INPUT CAPABILITIES AND PRODUCT ADOPTION

Johannes Boehm[a], Swati Dhingra[b], John Morrow[c]

Preliminary Draft. Do not cite.

February 15, 2015

Abstract. Multi-product firms dominate production and exporting, and product turnover contributes substantially to aggregate growth. This paper examines the sources of core competencies in product adoption for Indian manufacturing establishments. We show that firms are twice as likely to add products in their Top 5 upstream and downstream sectors. Similarity to a sector's input (but not output) structure predicts product adoption, controlling for average adoption rates of all products of the firm's sector. These results show that vertical linkages drive product adoption and that within sectors, firms' product capabilities depend on economies of scope rather than product market complementarities. Unlike single product firms, multi-product firms can internalize product adoption to focus on products that have upstream or downstream linkages to existing products.

JEL Codes: L1, L2, M2, O3.
Keywords: Multiproduct firms, product adoption, vertical linkages.

Acknowledgments. We are grateful to Kalina Manova and John Sutton, as well as seminar participants at Sciences Po for helpful comments and Simen Gaure for help with the R package lfe. Boehm thanks the Sciences Po-Banque de France partnership for financial assistance and Morrow thanks the LSE-TISS-TATA Grant for financial assistance.

[a]Sciences Po. Corresponding Author: johannes.boehm@sciencespo.fr.
[b]Centre for Economic Performance, LSE.
[c]Centre for Economic Performance, University of Essex.

1. Introduction

Multiproduct firms dominate production and export activity, and their continual product turnover contributes substantially to aggregate output growth. In the United States, multi-product firms account for over 90 per cent of manufacturing output and multi-product exporters account for over 95 per cent of exports. About 89 per cent of multi-product firms vary their product mix within five years and these changes in the product mix make up a third of the increase in US manufacturing output (Bernard et al. 2007, 2010a).

Adding or dropping a product is an important event for firms. Product churning affects over a third of US firms' products and output. A growing literature shows firms continually adapt their product mix to focus on their core competencies (Bernard et al. 2005, 2011; Eckel and Neary 2010; Mayer et al. 2009; Iacovone and Javorcik 2010). Over time, firms tend to move towards a few successful products, and focusing on these selected products enables them to respond to changes in their economic environment. While the literature highlights the importance of core competencies, it does not explain their source. To better understand the nature of firms' product decisions, this paper focuses on the role of input-output linkages in firms' product adoption decisions.

A multi-product firm can internalize the supply and demand linkages across its products, and might focus on products that take advantage of economies of scope or product market complementarities (Teece, 1982). We evaluate these possibilities for manufacturing establishments in India. During the 2000s, manufacturing value added in India was on the rise and product turnover contributed 28 per cent to net sales growth, dovetailing with Harrison et al. (2013) who find that learning (increases in average, unweighted firm productivity) was the main driver of productivity growth. Indian firms actively changed their product mix, and we study detailed data on firm inputs, finding that vertical linkages are an important determinant of the product mix of firms. Controlling for firm and sector effects each period, firms are much more likely to move into upstream or downstream products from their main product line. Controlling for the rate at which firms in each sector adopt products from every other sector and using detailed firm level input and output data to measure idiosyncratic firm similarity to other sectors, we find that input similarity matters strongly for product adoption while output similarity does not. Thus firms' product capabilities lie in developing goods upstream from the firm, which suggests that economies of scope rather than product market complementarities drive product adoption.

In summary, our main findings are that:

- Vertical linkages matter:
  - Firms are 100% more likely to add products in the sectors of their Top 5 suppliers and buyers.
  - Similar results obtain when the likelihood of product adoption is estimated using expenditure shares of suppliers and buyers across all sectors.
  - Product additions along vertical linkages are associated with a higher contemporaneous growth in firm profitability.
- A firm's idiosyncratic upstream linkages predict product adoption:
  - Controlling for all sector pair rates of product adoption, a firm's idiosyncratic similarity to the inputs of a sector predicts adoption of that sector's products.
  - Product additions are associated with a higher profitability growth when the added product and the firm share input characteristics.

