(age)	-0.00664	-0.0140**	-0.005	-0.0104*	-0.00481
	(0.00491)	(0.00563)	(0.004)	(0.00626)	(0.00518)
Log(Employment)	0.0032	0.00838*	0.0017	0.00738	0.00297
	(0.0051)	(0.00501)	(0.005)	(0.00623)	(0.00583)
Foreign shares	0.00021	0.000200	0.00019	0.000174	0.000189
	(0.000180)	(0.000192)	(0.00018)	(0.000246)	(0.000204)
Log(R&D)	0.00042	0.00121*	0.00057	0.000661	0.000362
	(0.0005)	(0.000706)	(0.00056)	(0.000979)	(0.000631)
HHI	0.01568**	0.00586	0.015**		0.0122*
	(0.00770)	(0.00712)	((0.0064)		(0.00643)
Multi-product firm	0.0043		0.004		0.00416
	(0.0041)		(0.0040)		(0.00448)
Log(Labour prod.)	0.00972		0.0113		0.0133
	(0.00843)		(0.008)		(0.00908)
High tech. sector	0.044***		0.0435***		0.0407***
	(0.0089)		(0.0089)		(0.0087)
Constant	0	0	0	0.389***	0
		(0)	(0)	(0.151)	(0)
Time fixed effects	yes	yes	yes	yes	yes
Observations	11,259	11,523	11,175	10,676	10,385
Number of firms	1,899	1,950	1888	1,801	1,750
No. of instruments	50	50	50	39	46
AR(1)	0.000	0.000	0.000	0.000	0.000
AR(2)	0.186	0.555	0.230	0.660	0.301
Hansen p val	0.062	0.055	0.706	0.073	0.130
Sargan p val	0.07	0.173	0.091	0.169	0.258

Note: VS2= import of raw materials + import of stores and spares+ import of capital goods+ / raw material expenses+ SS consumed+ GFA. VS3= import of raw materials + import of stores and spares / raw material expenses+ SS consumed. Standard errors are two-step robust, clustered on firms.

Another measure of importance, particularly for trade policy, is the import content of exports measured as the ratio of imported intermediates to total exports. Here, under imported intermediates, we include import of raw materials, stores and spares, capital goods and services. Results using this measure as a proxy for GVC integration are shown in table 6.2 Six alternate models are presented. We observe that in line with out previous results, lagged values of firm sophistication are positively and significantly related to current year's firm sophistication. Older firms produce less sophisticated goods than younger firms, while larger firms produce more sophisticated goods than smaller firms. The sign on log import content of exports is positive and significant, implying that as the share of intermediates imports in exports increase, product sophistication of the firm rises.

Table 6.2; Results with Import content of exports capturing GVC integration

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	model1	model2	model3	model4	model5	Model6
L.Log(firm soph.)	0.786***	0.767***	0.774***	0.791***	0.781***	0.789***
	(0.0398)	(0.0434)	(0.0409)	(0.0410)	(0.0434)	(0.0409)
L2.Log(firm soph.)	0.132***	0.134***	0.132***	0.125***	0.110**	0.109**
	(0.0395)	(0.0404)	(0.0415)	(0.0462)	(0.0471)	(0.0476)
Log(Import content of exports)	0.0255***	0.0197**	0.0215**	0.0242***	0.0154*	0.0193**
	(0.00903)	(0.00856)	(0.00880)	(0.00874)	(0.00838)	(0.00822)
Log(age)	-0.0166***	-0.0151***	-0.0123**	-0.0161***	-0.00911*	-0.0113**
	(0.00537)	(0.00506)	(0.00500)	(0.00495)	(0.00497)	(0.00494)
Log(Employment)	0.0110**	0.0103**	0.00915*	0.00935*	0.00398	0.00427
	(0.00500)	(0.00490)	(0.00526)	(0.00542)	(0.00538)	(0.00550)
Foreign Firm		0.0215	0.0230	0.0238	0.0163	0.0233
		(0.0160)	(0.0160)	(0.0161)	(0.0160)	(0.0155)
HHI			-0.00171	-0.00271	0.000819	0.000950
			(0.00731)	(0.00721)	(0.00700)	(0.00736)
Multi-product firm			-0.00757*	-0.00779**	-0.00584	-0.00322
			(0.00395)	(0.00393)	(0.00394)	(0.00393)
Log(R&D)				0.00103	0.00112	0.000804
				(0.000760)	(0.000770)	(0.000806)
Log(Labour Prod.)					0.0115	0.00514
					(0.00827)	(0.00822)
High tech. sector						0.0193*
						(0.0108)
Constant	0	0	0	0.293*	0.454***	0
	(0)	(0)	(0)	(0.163)	(0.171)	(0)
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	11,802	11,743	11,380	11,380	11,380	11,380
Number of firms	1,998	1,979	1,916	1,916	1,916	1,916
No. of instruments	33	39	41	41	45	42
AR(1)	0.000	0.000	0.000	0.000	0.000	0.000
AR(2)	0.503	0.496	0.539	0.679	0.934	0.957
Hansen p-val	0.310	0.113	0.130	0.151	0.058	0.074
Sargan p-val	0.590	0.561	0.576	0.417	0.186	0.158

Note: Import content of export= import of raw materials + import of stores and spares+ import of capital goods+ import of services /total exports. Standard errors are two-step robust, clustered on firms.

7. Governance and firm-level sophistication; Empirical analysis

To measure the impact of governance on firm-level sophistication, we first estimate equation 3 (as below), with difficulty in codification of transactions and supplier competence levels capturing governance.

$$\log(PSI_{ijt}) = \alpha_0 + \log(PSI_{ijt})_{t-1} + \log(PSI_{ijt})_{t-2} + \beta_1 Log(Difficulty\ in\ codification\ of\ transactions)_{ijt} + \beta_2 \log(Supplier\ competence) + \beta_2 \log(innovTC_{ijt}) + \beta_3 \log(LabourProd_{ijt}) + \beta_4 HHI_{jt} + \beta_5 X_{ijt} + a_i + a_t + e_{ijt} \ldots\ldots\ldots\ldots \text{ Eq. 3.}$$

Here, the share of skilled labour is used as an inverse proxy for codification and an ICT index is used as a proxy for supplier competence. We expect difficulty in codification of tasks (and our proxy of share of skilled labour) to be positively related to supplier firms' sophistication levels. Similarly, we expect that as supplier competence increases, the supplier can produce more sophisticated goods and its average sophistication level will rise.

Next, we substitute the actual values of codification and competence variables in equation 3 with governance dummies for captive, hierarchical and market/modular firms. In equation 4, we expect a negative values for β_1, β_2 and β_3, since firms in Relational chain (the base governance category) have both a high share of skilled labour and a high ICT index relative to other firms in the industry.

$$\log(PSI_{ijt}) = \alpha_0 + \log(PSI_{ijt})_{t-1} + \log(PSI_{ijt})_{t-2} + \beta_1 Captive_{ijt} + \beta_2 Hierarchical_{ijt} + \beta_3 Market/Modular_{ijt} + \beta_4 \log(innovTC_{ijt}) + \beta_5 \log(LabourProd_{ijt}) + \beta_6 HHI_{jt} + \beta_7 X_{ijt} + a_i + a_t + e_{ijt} \ldots\ldots\ldots\ldots \text{Eq.4}$$

It is important to note that, as these suppliers become more sophisticated, they may increase investment into ICT to improve their capabilities and hire more skilled labour. Therefore, by successfully upgrading their products, suppliers may be moving into a more power-symmetrical governance structure. To deal with this problem of reverse causality running from firm sophistication to governance, we estimate equation 3 and 4 using the System-GMM estimator and treat all governance variables as endogenous. The results are presented in section 7.1.

7.1 Results

Table 7 presents four specifications of equation 3. In model 1, firm-level product sophistication is regressed on lagged values and we find that sophistication in period t is positively and significantly affected by sophistication in period t-1. We use the log of share of skilled labour and log of ICT index to capture governance, and find that both variables positively impact firm sophistication. This implies that supplier firms in Relational chains, which are expected to have higher share of skilled labour and higher competence compared to firm in other governance structures, offer higher opportunities for product upgrading. We have controlled for firm's own R&D efforts, labour productivity, foreign ownership, industry concentration, age and size. In model 2, we further control for product scope and confirm results of model 1.