Studies in the business literature – that we discuss at length below – document that firms are more likely to add products that use similar technologies to the firm's main product. We contribute to the literature on vertical linkages and multiproduct firms in three ways. First, we show that product additions are driven by a firm's ability to internalize supply and output linkages across products, controlling for firm-time and product-time fixed effects. Second, using firm level input and output data and controlling for all pairs of sector level product adoption, we show that input similarity to a sector predicts product adoption while output similarity does not. This addresses concerns that the results are driven by sector-pair factors that make the co-adoption of certain products more likely. Third, we provide the first systematic study of vertical linkages within multi-product firms in the context of a developing country. This is important because supply-side bottlenecks are particularly acute in developing countries (The World Bank, 2013), and we provide estimates for the extent to which multi-product firms are able to overcome these bottlenecks by internalizing vertical linkages.

In related work, Aw and Lee (2009) focus on four Taiwanese electronics industries and estimate cost functions to arrive at the incremental marginal cost of the core product when the firm adds a new product. This provides a cost-based measure of supply linkages across products, and they find that firms move towards specializing in core products in the nineties. After controlling for plant characteristics, multi-product plants tend to drop products that are dissimilar to their core products. While Aw and Lee focus on one sector, we study supply linkages for all manufacturing industries and show that supply linkages are important across several manufacturing products.

Although there is limited work relating vertical linkages to product adoption, there are a growing number of studies relating linkages to productivity (see the forthcoming handbook chapter by Combes and Gobillon (2014)). In particular, Lopez and Sudekum (2009) find that upstream, but not downstream, linkages are associated with higher productivity, perhaps in part due to the stronger effect of upstream linkages on product adoption that we find.

Studying the liberalization period after 1991, Goldberg et al. (2009) find that Indian firms engaged in very little product dropping, relative to their US counterparts. They suggest this is due to remnants of a licensing regime that made it difficult to drop unprofitable products or due to rapid growth in the post-reform era. Lall et al. (2004) estimate sources of agglomeration in India for 1994-95, finding no evidence for inter-industry urbanization economies, where we might expect input-output linkages to matter the most. We find that by the 2000s, Indian firms dropped products at rates similar to those for US firms. Firms concentrated on their core products during this period and product churning was driven by supply linkages across products. We therefore capture the process of product turnover in an environment that moved from constraints on product rationalization to a more dynamic economy.[1]

Our findings echo previous work in the business literature which suggests firms tend to add products that have technological linkages to their core products, measured by the extent of commonality in patents used across industries. Scherer (1982) estimates technology flows from data on the proportion of patents filed in origin industries used in destination industries and interindustry economic transfers drawn from the input-output tables. Using this measure, Robins and Wiersema (1995) show firm performance is positively related to technological relatedness, controlling for industry and firm characteristics in a cross-section of 120 publicly listed US firms. Bowen and Wiersema (2005) find a positive correlation between technological relatedness of a firm's products and import penetration in core products, after controlling for industry and firm characteristics of publicly-listed US firms from 1985 to 1994. Bryce and Winter (2009) find that industry relatedness (based on the joint presence of industry activity within firms) predicts entry into related industries by building new facilities rather than by acquisition, suggesting relatedness captures specific knowledge enabling economies of scope.

[1]This is also supported by Vandenbussche and Viegelahn (2014), who show that Indian firms move away from inputs facing domestic anti-dumping measures by decreasing sales of products using these inputs.

Focusing on input relatedness, Fan and Lang (2000) construct measures for vertical relatedness and complementarity across industries using the input output tables for US firms. Vertical relatedness refers to the value of input transfer between industries and complementarity is captured by the degree of overlap in the industries' input and output markets. They find that the vertical relatedness and complementarity of US firms' products increased between 1979 to 1997, but relatedness and complementarity were not always positively related to the valuation of firms.[2] Liu (2010) uses relatedness measures to determine the impact of import competition on product lines, finding that for publicly listed US firms, import competition leads firms to drop peripheral products to refocus on core production. Rondi and Vannoni (2005) find European manufacturing firms expanded their output of core industries after integration of the European Union. They find highly diversified firms are more likely to focus on products that share intermediate inputs and markets with their core products.

The remainder of the paper proceeds to describe Indian multi-product firm data and the basic patterns and dynamics of products produced. Section 3 estimates the role of vertical linkages in product adoption and Section 4 concludes.

2. Product Turnover among Indian Manufacturing Firms

2.1. **Data.** We use annual data on manufacturing firms from the Indian Annual Survey of Industry (ASI), which is conducted by the Indian Ministry of Statistics and Programme Implementation, and is the Indian government's main source of industrial statistics on the formal manufacturing sector. The ASI consists of two parts: a census of all manufacturing plants that are larger than 100 employees, and a random sample of one fifth of all plants that employ between 20 and 100 workers.[3] The ASI's sampling methodology and product classifications have changed several times over the course of its history. In order to ensure consistency, we focus on the time frame of the fiscal years (May to April) 2000/01 to 2007/08. The unique aspect of the ASI is that it contains detailed information on both intermediate inputs and outputs at the plant level, which allows us to link the firm's input characteristics to their product mix decisions.

2.2. **Definition of Products.** At their finest levels, BRS have 1,440 5-digit SIC products for US firms under 455 4-digit SIC industries. GKPT have 1,886 "products" under

[2]Using plant-level observations from the Longitudinal Research Database, Schoar (2002) finds firms diversifying into more 2-digit SIC industries experience a net reduction in productivity.

[3]While the ASI is meant to sample at the plant level, firms can (and typically will) use the same survey form for several of their plants. Hence, we generally interpret the observations to be at the firm level.

108 4-digit NIC industries in 22 2-digit NIC sectors. Compared to them, we have 5,204 products in 5-digit ASIC industries and 1,108 products in 4-digit ASIC which should be comparable with the finest levels in BRS and GKPT. These products are in 262 3-digit ASIC industries and 64 2-digit ASIC sectors, as shown in Table 1. We also show results for nine broad sectors (1-digit ASIC). The ASI product definition is finer and we can compare with the BRS and GKPT results by working with more aggregate (4-digit) ASIC classifications. To study the determinants of product turnover, we will examine how firms add and drop 4-digit ASIC products across different 2-digit ASIC industries.

(Table 1: SIC/ASIC/Prowess comparison approximately here)

2.3. **Multi-Product and Multi-Industry Firms.** We define a multi-product firm as a plant-year observation where the plant produces two or more 4-digit products. Table 2 shows the prevalence of multi-product and multi-industry firms in the sample. Multi-product firms account for 39% of observations (41% if products are defined at the five-digit level), similar to BRS and GKPT's datasets (39% and 47%, respectively). As is well known, multi-product firms tend to be larger: they account for 71% of sales. Multi-sector firms account for 19% (2-digit) and 8% (1-digit) of the observations in the sample, but 49% (32% respectively) of sales. GKPT's sample of publicly listed firms gives similar results, 24% of firms are multi-sector firms and their share in total sales is 54%.[4]

(Table 2: Frequency of SP/MP approximately here)

2.4. **Core competencies.** Table 3 shows the sales distribution of products within firms. Firms generate a large proportion of their sales revenue from their primary products. This suggests that they have 'core competencies'.[5]

(Table 3: Sales shares by product rank approximately here)

2.5. **Product Turnover.** We now turn to documenting product and industry turnover among the ASI firms. Table 18 shows the fraction of firms that change their product/industry scope over a one-year, three-year, and five-year horizon. Given the nature of the ASI sampling methodology, our panel is not balanced; an n-year horizon hence consists of all observation pairs that are n years apart from each other. The product scope changes are forward-looking: a plant that produces one product in year t and

[4]Table 9 in Appendix B compares sales shares in our sample and GKPT's.

[5]The concentration of sales is similar to the findings of GKPT (Table 2) that uses the CMIE data on publicly listed firms, so we are confident of the soundness of the data.

the same product together with a new one in year $t+1$ would be counted as an 'add only' for a single-product firm at the one-year horizon.

Looking at the 4-digit ASIC category, we find that 65% of all firms make some change in their product range in a 5-year horizon. Even at the highly aggregate 2-digit ASIC category (which has 64 product categories), we find that 26% of all firms make some change in their product range. Product adding and product dropping are prevalent, and a large fraction of firms engage in both activities. Product turnover in the ASI data is broadly similar to BRS in the US, but higher than GKPT.[6]

(Table 18: Frequency of addings/droppings approximately here)

3. Do firms move upstream and/or downstream?

What determines which products firms adopt? In this section, we document a robust relationship between a firm's vertical linkages and the direction of product adoption. We document that firms are more likely to add products that are vertically related to their existing product line. Firms are more likely to add products in sectors that are upstream or downstream to their main products. We then show that controlling for the rates at which firms in a particular industry adopt every other industry's products, idiosyncratic input patterns predict product adoption while output patterns do not. We then show that these product adoptions are associated with significant performance premia.