In model 3, we replace share of skilled labour and ICT index with dummy variables capturing skilled labour and competence of the firm relative to the five-digit industry it is operating under. High-skilled firm is a dummy variable equal to 1 for all firms whose share of skilled-labour is above the median share in its two-digit industry. Similarly, High-competence firm is a dummy variable equal to 1 for all firms whose ICT index is greater than the median value of ICT index in its two-digit industry. Results from model 3 show that high-skilled firms have significantly higher sophistication than low-skilled firms. The coefficient on high-competence firm is also positive but not significant. In model 4, we add industry dummies to model 3 to capture industry specific effects. We find that results remain robust.

In table 7.1, we present results using combinations of the high-skilled firm dummy and the high competence dummy to define different governance structures. Model 1 regress firm sophistication on lagged values, age, size and governance. In model 2, we control for firm's innovative capacity.. Models 3, 4 and 5 introduce labour productivity, product scope and HHI, and foreign ownership as controls respectively. 6 runs model 5 with a different lag specification.

For our main variables; dummies capturing governance structures, we find that firms in Captive chains have roughly 5% lower sophistication level than firms in relational chains. This result is statistically significant and consistent across the models. Captive firms are suppliers with relatively low competence and low share of skilled labour, and therefore end up competing in the production of less-sophisticated goods. Production in such firms is tightly monitored by the Lead firms and while knowledge flows may take place from the Lead firm to the supplier firms, they maybe confined to perform a narrow range of activities, such as production of standard consumer goods (Strasser, 2015) or mere assembly (Pietrobelli and Rabellotti, 2008). The control of the Lead-firm remains high in Captive firms, resulting in a risk of 'lock-in' of GVC suppliers in the production of low-value-added goods, as we observe in the case of Indian GVC suppliers.

Table 7: Results: Dependent variable- Log (Firm sophistication)

VARIABLES	(1) model1	(2) model2	(3) model3	(4) model4
L.Log(firm soph.)	0.808***	0.811***	0.791***	0.796***
	(0.0444)	(0.0446)	(0.0464)	(0.0425)
L2.Log(firm soph.)	0.0862	0.0845	0.0897*	0.0927*
	(0.0571)	(0.0576)	(0.0472)	(0.0484)
Log(R&D)	0.000303	0.000289	0.00144*	0.000589
	(0.000863)	(0.000909)	(0.000818)	(0.000834)
Log(Supplier competence)	0.0161*	0.0162*		
	(0.00968)	(0.00991)		
Log(Share of skilled labour)	0.0119***	0.0122***		
	(0.00367)	(0.00378)		
High Competence firm			0.0151	0.0120
			(0.0106)	(0.0107)
High skilled firm			0.0129**	0.0138**
			(0.00513)	(0.00558)
Log(employment)	0.0126**	0.0130**	0.00893	0.0102
	(0.00596)	(0.00632)	(0.00615)	(0.00627)
Foreign shares	-0.000167	-0.000165	-0.000144	-9.76e-05
	(0.000203)	(0.000204)	(0.000203)	(0.000191)
Log(age)	-0.00654	-0.00623	-0.0113*	-0.00954*
	(0.00466)	(0.00477)	(0.00592)	(0.00552)
Multi-product firm		-0.000156		
		(0.00405)		
HHI	0.0284***	0.0286**	0.00631	0.0263***
	(0.0110)	(0.0111)	(0.00641)	(0.00977)
Log(Labour prod.)	0.0184	0.0179	0.0158*	0.00825
	(0.0165)	(0.0169)	(0.00931)	(0.0122)
Constant	0	0.380**	0	0
	(0)	(0.173)	(0)	(0)
Industry fixe effects	yes	yes	no	yes
Times fixed effects	yes	yes	yes	yes
Observations	9,186	8,903	9,333	9,333
Number of firms	1,736	1,680	1,740	1,740
No. of instruments	60	61	39	60
Ar(1)	0.000	0.000	0.000	0.000
AR(2)	0.098	0.10	0.09	0.113
Hansen p-val	0.07	0.055	0.242	0.10
Sargan p-val	0.153	0.09	0.430	0.253

Note: constants and year dummies are not reported. Firm sophistication is calculated using prody_tvar.

Standard errors are two-step cluster robust.

Table 7.1;Results: Dependent variable- Log(Firm sophistication)

VARIABLES	model1	model2	Model3	Model4	Model5	Model6
L.Log(firm soph.)	0.786***	0.792***	0.787***	0.795***	0.791***	0.794***
	(0.0450)	(0.0443)	(0.0461)	(0.0427)	(0.0434)	(0.0456)
L2.Log(firm soph.)	0.0891*	0.0932*	0.0891*	0.0726	0.0780	0.0904*
	(0.0509)	(0.0522)	(0.0486)	(0.0539)	(0.0536)	(0.0544)
Log(Employment)	0.0169***	0.0163***	0.0127**	0.0166***	0.0179***	0.0170***
	(0.00599)	(0.00599)	(0.00579)	(0.0059)	(0.00601)	(0.0065)
Captive	-0.0451*	-0.0484**	-0.0479**	-0.0527**	-0.0552**	-0.0518**
	(0.0238)	(0.0242)	(0.0241)	(0.0246)	(0.0247)	(0.0257)
Hierarchical	-0.00369	-0.00654	-0.00856	-0.0141	-0.0210	-0.0173
	(0.0175)	(0.0173)	(0.0178)	(0.0180)	(0.0177)	(0.0184)
Market/Modular	-0.00981	-0.0139	-0.0110	-0.0155	-0.0188	-0.0196
	(0.0203)	(0.0198)	(0.0192)	(0.0198)	(0.0197)	(0.0197)
Log(age)	-0.0166**	-0.019***	-0.0153**	-0.015***	-0.015***	-0.0141**
	(0.00670)	(0.00659)	(0.00645)	(0.00598)	(0.00594)	(0.00621)
Log(R&D)		0.00154*	0.00178**	0.000810	0.000553	0.000642
		(0.00086)	(0.000890)	(0.00092)	(0.000941)	(0.00093)
Log(Labour prod.)			0.0114	-0.000639	-0.000837	-0.000214
			(0.00952)	(0.0125)	(0.0124)	(0.0130)
Multi-prod. firm				-0.00347	-0.00423	-0.00409
				(0.00403)	(0.00406)	(0.00426)
HHI				0.0287***	0.0294***	0.0263**
				(0.0101)	(0.0102)	(0.0107)
Foreign Shares					-0.000115	-0.000117
					(0.00207)	(0.00020)
Constant	0	0.391**	0.490***	0.390***	0	0.332**
	(0)	(0.167)	(0.178)	(0.140)	(0)	(0.156)

Industry fixed effects	no	no	no	yes	yes	yes
Observations	9,383	9,383	9,383	9,097	9,047	9,047
Number of firmid	1,757	1,757	1,757	1,701	1,684	1,684
No. of instruments	41	44	49	70	73	64
Ar(1)	0.000	0.000	0.000	0.000	0.000	0.000
AR(2)	0.11	0.138	0.114	0.081	0.100	0.132
Hansen	0.51	0.540	0.574	0.522	0.527	0.170
Sargan	0.79	0.760	0.704	0.523	0.561	0.216

Note: : year dummies are not reported. Firm sophistication is calculated using prody_tvar. Standard errors are two-step cluster robust.

7.2 Robustness checks

Table 7.2 carries out some robustness checks to check the consistency of the results obtained using the governance variables. In model 1, we change the specification of the model. In model 2, we use a R&D dummy instead of R&D intensity to capture s firm's innovative capacity. High R&D is equal to 1 for all firms whose share of R&D expenditure in total sales is above the median value. In model 3, we use log of deflated sales as a proxy for size, instead of using log of employment. In model 4 ,we add an industry control for high technology sectors and in model 5, we use a different lag specification for model 4. We find that the coefficient on the variable Captive firm remains negative and significant and comparable to the results obtained in section 7.1. From table 7.1 and 7.2, we also observe that the coefficient on Hierarchical firm and Market/Modular firm is also negative but not significant.