3.1. **Vertical Linkages: Top Upstream and Downstream Sectors.** To study the role of vertical linkages, we first assign each firm j to the two-digit sector $n(j,t)$ where its sales revenue is the highest. We define $top5upstreamSector^k_{n(j)}$ to be one if sector k is one of the top five input supplying sectors for the sector of firm j at time t, holding the sector ranking constant over time. Likewise, we define a dummy variable $top5downstreamSector^k_{n(j)}$ that is one if sector k is one of the top five buyers of inputs from firm j's two-digit sector at time t. Note that while each firm's sector classification may change over time, we hold the rankings of top upstream and downstream sectors constant by calculating them from input and sales shares that have been aggregated over time. Furthermore, we exclude the sector itself when calculating top upstream and downstream sectors. We then estimate a linear probability model to explain the

[6]We contrast the product turnover summary statistics of the ASI with BRS and GKPT in more detail in Appendix A.

probability of product additions:

(3.1)

$$AddProd_{jt}^{k} = \beta_1 top5upstreamSector_{n(j)}^{k} + \beta_2 top5downstreamSector_{n(j)}^{k} + \alpha_{jt} + \alpha_t^k + \varepsilon_{jt}^{k}$$

The dependent variable $AddProd_{jt}^{k}$ is one if firm j adds at least one four-digit product of the two-digit sector k between years t and $t+1$.[7] All our results hold when instead using only product additions that occur with no dropping of a product in the same sector; hence our results are not driven by misclassification of products.

Equation (3.1) includes firm-year fixed effects α_{jt} to control for any factors that may cause the firm to add products in general, and sector-time fixed effects to control for sector-wide demand and supply shocks. In an additional specification, we also include firm-sector fixed effects α_j^k, thus identifying the prevalence of adding products in upstream and downstream sectors only from variation in the firm's main sector over time.

Table 4 shows the results from estimating equation (3.1) using ordinary least squares. The core result is that firms are more likely to add products in sectors that are vertically linked to their sector. The unconditional probability of adding a product in any given sector is .005; coefficients of .005 on $top5upstreamSector_{n(j)}^{k}$ and $top5downstreamSector_{n(j)}^{k}$ (column 3) mean that the firm is twice as likely to add a product in one of its top five upstream and downstream sectors. Hence, firms are more likely to move to sectors that are vertically linked to their own sector. Column 4 of Table 4 shows that this also holds when identifying the effect purely from time variation: as a firm changes its sector classification, it becomes more likely to add products in sectors that are vertically linked to the new sector, but not to the old sector.

(Table 4: Moving upstream/downstream approximately here)

3.2. **Vertical Relatedness: The Role of Similarity.** We have just seen that firms tend to add products in sectors which are upstream or downstream to the firm's sector. But do firms have inherent product capabilities, such as economies of scope or strategic complementarities across products beyond the classification of their sector? To answer this, we can predict product adoption, controlling for the rate at which firms in a sector adopt products from another sector, and use idiosyncratic firm characteristics to understand product adoption. Accordingly, to understand if idiosyncratic economies of scope in input usage drive product adoption, we define a firm's input similarity to a sector. Similarly, to understand if strategic complementarities across products drive adoption, we define a firm's output similarity to a sector.

[7]Hence we only count product additions if we observe the firm in two subsequent years.

3.2.1. *Input and Output Similarity.* We consider a firm and a sector to be similar in input characteristics if their expenditure shares on inputs from other sectors are highly correlated.[8] Our measure of input similarity is the normalized inner product of firm j's input expenditure shares (θ_{jt}) and sector k's input expenditure shares ($\overline{\theta}_{kt}$):

$$\text{inputSimilarity}_{jt}^{k} = \sum_{n=1}^{N} \theta_{jt}^{n} \overline{\theta}_{kt}^{n} / \sqrt{\left(\sum_{n=1}^{N} \left(\theta_{jt}^{n} \right)^2 \right) \left(\sum_{n=1}^{N} \left(\overline{\theta}_{kt}^{n} \right)^2 \right)}$$

where n indexes expenditure shares of spending on two-digit inputs. We construct aggregate intermediate input shares by aggregating up the micro-data, treating a plant as belonging to the two-digit sector where the value of its produced goods is the highest. $\text{inputSimilarity}_{jt}^{k}$ ranges from zero when firm j and sector k have no two-digit inputs in common to one when the input expenditure shares of firm j and sector k are identical.