Table 7.2: Dependent variable-Log(Firm sophistication)

VARIABLES	(1) model1	(2) model2	(3) model3	(4) model4	(5) model5
L.Log(firm soph.)	0.800***	0.796***	0.799***	0.800***	0.797***
	(0.0486)	(0.0465)	(0.0439)	(0.0420)	(0.0418)
L2.Log(firm soph.)	0.0990*	0.101**	0.0630	0.0648	0.0635
	(0.0517)	(0.0496)	(0.0549)	(0.0540)	(0.0529)
Log(R&D)	0.00179*		0.000792	0.000784	
	(0.000932)		(0.000930)	(0.000928)	
High R&D		0.0238**			
		(0.0117)			
Captive firm	-0.0493*	-0.0493*	-0.0764***	-0.0550**	-0.0522**
	(0.0260)	(0.0253)	(0.0240)	(0.0243)	(0.0243)
Hierarchical firm	-0.0104	-0.0103	-0.0238	-0.0127	-0.0104
	(0.0187)	(0.0187)	(0.0188)	(0.0179)	(0.0182)
Market/Modular firm	-0.0164	-0.0162	-0.0252	-0.0164	-0.0110
	(0.0200)	(0.0198)	(0.0196)	(0.0197)	(0.0205)
Multi-product firm	-0.00612	-0.00615	-0.00271	-0.00394	-0.00378
	(0.00442)	(0.00443)	(0.00424)	(0.00401)	(0.00403)
Log(age)	-0.0136**	-0.0131*	-0.0125**	-0.0175***	-0.0154**
	(0.00663)	(0.00718)	(0.00607)	(0.00564)	(0.00614)
HHI	0.00833	0.00815	0.0307***	0.0271***	0.0296***
	(0.00676)	(0.00672)	(0.0100)	(0.00986)	(0.0102)
Log(Labour prod.)	0.00846	0.00944	-0.0231	-0.00616	-0.00382
	(0.0103)	(0.0111)	(0.0151)	(0.0113)	(0.0128)
LLog(employment)	0.0132**	0.0124*		0.0186***	0.0187***
	(0.00662)	(0.00660)		(0.00574)	(0.00574)
Firm size			0.0156**		
			(0.00663)		
High tech. sector				0.0593**	0.0381*
				(0.0243)	(0.0218)
Constant	0	0	0	0.363***	0
	(0)	(0)	(0)	(0.136)	(0)
Industry fixed effects	No	No	Yes	Yes	Yes
Time fixed effects	yes	yes	yes	yes	yes
Observations	9,097	9,097	9,097	9,097	9,097
Number of firms	1,701	1,701	1,701	1,701	1,701

Note: constants and year dummies are not reported. Firm sophistication is calculated using prody_tvar.

8. Conclusions

By upgrading products, developing country supplier firms can climb up the value-chain ladder to higher value-added and more rent generating activities. Through our empirical investigation of Indian firms, we have attempted to identify some key drivers that can help firms transition from production of low sophisticated goods to better and more complex products. In particular, we are interested in how Global Value Chains can impact product upgrading.

To capture product upgrading at the firm-level, the study combined two databases; Prowess- the firm-level dataset for India- and WITS (UNCOMTRADE). First, we estimated Hausmann's Sophistication Index (2007) for each 4-digit HS product using product-level export data and GDP per capita data for different countries from WITS. Next, we matched products Prowess with products in WITS, and used product-level sales data to calculate the average extent of Indian firms' product sophistication.

Then, to analyse the impact of GVCs on firm-level product-sophistication, we set out three research questions; 1) What is the impact of participating in GVCs on firm-level product upgrading? 2) For GVC firms, does increasing integration lead to product upgrading and 3) Does governance matter for product upgrading?

To answer these questions, we used an unbalanced panel of Indian manufacturing firms for the period 2000/2001-2014/2015, and employed methodologies of system GMM to deal with issues of omitted variables bias, reverse causality and endogeneity. We checked the validity of theGMM estimations using Hansen's test of over-identifying restrictions, Difference-in-Hansen's tests and auto-correlation tests. We also perform a variety of checks and sensitivity analysis to ensure that our results are robust.

We find that on average Indian GVC firms have a higher sophistication level than non-GVC firms. Importing intermediates allows firms to learn from the technology embodied in foreign inputs, while exporting to more advanced lead firms can lead to flow of knowledge, skills and training. Since GVC firms are two-way traders, they learn by both importing and exporting, and the resulting complementarities enables them to produce more sophisticated goods than one-way traders or domestic firms.

Using different vertical specialisation measures as a proxy for magnitude of GVC participation, our results find that increasing integration into GVCs also positively and significantly impacts firm-level sophistication. However, developing country supplier firms should not blindly race to link since the GVC governance structure can crucially shape suppliers' upgrading opportunities. For the case of Indian firms, we find empirical evidence that firms linking into Captive chains produce significantly less sophisticated goods than firms linking into Relational chains. Captive firms operate under high power asymmetries and may face the risk of being 'locked-in' by the Lead firm.

Along with external factors such as participation and governance in GVCs, firms' internal efforts into upgrading are also important. We find that firms can significantly improve their product-sophistication level by investing in their innovative capacity. This would allow them to better absorb and implement foreign knowledge or technology. We also find empirical evidence of firm characteristics such as firm age and size, as well as industry concentration, playing important roles in determining extent of firm sophistication.

The Indian case study does render empirical support to the hypothesis that benefits by linking into GVCs are not automatic or uniform. Indian GVC firms that internally invested in developing their own ICT competence, innovative capacity and share of skilled labour produced significantly more sophisticated goods than other Indian GVC firms. Our study therefore has important policy implications, not only in terms of linking into GVCs but also how to 'gainfully' link.

REFERENCES

Albornoz, Facundo, Darío Milesi, and Gabriel Yoguel. "New Economy in old sectors: some issues coming from two production networks in Argentina."DRUID summer conference. 2002.

Altomonte, Carlo, and Gábor Békés. "Trade complexity and productivity."IEHAS Discussion Papers 914 (2009).

Alvarez, Roberto, and Ricardo A. Lopez. "Exporting and performance: evidence from Chilean plants." Canadian Journal of Economics/Revue canadienne d'économique 38.4 (2005): 1384-1400.

Amador, Joao LM, and Sónia Cabral. "Global value chains: surveying drivers and measures." (2014).

Amiti, Mary, and Donald R. Davis. "Trade, firms, and wages: Theory and evidence." The Review of economic studies 79.1 (2012): 1-36.

Amiti, Mary, and Jozef Konings. "Trade liberalization, intermediate inputs, and productivity: Evidence from Indonesia." The American Economic Review 97.5 (2007): 1611-1638.

Andersson, Martin, Hans Lööf, and Sara Johansson. "Productivity and international trade: Firm level evidence from a small open economy." Review of world economics 144.4 (2008): 774-801.

Ando, Mitsuyo, and Fukunari Kimura. "The formation of international production and distribution networks in East Asia." International Trade in East Asia, NBER-East Asia Seminar on Economics, Volume 14. University of Chicago Press, 2005.

Athukorala, Prema-chandra and Jayant Menon. 2010. Global Production Sharing, Trade Patterns and Determinants of Trade Flows. Working Papers in Trade and Development. No. 2010/06, July, Australian National University.

Athukorala, Prema-chandra, and Nobuaki Yamashita. "Production fragmentation and trade integration: East Asia in a global context." The North American Journal of Economics and Finance 17.3 (2006): 233-256.

Athukorala, Prema-chandra, and Shahbaz Nasir. "Global production sharing and south-south trade." Indian Growth and Development Review 5.2 (2012): 173-202.

Athukorala, Prema-chandra. 2010. Production Networks and Trade Patterns in East Asia: Regionalization or Globalization?. ADB Working Paper Series on Regional Economic Integration, No. 56, August.

Aw, B-Y., and Amy Ruey-meng Hwang. "Productivity and the export market: A firm-level analysis." Journal of development economics 47.2 (1995): 313-332.

Baldwin, John R., and Wulong Gu. "Trade liberalization: Export-market participation, productivity growth, and innovation." Oxford Review of Economic Policy 20.3 (2004): 372-392.

Baldwin, John Russel, and Beiling Yan. Global value chains and the productivity of Canadian manufacturing firms. Statistics Canada, 2014.

Baldwin, Richard, and Anthony J. Venables. "Spiders and snakes: offshoring and agglomeration in the global economy." Journal of International Economics 90.2 (2013): 245-254.

Baldwin, Richard. Trade and industrialisation after globalisation's 2nd unbundling: How building and joining a supply chain are different and why it matters. No. w17716. National Bureau of Economic Research, 2011.

Banga, Rashmi, and Bishwanath Goldar. "Contribution of services to output growth and productivity in Indian manufacturing: pre-and post-reforms."Economic and Political weekly (2007): 2769-2777.

Banga, Rashmi. "Linking into global value chains is not sufficient: do you export domestic value added contents?." Journal of Economic Integration(2014): 267-297.

Barrientos, Stephanie, and Andrienetta Kritzinger. "Squaring the circle: Global production and the informalization of work in South African fruit exports." Journal of International Development 16.1 (2004): 81-92.

Barrientos, Stephanie. "Contract labour: The 'Achilles heel'of corporate codes in commercial value chains." Development and Change 39.6 (2008): 977-990.

Bell, M. & Pavitt, K. (1993) Technological accumulation and industrial growth: Contrasts between developed and developing countries, Industrial and Corporate Change, 2(2), pp. 157–210.

Bell, Martin, and Michael Albu. "Knowledge systems and technological dynamism in industrial clusters in developing countries." World development27.9 (1999): 1715-1734.

Bernard, Andrew B., et al. "Plants and productivity in international trade." The American Economic Review 93.4 (2003): 1268-1290.

Bernard, Andrew B., et al. "Plants and productivity in international trade." The American Economic Review 93.4 (2003): 1268-1290.

Bernard, Andrew B., J. Bradford Jensen, and Robert Z. Lawrence. "Exporters, jobs, and wages in US manufacturing: 1976-1987." Brookings papers on economic activity. Microeconomics 1995 (1995): 67-119.

Biscourp, Pierre, and Francis Kramarz. "Employment, skill structure and international trade: Firm-level evidence for France." Journal of International Economics 72.1 (2007): 22-51.

Cattaneo, Olivier, Gary Gereffi, and Cornelia Staritz, eds. Global value chains in a postcrisis world: a development perspective. World Bank Publications, 2010.

Cattaneo, Olivier, Gary Gereffi, and Cornelia Staritz, eds. Global value chains in a postcrisis world: a development perspective. World Bank Publications, 2010.

Chen, Xikang, et al. "Domestic value added and employment generated by Chinese exports: A quantitative estimation." China Economic Review 23.4 (2012): 850-864.

Cincera, Michele, and Reinhilde Veugelers. "Young leading innovators and the EU's R&D intensity gap." Economics of Innovation and New Technology22.2 (2013): 177-198.

Clark, Don P. "Country and industry-level determinants of vertical specialization-based trade." International Economic Journal 20.2 (2006): 211-225.

Clerides, Sofronis K., Saul Lach, and James R. Tybout. "Is learning by exporting important? Micro-dynamic evidence from Colombia, Mexico, and Morocco." Quarterly journal of Economics (1998): 903-947.

Coe, Neil M., et al. "'Globalizing'regional development: a global production networks perspective." Transactions of the Institute of British geographers29.4 (2004): 468-484.

Coe, Neil M., Peter Dicken, and Martin Hess. "Global production networks: realizing the potential." Journal of economic geography (2008): lbn002.

Cooke, P., Uranga, M. & Etxebarria, G. (1997) Regional innovation systems: Institutional and organisational dimensions, Research Policy, 26(4–5), pp. 475–491.

Dallas, Mark P. "'Governed'trade: global value chains, firms, and the heterogeneity of trade in an era of fragmented production." Review of International Political Economy 22.5 (2015): 875-909.

Damijan, Jože Pavlič, Sašo Polanec, and Janez Prašnikar. Self-selection, export market heterogeneity and productivity improvements: Firm level evidence from Slovenia. Katholieke Universiteit Leuven, 2004.

Daudin, Guillaume, Christine Rifflart, and Danielle Schweisguth. "Who produces for whom in the world economy?." Canadian Journal of Economics/Revue canadienne d'économique 44.4 (2011): 1403-1437.

De Backer, Koen, Sebastien Miroudot, and Alexandros Ragoussis. "Manufacturing in global value chains." Manufacturing Europe's Future", Bruegel Blueprint Series 21 (2013).

De Loecker, Jan. "Do exports generate higher productivity? Evidence from Slovenia." Journal of international economics 73.1 (2007): 69-98.

Dolan, C.S. and Sutherland, K., 2003, 'Gender and employment in the Kenya horticulture value chain', Globalisation, Production and Poverty Discussion Paper 8, Norwich: University of East Anglia, www.gapresearch.org (accessed October 2003)

Ebenstein, Avraham, et al. "Estimating the impact of trade and offshoring on American workers using the current population surveys." Review of Economics and Statistics 96.4 (2014): 581-595.

Ernst, D. & Kim, L. (2001) Global production networks, knowledge diffusion, and local capability formation: A conceptual framework. East-West Centres working paper no 19, East-West Centre, Honolulu

Fafchamps, Marcel, Said El Hamine, and Albert Zeufack. "Learning to export: Evidence from Moroccan manufacturing." Journal of African Economies 17.2 (2008): 305-355.

Feenstra, Robert C., and Gordon H. Hanson. "The impact of outsourcing and high-technology capital on wages: estimates for the United States, 1979-1990." Quarterly Journal of Economics (1999): 907-940.

Feenstra, Robert C., and J. Bradford Jensen. "Evaluating estimates of materials offshoring from US manufacturing." Economics Letters 117.1 (2012): 170-173.

Fernandes, Ana M., and Alberto Isgut. "Learning-by-doing, learning-by-exporting, and productivity: evidence from Colombia." World Bank policy research working paper 3544 (2005).

Figueiredo, Paulo N. Technological learning and competitive performance. Edward Elgar Publishing, 2001.

Fontagné, Lionel, Michael Freudenberg, and Guillaume Gaulier. "Disentangling horizontal and vertical intra-industry trade." Paris: CEPII(2005): 2005-10.

Gaulier, Guillaume, Françoise Lemoine, and Deniz Ünal-Kesenci. "China’s integration in East Asia: Production sharing, FDI & high-tech trade."Economic Change and Restructuring 40.1-2 (2007): 27-63.

Gereffi, Gary, and Miguel Korzeniewicz. Commodity chains and global capitalism. No. 149. ABC-CLIO, 1994.

Gereffi, Gary, John Humphrey, and Raphael Kaplinsky. "Introduction: Globalisation, value chains and development." IDS bulletin 32.3 (2001): 1-8.

Gereffi, Gary, John Humphrey, and Timothy Sturgeon. "The governance of global value chains." Review of international political economy 12.1 (2005): 78-104.

Gereffi, Gary. "International trade and industrial upgrading in the apparel commodity chain." Journal of international economics 48.1 (1999): 37-70.

Gereffi, Gary. "The global economy: organization, governance, and development." The handbook of economic sociology 2 (2005): 160-182.

Giuliani, Elisa, Carlo Pietrobelli, and Roberta Rabellotti. "Upgrading in global value chains: lessons from Latin American clusters." World development 33.4 (2005): 549-573.

Goldar, Bishwanath. "Employment growth in organised manufacturing in India." Economic and Political weekly (2000): 1191-1195.

Goldberg, Pinelopi K., et al. Imported intermediate inputs and domestic product growth: Evidence from India. No. w14416. National Bureau of Economic Research, 2008.

Görg, Holger, and Aoife Hanley. "Labour demand effects of international outsourcing: Evidence from plant-level data." International Review of Economics & Finance 14.3 (2005): 365-376.

Granovetter, Mark. "Economic action and social structure: The problem of embeddedness." American journal of sociology (1985): 481-510.

Greenaway, David, and Richard Kneller. "Firm heterogeneity, exporting and foreign direct investment." The Economic Journal 117.517 (2007): F134-F161.

Grossman, Gene M., and Esteban Rossi-Hansberg. "Trading tasks: A simple theory of offshoring." The American Economic Review 98.5 (2008): 1978-1997.

Grossman, Gene, and Elhanan Helpman. "Innovation and growth in the world economy." (1991).

Guerrieri, Paolo, and Carlo Pietrobelli. "Old and New Forms of Clustering and Production Networks in Changing Technological Regimes Contrasting Evidence from Taiwan and Italy." Science Technology & Society 11.1 (2006): 9-38.

Hagemejer, Jan, and Marcin Kolasa. "Internationalisation and Economic Performance of Enterprises: Evidence from Polish Firm-level Data." The World Economy 34.1 (2011): 74-100.

Harrison, Ann, and Gordon Hanson. "Who gains from trade reform? Some remaining puzzles." Journal of development Economics 59.1 (1999): 125-154.

Hasan, Rana, Devashish Mitra, and Krishnarajapet V. Ramaswamy. "Trade reforms, labor regulations, and labor-demand elasticities: Empirical evidence from India." The Review of Economics and Statistics 89.3 (2007): 466-481.

Hijzen, Alexander, Tomohiko Inui, and Yasuyuki Todo. "Does Offshoring Pay? Firm-Level Evidence from Japan." Economic Inquiry 48.4 (2010): 880-895.

Hoda, Anwarul, and Durgesh Kumar Rai. "Trade and Investment Barriers affecting International Production Networks in India." ICRIER working paper 281. ICRIER New Delhi, 2014.

Hummels, David, et al. "The wage effects of offshoring: Evidence from Danish matched worker-firm data." The American Economic Review 104.6 (2014): 1597-1629.