The products of two sectors may also be inherently similar in their production process, such that firms who produce one are also likely to start producing products from the second.[9] Analogous to our input similarity index, we construct an output similarity index $\text{outputSimilarity}_{jt}^{k}$ as the normalized inner product between sales of firm j at time t (denoted σ_{jt}) and sector k's sales shares at time t (denoted $\overline{\sigma}_{kt}$), where n indexes the sector of purchasers:

$$\text{outputSimilarity}_{jt}^{k} = \sum_{n=1}^{N} \sigma_{jt}^{n} \overline{\sigma}_{kt}^{n} / \sqrt{\left(\sum_{n=1}^{N} \left(\sigma_{jt}^{n} \right)^2 \right) \left(\sum_{n=1}^{N} \left(\overline{\sigma}_{kt}^{n} \right)^2 \right)}.$$

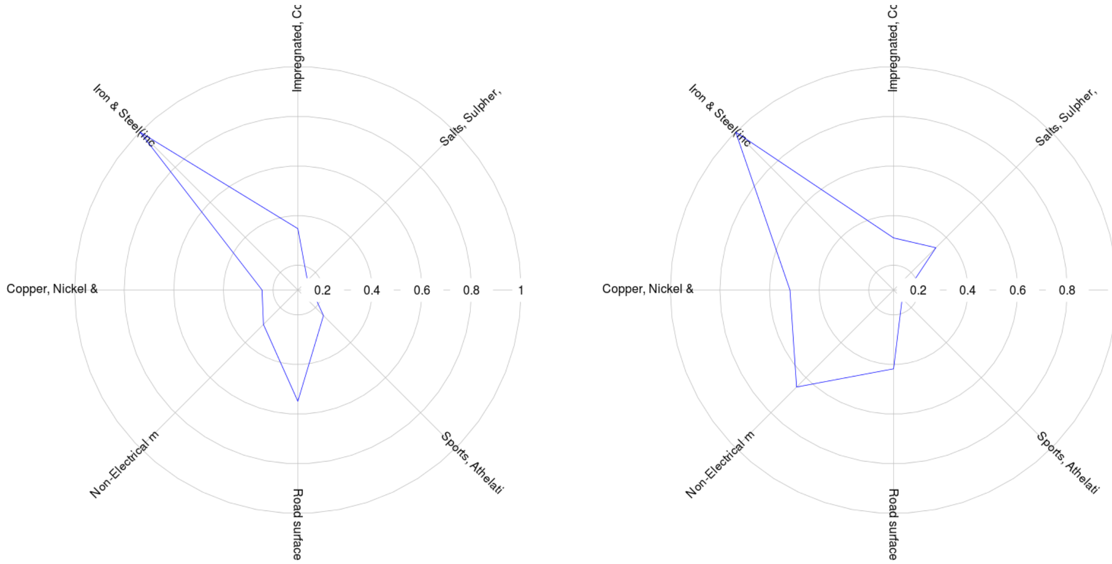
Input similarity for the Iron and Steel industry as a whole is depicted in Figure 3.2a for all industries with a similarity greater than .1. In contrast, output similarity for the Iron and Steel industry is depicted in Figure 3.2b for all industries with a similarity greater than .1. In both cases, similarity with Iron and Steel is 1, while output similarity is more concentrated than input similarity.

Table 5 shows the results of predicting product adoption, controlling for sector pair-specific product adoption rates through sector-pair fixed effects. Firms tend to add products that require a mix of inputs that is similar to the firm's existing basket of intermediate inputs. This suggests firms have inherent capabilities that they use to position themselves strategically in the product spectrum. The results show that

[8]This approach to vertical relatedness is distinct from 'circular flows' between industries (e.g. Fan and Lang (2000)). Our approach decomposes inputs and outputs in line with our definitions of top suppliers and buyers.

[9]Bernard et al. (2010) document a sizable amount of co-production particularly between textile and apparel producers, and between machinery, metal, and electronics producers.

Figure 3.1. Similarity for the Iron and Steel Industry



(A) Input Similarity (B) Output Similarity

indeed firms have idiosyncratic input capabilities, but not output capabilities, that drive product adoption. adoption, which is consistent with the high rate of adoption of upstream product lines. Quantitatively, a one standard deviation increase in the input similarity index is associated with a 15% increase in the probability of product addition.