Hummels, David, Jun Ishii, and Kei-Mu Yi. "The nature and growth of vertical specialization in world trade." Journal of international Economics 54.1 (2001): 75-96.

Humphrey, John, and Hubert Schmitz. "How does insertion in global value chains affect upgrading in industrial clusters?." Regional studies 36.9 (2002): 1017-1027.

Johnson, Robert C., and Guillermo Noguera. "Accounting for intermediates: Production sharing and trade in value added." Journal of international Economics 86.2 (2012): 224-236.

Jones, Ronald W., and Henryk Kierzkowski. "A framework for fragmentation." (2001).

Kadarusman, Yohanes, and Khalid Nadvi. "Competitiveness and technological upgrading in global value chains: evidence from the Indonesian electronics and garment Sectors." European Planning Studies 21.7 (2013): 1007-1028.

Kaplinsky, Raphael, and Mike Morris. A handbook for value chain research. Vol. 113. Ottawa: IDRC, 2001.

Kaplinsky, Raphael. "Global Value Chains: Where They Came From, Where They Are Going and Why This Is Important." Innovations, Knowledge and Development Working paper (68) (2013).

Kaplinsky, Raphael. "Globalisation and unequalisation: What can be learned from value chain analysis?." Journal of development studies 37.2 (2000): 117-146.

Kasahara, Hiroyuki, and Beverly Lapham. "Productivity and the decision to import and export: Theory and evidence." Journal of International Economics89.2 (2013): 297-316.

Keane, Jodie. "Global value chain analysis." (2014).

Kien, Tran Nhuan, and Yoon Heo. "Doi moi policy and socio-economic development in Vietnam, 1986–2005." International area studies review 11.1 (2008): 205-232.

Kimura, Fukunari and Ayako Obashi. 2011. Production Networks in East Asia: What We Know So Far, ADBI Working Paper Series, No. 320, Tokyo: Asian Development Bank Institute.

Kishimoto, Chikashi. "Clustering and upgrading in global value chains: the Taiwanese personal computer industry." Chapters (2004).

Knorringa, Peter, and Lee Pegler. "Globalisation, firm upgrading and impacts on labour." Tijdschrift voor economische en sociale geografie 97.5 (2006): 470-479.

Koopman, Robert, et al. Give credit where credit is due: Tracing value added in global production chains. No. w16426. National Bureau of Economic Research, 2010.

Krugman, Paul R. Geography and trade. MIT press, 1991.

Kumar, R. Sudhir, and MH Bala Subrahmanya. "Influence of subcontracting on innovation and economic performance of SMEs in Indian automobile industry." Technovation 30.11 (2010): 558-569.

L. Bazan, L. Navas-Aleman, The underground revolution in the Sinos Valley: A comparison of upgrading in global and national value chain, (2004)

Lakhani, Tashlin, Sarosh Kuruvilla, and Ariel Avgar. "From the firm to the network: Global value chains and employment relations theory." British Journal of Industrial Relations 51.3 (2013): 440-472.

Lall, S. (1992) Technology capabilities and industrialization, World Development, 20(2), pp. 165–186.

Lundvall, B. A. (2007) National innovation systems: Analytical concept and development tool, Industry and Innovation, 14(1), pp. 95–119.

Martins, Pedro S., and Yong Yang. "The impact of exporting on firm productivity: a meta-analysis of the learning-by-exporting hypothesis."Review of World Economics 145.3 (2009): 431-445.

Melitz, Marc J. "The impact of trade on intra-industry reallocations and aggregate industry productivity." Econometrica 71.6 (2003): 1695-1725.

Mion, Giordano, and Linke Zhu. "Import competition from and offshoring to China: A curse or blessing for firms?." Journal of International Economics89.1 (2013): 202-215.

Mitra, Devashish, and Priya Ranjan. "Offshoring and unemployment: The role of search frictions labor mobility." Journal of International Economics 81.2 (2010): 219-229.

Mudambi, Ram. "Location, control and innovation in knowledge-intensive industries." Journal of economic Geography 8.5 (2008): 699-725.

Muûls, Mirabelle, and Mauro Pisu. "Imports and Exports at the Level of the Firm: Evidence from Belgium." The World Economy 32.5 (2009): 692-734.

Nadvi, K. (2011) Labour standards and technological upgrading: Competitive challenges in the global football industry, International Journal of Technological Learning, Innovation and Development, 4(1–3), pp. 235–257.

Nadvi, Khalid, et al. "Vietnam in the global garment and textile value chain: impacts on firms and workers." Journal of international development 16.1 (2004): 111-123.
Navaretti, Giorgio Barba, Anthony Venables, and Frank Barry. Multinational firms in the world economy. Princeton University Press, 2004.
Nelson, R. (1993) National Innovation Systems: A Comparative Analysis (New York: University Press).
Ng, Francis, and Alexander J. Yeats. Good governance and trade policy: are they the keys to Africa's global integration and growth?. Vol. 2038. World Bank Publications, 1999.
Park, Albert, et al. "Exporting and firm performance: Chinese exporters and the Asian financial crisis." The Review of Economics and Statistics 92.4 (2010): 822-842.
Pietrobelli, Carlo, and Roberta Rabellotti. "Global value chains meet innovation systems: are there learning opportunities for developing countries?." World Development 39.7 (2011): 1261-1269.
Ponte, Stefano, and Peter Gibbon. "Quality standards, conventions and the governance of global value chains." Economy and society 34.1 (2005): 1-31.
Puzzello, Laura. "A proportionality assumption and measurement biases in the factor content of trade." Journal of International Economics 87.1 (2012): 105-111.
Revenga, Ana. "Employment and wage effects of trade liberalization: the case of Mexican manufacturing." Journal of labor Economics 15.S3 (1997): S20-S43.
Robinson, P., and H. Rainbird. "Beyond the Organisation and the Nation State… and the Employment Relationship, Same as it ever was? Global Value Chains and the Challenges for Industrial Relations Research." British Journal of Industrial Relations 50th Anniversary Conference, London School of Economics. 2011.
Sato, Yuri, et al. "Capability matrix: a framework for analyzing capabilities in value chains." (2009).
Schmitt, Alexander, and Johannes Van Biesebroeck. "Governing supply relationships: Evidence from the automotive sector." (2015).
Schmitz, Hubert, and Peter Knorringa. "Learning from global buyers." Journal of development studies 37.2 (2000): 177-205.
Schor, Adriana. "Heterogeneous productivity response to tariff reduction. Evidence from Brazilian manufacturing firms." Journal of Development Economics 75.2 (2004): 373-396.
Sen, Kunal. "International Trade and Manufacturing Employment Outcomes in India." A com (2008).
Sen, Rahul and Sadhana Srivastava. 2011. Integrating into Asia's International Production Networks: Challenges and prospects for India. Studies in Trade and Investment 75. India: A New Player in Asian Production Networks?, Edited by Witada Anukoonwattaka and Mia Mikic. UNESCAP.
Serti, Francesco, and Chiara Tomasi. "Self-selection and post-entry effects of exports: Evidence from Italian manufacturing firms." Review of World Economics 144.4 (2008): 660-694.
Sydor, Aaron. Global Value Chains: Impacts and Implications. Minister of Public Works and Government Services, 2011.
Taylor, Phil, Kirsty Newsome, and Al Rainnie. "'Putting labour in its place': global value chains and labour process analysis." Competition & Change17.1 (2013): 1-5.
Topalova, Petia, and Amit Khandelwal. "Trade liberalization and firm productivity: The case of India." Review of economics and statistics 93.3 (2011): 995-1009.
Van Biesebroeck, Johannes. "Exporting raises productivity in sub-Saharan African manufacturing firms." Journal of International economics 67.2 (2005): 373-391.
Veeramani, C. "Anatomy of India's Merchandise Export Growth, 1993-94 to 2010-11." Economic and Political Weekly 47.1 (2012): 94-104.
Verhoogen, Eric A. "Trade, quality upgrading and wage inequality in the Mexican manufacturing sector." Vol (2007).
Vogel, Alexander, and Joachim Wagner. "Higher productivity in importing German manufacturing firms: self-selection, learning from importing, or both?." Review of World Economics 145.4 (2010): 641-665.
Wagner, Joachim. "Exports and productivity: A survey of the evidence from firm-level data." The World Economy 30.1 (2007): 60-82.
Wagner, Joachim. 2002. The Causal Effects of Exports on Firm Size and Labor Productivity: First evidence from a Matching Approach, Economics Letters, Vol 77(2), pp. 287-92.

Wood, G. T. and Lane, C. (2011). 'Institutions, change and diversity'. In C. Lane and G. T. Wood (eds.), Capitalist Diversity and Diversity with Capitalism. London: Routledge, pp. 1–31.

Yeats, Alexander J. "Just how big is global production sharing?." World Bank Policy Research Working Paper 1871 (1998).

Appendix

A1. NIC industries our sample

NIC industry code	NIC Industry name
10	Food products
11	Beverages
12	Tobacco products
13	Textiles
14	Wearing apparel
15	Leather and products
16	Wood and products
17	Paper and products
18	Printing and reproduction of media
20	Chemical products
21	Pharmaceuticals
22	Rubber and plastics
23	Non-metallic minerals
24	Basic metals
25	Fabricated metal products
26	Computer, electronics and optical products
27	Electrical equipment
28	Machinery and equipment
29	Motor vehicles, trailer, semi-trailer
30	Other transport
31	Furniture
32	Other manufactures

A.2 Matching Prowess products with HS products

Since the product classification in CMIE prowess cannot be directly linked to any standard international classification, we undertake a procedure of matching to re-classify all prowess products according to HS classification.

To match product-sophistication data at HS 4-digit level to prowess ID codes, we assign correspondences between the two by hand, exploiting the fact that both prowess's classification and HS classification are closely related to ISIC classification. This means that both classifications have similar names and ordering of products. To start of, we map every 14 digit prowess code to HS at 2, 4 or 6 digit level. We find that we are able to match around 80% of the products to four-digit HS classification. While matching to six-digit would be ideal, Indian firms do not report detailed product names and therefore are best matched to HS at four digit level. Matching to six-digit would lead to loss of accuracy, and is therefore avoided.

Our mapping uses both product names in the two datasets as well as numerical ordering to generate correspondences. Tables 1,2 and 3 show the matching procedure with the example of HS code 61: Article of apparel & clothing access, knitted or crocheted. We find that in Prowess (see table 1), this sector is given the code 60. Consider the PRID 6070101000000, which we have assigned a standardised name; 'Men's overcoats etc., knitted or crocheted', following the standardisation procedure explained above. The PRID 6070101000000 is followed by 6070102000000 –'Women's overcoats' and 6070103000000 –'Men's suit's etc'. These names can be very easily matched to HS code 6101, 6102 and 6103, which follow the same ordering and assign extremely similar names. See tables 1 and 2.

1.1. Table 1: Prowess product names under articles of Apparel, clothing accessories, knitted or crocheted.

14 digit PRID	**Standardized prowess name**
60701010000000	Men's overcoats, etc. knitted or crocheted
60701020000000	Women's overcoats, etc. knitted or crocheted
60701030000000	Men's suits, trousers, etc. knitted or crocheted
60701040000000	Women's suits, dresses etc. knitted or crocheted
60701050000000	Men's shirts, knitted or crocheted.

1.2. Table 2: HS Product names under articles of Apparel, clothing accessories, knitted or crocheted.

HS CODE	HS code name
6101	Men's or boys' overcoats, car-coats, etc. knitted or crocheted
6102	Women's or girls' overcoats, car-coats, knitted or crocheted
6103	Men's or boys' suits, ensembles, jackets, blazers, knitted or crocheted
6104	Women's or girls' suits, ensembles, dresses etc., knitted or crocheted
6105	Men's or boys' shirts, knitted or crocheted. knitted or crocheted

It is quite clear that the matching from Prowess to HS, in the case of Articles of Apparel, clothing accessories, knitted or crocheted, will be as follows:

1.3. Table 3: Example for matching HS products to Prowess products

14 digit PRID	**Standardized prowess name**	**Matched to HS code**
60701010000000	Men's overcoats, etc. knitted or crocheted	6101
60701020000000	Women's overcoats, etc. knitted or crocheted	6102
60701030000000	Men's suits, trousers, etc. knitted or crocheted	6103
60701040000000	Women's suits, dresses etc. knitted or crocheted	6104
60701050000000	Men's shirts, knitted or crocheted.	6105

However, the mapping is not merely matter of harmonising names as it is not necessary that a prowess code maps to only one 4 digit HS. Consider the example of 'silk and silk textiles' in table 4 and 5. While HS distinguishes between three types of silk yarn (5004, 5005,5006) in table 5, Prowess has only one category of silk yarn (table 4). Therefore, when matching 'silk yarn', we match to the three HS codes (table 6). On the other hand CMIE distinguishes between woven and processed fabrics of silk while HS lumps them under the same category- 5007; "woven fabrics of silk or of silk waste". Therefore, we match both 'woven fabrics of silk' and 'Silk fabrics processed' to 5007 (table 6). Using this procedure, we match 2407 product codes in Prowess to Four-Digit HS codes.

1.4. Table 4: Prowess product names under 'silk and silk textiles'

PRID	**PRID NAME**
601010000000000	Silk worm cocoons
601020000000000	Raw silk
601030000000000	Silk waste
601040000000000	Silk yarns
601050000000000	Woven fabrics of silk
601060000000000	Silk fabrics, processed

1.5. Table 5: HS product names under 'silk and silk textiles'

HS code	HS name
5001	Silk-worm cocoons suitable for reeling
5002	Raw silk (not thrown)
5003	Silk waste (including cocoons unsuitable for reeling, yarn waste and garneted stock)
5004	Silk yarn (other than yarn spun from silk waste)not put up for retail sale
5005	Yarn spun from silk waste, not put up for retail sale
5006	Silk yarn and yarn spun from silk waste, put up for retail sale; silk-worm gut
5007	Woven fabrics of silk or of silk waste

1.6. Table 6: Matching for silk and silk textiles

PRID	Product name	Matched to Hs code
60101000000000	Silk worm cocoons	5001
60102000000000	Raw silk	5002
60103000000000	Silk waste	5003
60104000000000	Silk yarn	5004,5005,5006
60105000000000	Woven fabrics of silk	5007
60106000000000	Silk fabrics processed	5007

Having matched the codes to HS category, we check and validate our matching using Barrows's (2016) PRID-HS concordance [19]. If a product in Prowess is matched to multiple HS codes, such as in the example above, we simply take the average sophistication level of all the HS codes to reach at the sophistication level of the PRID reported by the firm. So for the

[19] This concordance was obtained from Geoffery Barrows, used in 'Does trade make firms cleaner?' Theory and Evidence from Indian Manufacturing (Barrows, 2014)

case of 60104000000000, the sophistication level of the PRID would be the simple average of sophistication level of HS 5004, 5005 and 5006.

A.3 Countries with no missing data on exports or GDP per capita (PPP) in the period 2001-2014

Albania	Italy	Belize	Malaysia
Algeria	Jamaica	Benin	Maldives
Argentina	Japan	Bolivia	Malta
Armenia	Jordan	Botswana	Mauritius
Australia	Kazakhstan	Brazil	Mexico
Austria	Korea, Rep.	Bulgaria	Moldova
Azerbaijan	Latvia	Burundi	Morocco
Bahamas, The	Lebanon	Cambodia	Mozambique
Bahrain	Lithuania	Canada	Namibia
Barbados	Luxembourg	Central African Republic	Netherlands
Belarus	Madagascar	Chile	New Zealand
Belgium	Malawi	China	Nicaragua
Finland	Sao Tome and Principe	Colombia	Niger
France	Saudi Arabia	Cote d'Ivoire	Norway
Gambia, The	Senegal	Croatia	Oman
Georgia	Singapore	Cyprus	Paraguay
Germany	Slovak Republic	Czech Republic	Peru
Greece	Slovenia	Denmark	Philippines
Guatemala	South Africa	Dominican Republic	Poland
Guyana	Spain	Ecuador	Portugal
Hong Kong, China	Sri Lanka	Egypt, Arab Rep.	Romania
Hungary	Suriname	El Salvador	Russian Federation
Iceland	Sweden	Estonia	Rwanda
India	Switzerland	Ethiopia	Samoa
Indonesia	Tanzania	Ireland	Thailand
Israel	United States	Tonga	Uganda
United Arab Emirates	Uruguay	Turkey	Ukraine
United Kingdom	Vietnam	Zambia	