(Table 5: Add with Sector-Pair FE approximately here)

3.2.2. *Upstream and Downstream Similarity.* Two sectors are strongly vertically linked if the intermediate input flows between them are large. For each pair of two-digit sectors we define $Upstream^k_{jt}$ as the input expenditure share of firm j's sector on sector k, which is high if sector k is an important supplier for firm j's sector. We also define $Downstream^k_{jt}$ as the share of firm j's revenues coming from sector k. When firm j's products span more than one sector, we take the maximum of each of the sectors' expenditure/sales shares.[10] We calculate the two measures from the microeconomic data on intermediate inputs; in doing so, we aggregate across years and do not distinguish between imports and domestically sourced intermediate inputs.

[10]In this case, let $y_{m,n}$ be the sales of sector m to sector n as inputs, and denote the set of sectors of firm j at time t by $I(j,t)$, then the vertical linkage measures are

$$\text{Upstream}^k_{jt} = \max_{m \in I(j,t)} y_{k,m} / \sum_l y_{l,m}, \quad \text{Downstream}^k_{jt} = \max_{m \in I(j,t)} y_{m,k} / \sum_l y_{m,l}.$$

3.2.3. *Specification and Results.* We predict the direction of product change using the similarity between the firm's current basket of intermediate inputs and the sector in which the firm may add a product, controlling for vertical linkages between that sector and the firm's current product mix. Our empirical specification is:

$$\begin{aligned} AddProd_{jt}^{k} &= \beta Upstream_{jt}^{k} + \gamma Downstream_{jt}^{k} \\ &\quad + \delta \text{inputSimilarity}_{jt}^{k} + \lambda \text{outputSimilarity}_{jt}^{k} + \alpha_{jt} + \alpha_{t}^{k} + \varepsilon_{jt}^{k}. \end{aligned} \tag{3.2}$$

We include a firm-year fixed effect α_{jt} which controls for any factors that may influence the firm's decision to change the product scope and a sector-year fixed effect α_t^k to control for sector k supply and demand shocks. The presence of firm-year fixed effects means that the variation we exploit is within firm-year observations (across targeted sectors), with α_t^k eliminating the effect of sector-wide demand and supply shocks.

(Table 6: Input capabilities approximately here)

Table 6 shows the result of estimating equation (3.2) using ordinary least squares, with standard errors clustered at the firm level in columns 3 and 4. The estimate of the input similarity index is positive and significant at the 1% level throughout all specifications, which leads us to conclude that the firm's current basket of intermediate inputs is affecting the direction of product additions, with firms preferring to add products in sectors that share common intermediate inputs with the firm's existing activities.

The coefficient of the output similarity index is also positive and statistically significant: firms are more likely to add products in sectors where firms have similar patterns of sales shares. This includes the firm's own two-digit sector, but may also indicate the presence of technological similarities between sectors.

The finding that the coefficients on both $Upstream_{jt}^k$ and $Downstream_{jt}^k$ are positive and significant reaffirms the results of Table 4. Firms are more likely to add products from sectors that are vertically related to their existing products.

3.3. **Robustness.** One concern about specification (3.2) is that the positive coefficient on the similarity index may be driven by anticipation effects: firms may know that they want to add a product, and start buying intermediate inputs already before actually producing the new good.[11] In that case, causality would run from the product addition to the similarity index.

[11]That said, the survey questionnaire asks for the intermediate inputs *consumed in that year*, hence anticipatory purchases should in principle not show up.

We employ two different strategies to find out whether this may be driving the results. First, we run specification (3.2) with the similarity index replaced by its lags. The use of lags would mean that the firm would have to do the anticipatory purchases long before starting to produce the new good in order for reverse causality to be present. Second, we construct an alternative version of the similarity index where we use the firm's input shares at the first observation (instead of at time t). Hence, in this approach, we hold the input capabilities of the firm constant over time. Any time variation in the input similarity index comes from time-variation in the sector k's aggregate intermediate input shares.

Table 13 (see Appendix) shows the results of these alternative specifications. Columns (1) uses the input and output similarity indexes constructed from the firm's initial input shares. Columns (2) and (4) lag the similarity indexes by one and two observations, columns (3) and (5) by one and two years. Throughout all specifications, the coefficient on the input similarity index remains positive and statistically significant at the 1% level. Hence, while anticipation effects may be present, they are not the main driver of the results in Table 6.