A.4 construction of different sophistication indicators

Indicator type	Indicator	Avg. Export over time	Avg. GDP over time	Formula
Hausmann's PRODY indicator	Prody_tvar	NO	NO	$\sum_c \frac{X^k_{ct}/X^\bullet_{ct}}{\sum_c (X^k_{ct}/X^\bullet_{ct})} Y_{ct}$
Variant of PRODY	Prody_mtrd	YES	NO	$\sum_c \frac{\overline{X}^k_c/\overline{X}^\bullet_c}{\sum_c (\overline{X}^k_c/\overline{X}^\bullet_c)} Y_{ct}$
Variant of PRODY	Prody_mgdp	NO	YES	$\sum_c \frac{X^k_{ct}/X^\bullet_{ct}}{\sum_c (X^k_{ct}/X^\bullet_{ct})} \overline{Y}_c$
Variant of PRODY	Prody_Mean 1	YES	YES	$\sum_c \frac{\overline{X}^k_c/\overline{X}^\bullet_c}{\sum_c (\overline{X}^k_c/\overline{X}^\bullet_c)} \overline{Y}_c$
Variant of PRODY	Prody_Mean 2	NO	NO	$\overline{PRODY\_\text{tvar}}^k_t$ i.e. average of prody_tvar over time.
Lall's soph indicator	$Lall^k$	no	no	$100 \cdot \left(\frac{SI^k_t - \min(SI^{k\in K}_t)}{\max(SI^{k\in K}_t) - \min(SI^{k\in K}_t)}\right)$ with $SI^k_t = \sum_{g=1}^{G=10}\left(\overline{Y_{gt}} \cdot \frac{X^k_{gt}}{X^k_{\bullet t}}\right)$
Michaley soph. indicator	Mic1^k	no	no	$\sum_c Y_{ct}\left(\frac{X^k_{ct}}{X^k_{\bullet t}}\right)$
Michaley soph. indicator	Mic2^k	yes	yes	$\widehat{\beta^k}$, taken from $\frac{\overline{X^k_c}}{\overline{X^k_\bullet}} = \alpha + \beta^k \overline{Y_c} + \epsilon^k_c, \quad \forall k$

1.7. Note: Y represent GDP per capita, X represents exports.

A4.1 Ranking of different product sophistication indicators

Product name	HS code	Prody_ mean2	Prod lall	Prody mean 1	Prody_ tvar 2014	Prody_m gdp	Prody_ mtrd	Mic 1
Pharmaceutical products.	30	33836.069	1	1	1	1	1	1
Organic chemicals.	29	32157.892	14	3	2	2	3	11
Photographic or cinematographic goods.	37	32153.765	8	2	3	3	2	6
Nickel and articles thereof.	75	31897.361	4	4	10	4	4	2
Clocks and watches and parts thereof.	91	30515.018	24	10	5	5	10	22
Plastics and articles thereof.	39	30444.344	11	5	6	6	5	10
Nuclear reactors, boilers, mch. & mechanical appliance; parts	84	30401.056	19	6	7	7	5	17
Impregnated, coated, cover/laminated textile fabric etc.	59	30258.7	32	7	4	8	6	27
Miscellaneous chemical products.	38	30204.959	7	8	8	9	8	7
Optical instruments, photo, measurement, checking, precision, etc.	90	29386.105	12	12	11	10	12	12

A.4.2. Spearman rank correlation test

H0: correlation coefficient is zero. P-val indicated in brackets. Pval=0 implies H0 can be rejected.

Indicators	Prody_tvar	MIC2	MIC1	Lall	Prody_mtrd	Prody_mean1
MIC2	0.7288	1				
	(0)					
MIC1	0.7413	0.9585	1			
	(0)	(0)				
Lall	0.6604	0.9058	0.9338	1		
	(0)	(0)	(0)			
Prody_mtrd	0.963	0.7387	0.7295	0.6429	1	
	(0)	(0)	(0)	(0)		
Prody_mean1	0.9487	0.7619	0.755	0.6956	0.9695	1
	(0)	(0)	(0)	(0)	(0)	
Prody_mgdps	0.9732	0.736	0.7524	0.7013	0.9219	0.9572
	(0)	(0)	(0)	(0)	(0)	(0)
Prody_mean2	0.9524	0.7625	0.7556	0.6965	0.9653	0.995
	(0)	(0)	(0)	(0)	(0)	(0)

A.5 Construction of Governance

Governance	Construction
Share of skilled labour in total labour	Managerial remuneration/ total labour compensation
High skill level dummy	=1 if Share of skilled labour in total labour is above median value in two digit industry
Software assets	Gross software assets/ sales
Technology assets	Gross plant, machinery, computers and electrical installations/ sales
Infrastructure assets	Transport, communication equipment and infrastructure / sales
ICT index	Supplier competence or ICT index created using PCA on data related to software assets, technology assets and infrastructure assets. Retained component 1 following Kaiser's rule of retaining components with eigen vectors >1. Component 1 explained about 56% percent of variation in the variables. Using weights in component 1 , 0.55 for software assets, 0.64 for infrastructure assets and 0.52 for technology assets, the ICT index is created. Validity is checked using KMO stat (0.67) and Bartlett's test of spherecity with H0; no inter-correlation between variables. (p –val= 0.000).
High ICT dummy	=1 if ICT index of firm greater than median value in two digit industry
Market share of firm	Firms total sales/ sales in its five digit industry
High Market share	=1 if Market share of firm > median value of market share in its five digit industry
Market/ modular	=1 if High skill level dummy=0 and High ICT dummy=1
Market	=1 if Market/ Modular=1 and High Market share=0
Modular	=1 if Market/ Modular=1 and High Market share=1
Relational	=1 if High skill level dummy=1 and High ICT dummy=1
Captive	=1 if High skill level dummy=0 and High ICT dummy=0
Hierarchy	=1 if High skill level dummy=1 and High ICT dummy=0

A.6 Construction of Technological Capability

Innovative TC	**R&D intensity** = Share of R&D expenditure (total capital and current account expenditures) in total sales value of the firm.
Operational TC	**Technical know-how** = (expenditures on royalty+ technical knowhow fee + license fee)/ total sales value **Import of technology** = import of capital goods/ sales **TC index** created by PCA using data on Royalty, technical know-how fee and import of capital goods. Following Kaiser's rule, Component 1 is retained because its eigenvalue was found to be greater than 1. Component 1 explain around 45 % of variation in data and weighs heavily on import of capital goods (0.69), royalty (0.68) and less on technical-fee (0.23). Validity of PCA confirmed using Kaiser-Meyer-Olkin measure of sampling adequacy (kmo statistics >0.5) and Barlett's test of inter-correlation between variables (p stat =0.000) Index is scaled from 0 to 1.

A.7 Summary Statistics

Variable	No. of observations	Mean	Std. Dev.	Min	Max
Sales	69768.000	26.311	123.113	0.010	4773.070
Output	69768.000	24.829	118.758	0.010	4310.500
GVA	69762.000	11.874	64.531	0.000	2627.620
Export intensity	69660	12.35	23.32	0.000	100
Foreign ownership					
Foreign shares	68846.000	2.169	10.429	0.000	97.090
Foreign firms	68846.000	0.075	0.263	0.000	1.000
Types of firms					
Domestic firms	69768.000	0.396	0.489	0.000	1.000
Only exporters	69768.000	0.111	0.314	0.000	1.000
Only importers	69768.000	0.164	0.370	0.000	1.000
GVC	69768.000	0.328	0.470	0.000	1.000
Technological capability					
Innovative TC	69756.000	0.189	1.434	0.000	96.810
Operational TC	69768.000	0.508	6.434	0.000	427.920
Firm labour					
Labour	69058.000	1204.8	6309.565	1.000	504601.500
Labour productivity	69058.000	0.017	0.083	0.000	8.786
Product characteristics					
Multi-product firms	60901.000	0.507	0.500	0.000	1.000
Product sophistication level					

Prody_tvar	56129.000	27.734	9.689	0.000	100.000
Industry concentration					
HHI	69768.000	0.195	0.210	0.014	1.000
Firm size					
Big firm	69768.000	0.456	0.498	0.000	1.000
Medium firm	69768.000	0.314	0.464	0.000	1.000
Small firm	69768.000	0.230	0.421	0.000	1.000

A.8 Average sophistication across types of firms

	Prody_ tvar	**Mic2**	**Mic1**	**Lall**	**Prody_ mtrd**	**Prody_mean1**	**Prody_ mgdps**	**Prody_mean 2**
Non-GVC (mean)	21784.5	0.00186	29261	50.08	21479.4	21527.8	21831.45	21847.89
GVC (mean)	24556.15	0.00218	30479	52.153	24251.3	24216.1	24511.68	24517.46
Non- GVC (median)	22421.8	0.00208	30185	51.338	22085.6	22434.8	22546.3	22772.2
GVC (median)	25521.6	0.00255	31688	53.506	25196.9	25624.6	25658.5	25911.6
Domestic owned	22474.93	0.00191	29476	50.404	22169.9	22182.9	22484.8	22498.92
Foreign owned	25912.99	0.00264	32026	55.373	25579.3	25731.0	26050.53	26011.6
Single product firms	22989.51	0.00201	30003	51.473	22664.1	22672.7	22997.65	22994.49
Multi product firms	22155.33	0.00185	29185	49.848	21883.5	21912.8	22184.36	22211.35
High Oper.TC	24550.69	0.00240	31136	53.525	24187.1	24200.1	24547.66	24537.9
Low oper.TC	22232	0.00185	29257	50.001	21944.8	21965.1	22251.65	22271.1
High innov. TC	24798	0.00252	31496	54.333	24450.2	24489.4	24835.57	24803.31
Low innov. TC	22157.54	0.00181	29149	49.758	21865.6	21878.0	22164.97	22191.26

Note: Using t-test, we find that the means of each sophistication indicator is statistically significantly different across the categories considered; GVC-Non-GVC, Domestic-Foreign etc. Difference in means was found to be statistically significant at 1% across categories for all variables.

A.9 Results for propensity score specification and balancing property

1.8. Propensity score specification for 2012

The treatment is GVC.

GVC	Frequency	Percent	Cum.
0	3,246	67.79	67.79
1	1,542	32.21	100

Estimation of propensity score

GVC	Coeifficient	Std.Error	z	P>|z|	[95% confidence interval]	
L.log(R&D intensity)	0.0695555	0.0073363	9.48	0	0.0551765	0.0839345
L.Smallfirm	-1.439045	0.1448853	-9.93	0	-1.723015	-1.155075
L.Bigfirm	0.9102214	0.0858771	10.6	0	0.7419054	1.078537
L.Foreign firm	0.5841801	0.1369723	4.26	0	0.3157193	0.8526409
L. log(Labour prod)	0.1348936	0.0489862	2.75	0.006	0.0388825	0.2309048
L.log(Age)	0.1088703	0.0549987	1.98	0.048	0.0010747	0.2166658

Note: industry dummies are added but coefficients are not reported. Number of observations= 4253, LR Chi2(25)= 1064.24, Prob> chi1=0.0000, Pseudo R sqrd. = 0.19. The region of common support is [.01411234, .89084262]

```
Description of the estimated propensity score
in region of common support

                  Estimated propensity score
-------------------------------------------------------------
      Percentiles      Smallest
 1%    .0242914       .0141123
 5%    .0526933       .0143426
10%    .0687961       .0145548       Obs                4224
25%    .1332433        .014666       Sum of Wgt.        4224

50%    .3059817                      Mean           .3420494
                        Largest      Std. Dev.      .2270346
75%    .4971664       .8860056
90%    .6831677       .8872871       Variance       .0515447
95%    .7703075       .8877607       Skewness       .4498236
99%    .8304315       .8908426       Kurtosis       2.152779


******************************************************
Step 1: Identification of the optimal number of blocks
Use option detail if you want more detailed output
******************************************************


The final number of blocks is 10

This number of blocks ensures that the mean propensity score
is not different for treated and controls in each blocks

**********************************************************
Step 2: Test of balancing property of the propensity score
Use option detail if you want more detailed output
**********************************************************


The balancing property is satisfied


This table shows the inferior bound, the number of treated
and the number of controls for each block
```

Inferior of block of pscore	gvc 0	1	Total
.0141123	810	51	861
.1	471	76	547
.2	500	165	665
.3	158	69	227
.35	176	153	329
.4	311	253	564
.5	139	167	306
.6	119	233	352
.7	68	208	276
.8	26	71	97
Total	2,778	1,446	4,224

```
Note: the common support option has been selected
```

Propensity score specification for 2013

The treatment is GVC

GVC	Frequency	Percent	Cum.
0	2,910	66.65	66.65
1	1,456	33.35	100

Estimation of Propensity score

GVC	Coefficient	Std.Error	z	P>|z|	[95% confidence interval]	
L.log(R&D intensity	0.0715204	0.0075676	9.45	0	0.0566882	0.0863525
L.Bigfirm	1.115317	0.0963671	11.57	0	0.9264407	1.304193
L.Smallfirm	-1.281415	0.1697806	-7.55	0	-1.614179	-0.9486515
L.Foreign firm	0.7399234	0.1438771	5.14	0	0.4579294	1.021917
L. log(Labour prod)	0.2018197	0.0510871	3.95	0	0.101691	0.3019485
L.log(Age)	0.171592	0.0592646	2.9	0.004	0.0554355	0.2877486

Note: industry dummies are added but coefficients are not reported. Number of observations= 3878, LR CHI2(26)= 1048.56, Prob> chi1=0.0000, Pseudo R sqrd. = 0.208. The region of common support is [.01388497, .90609829].

```
Description of the estimated propensity score
in region of common support

                 Estimated propensity score
-------------------------------------------------------------
      Percentiles      Smallest
 1%      .026897       .013885
 5%     .0467077      .0139364
10%     .0686486      .0140201       Obs                3830
25%     .1323411      .0140879       Sum of Wgt.        3830

50%      .331103                     Mean           .3547185
                        Largest      Std. Dev.      .2360876
75%     .5342553       .899628
90%     .7257251      .9002046       Variance       .0557374
95%      .776267      .9032343       Skewness       .4159376
99%     .8620792      .9060983       Kurtosis       2.065212


******************************************************
Step 1: Identification of the optimal number of blocks
Use option detail if you want more detailed output
******************************************************


The final number of blocks is 8

This number of blocks ensures that the mean propensity score
is not different for treated and controls in each blocks

*************************************************************
Step 2: Test of balancing property of the propensity score
Use option detail if you want more detailed output
*************************************************************


The balancing property is satisfied


This table shows the inferior bound, the number of treated
and the number of controls for each block
```

Inferior of block of pscore	gvc 0	1	Total
.013885	214	5	219
.05	447	28	475
.1	503	83	586
.2	358	132	490
.3	351	202	553
.4	402	396	798
.6	171	427	598
.8	25	86	111
Total	2,471	1,359	3,830

```
Note: the common support option has been selected
```

Propensity score specification for 2014

The treatment is GVC

GVC	Frequency	Percent	Cum.
0	2,515	64.29	64.29
1	1,397	35.71	100

Estimation of Propensity score

GVC	Coefficient	Std.Error	z	P>|z|	[95% confidence interval]	
L.log(R&D intensity	0.0840938	0.007807	10.77	0	0.0687923	0.0993953
L.Bigfirm	0.9761691	0.0987459	9.89	0	0.7826308	1.169707
L.Smallfirm	-1.253409	0.1699056	-7.38	0	-1.586418	-0.9203998
L.Foreign firm	0.631445	0.148417	4.25	0	0.340553	0.922337
L. log(Labour prod)	0.2019672	0.0551495	3.66	0	0.0938761	0.3100583
L.log(Age)	0.1040501	0.0618371	1.68	0.092	-0.0171484	0.2252486

Note: industry dummies are added but coefficients are not reported. Number of observations= 3482, LR CHI2(26)= 976.08, Prob> chi1=0.0000, Pseudo R sqrd. = 0.21. The region of common support is [.01857418, .91467722].

```
Description of the estimated propensity score
in region of common support

                 Estimated propensity score
-------------------------------------------------------------
      Percentiles      Smallest
 1%     .0371013       .0185742
 5%     .0610064       .0193779
10%     .0865506       .0206635       Obs                3402
25%     .1669964       .0215498       Sum of Wgt.        3402

50%     .3420341                      Mean            .382178
                        Largest       Std. Dev.      .2411401
75%     .5618939        .908411
90%     .7583812       .9109284       Variance       .0581486
95%     .8102623       .9126443       Skewness       .3985924
99%     .8745001       .9146772       Kurtosis       2.019011


******************************************************
Step 1: Identification of the optimal number of blocks
Use option detail if you want more detailed output
******************************************************


The final number of blocks is 8

This number of blocks ensures that the mean propensity score
is not different for treated and controls in each blocks

**************************************************************
Step 2: Test of balancing property of the propensity score
Use option detail if you want more detailed output
**************************************************************


The balancing property is satisfied


This table shows the inferior bound, the number of treated
and the number of controls for each block
```

Inferior of block of pscore	gvc 0	gvc 1	Total
.0185742	412	19	431
.1	454	79	533
.2	188	44	232
.25	178	81	259
.3	344	183	527
.4	333	347	680
.6	159	386	545
.8	33	162	195
Total	2,101	1,301	3,402

```
Note: the common support option has been selected
```