4. Profitability and the Direction of Product Change

Define profitability as the ratio of the firm's sales value to input expenditure:

$$profitability_{jt} = \frac{\sum_k sales(j,t,k)}{wL_{jt} + rK_{jt} + I_{jt}} \tag{4.1}$$

where wL_{jt} denotes payments to labor, K_{jt} denotes capital, and I_{jt} denotes intermediate inputs. For the purpose of constructing profitability, we assume an interest rate of five percent.

4.1. **Vertical Linkages and Profitability.**

$$\begin{aligned} \Delta profitability_{jt} = {} & \beta_1 add_{jt-1} + \beta_2 addInOldSector_{jt-1} \\ & + \beta_3 newAddInTopUpstreamSector_{jt-1} \\ & + \beta_4 newAddInTopDownstreamSector_{jt-1} \\ & + \sum_k \alpha_t^k \mathbb{1}_{sales(j,t,k)>0} + \alpha_t + \alpha_j + \varepsilon_{jt} \end{aligned} \tag{4.2}$$

Equation (4.2) relates product additions to changes in profitability. The dependent variable, $\Delta profitability_{jt}$, is the change of log profitability between years $t-1$ and t. add_{jt-1} is a dummy that is one if firm j adds a four-digit product between years $t-1$ and t, its coefficient is the average percentage change in profitability associated with a firm that adds a product. $newAddInTopUpstreamSector_{jt-1}$ is a dummy that denotes an addition in a two-digit sector where the firm has had positive sales in period $t-1$. $newAddInTopUpstreamSector_{jt-1}$ and $newAddInTopDownstreamSector_{jt-1}$ denote dummies for additions in the firm's top five upstream and downstream sectors, respectively. Equation (4.2) also includes a set of sector-time dummies to control for sector-wide demand and supply shocks.

(Table 15: Input capabilities approximately here)

4.2. **Input Similarity and Profitability.**

$$\begin{aligned}\Delta profitability_{jt} = {} & \beta_1 \sum_k addProduct^k_{jt-1} \\ & + \beta_2 \sum_{k:\, s(j,t-1,k)>0} addProduct^k_{jt-1} \\ & + \beta_3 \sum_{k:\, s(j,t-1,k)=0} (addProduct^k_{jt-1})(inputSimilarity^k_{jt-1}) \\ & + \beta_4 \sum_{k:\, s(j,t-1,k)=0} (addProduct^k_{jt-1})(outputSimilarity^k_{jt-1}) \\ & + \sum_k \alpha^k_t \mathbb{1}_{sales(j,t,k)>0} + \varepsilon_{jt}\end{aligned} \tag{4.3}$$

Equation (4.3) studies the relationship between profitability growth and product additions. The left-hand side variable is the change in log profitability between year $t-1$ and t, as defined in equation (4.1). The right-hand side variables are constructed from product addition dummies (recall that $addProduct^k_{jt-1}$ is a dummy that is one if firm j adds a product belonging to sector k between periods $t-1$ and t) and the input and output similarity indices in period $t-1$. We divide product additions into four different groups, each of them associated with a variable on the right-hand side of equation (4.3), to identify premia associated with each group. In the order of notation: (1) additions in all sectors, (2) product additions in sectors where the firm already had positive sales at time $t-1$, (3) product additions in sectors where the firm had no sales revenue, weighted by input similarity index, and (4) product additions in sectors where the firm had no sales revenue, weighted by the output similarity index. We also include a set of sector-time dummies to control for sector-wide demand and supply shocks.

(Table 16: Input capabilities approximately here)

5. Conclusion

The firm's product scope and turnover is a popular subject for study among the literature in economics and business, yet there is little concrete evidence of the determinants of the direction of product turnover. This paper has established that vertical linkages are one important factor in this decision: firms are more likely to add products that require similar inputs to what the firm is already using. While this is compatible with capability theories of the firm, it remains a challenge to model and empirically test the mechanisms of such theories (Andreoni, 2014). In that regard, what we have shown is that even controlling for the rates at which firms in a particular industry adopt every other industry's products, idiosyncratic input patterns predict product adoption while output patterns do not. This points further studies of product choice mechanisms towards firms' upstream connections and intermediate input use rather than behavior in downstream markets. Idiosyncratic upstream and downstream linkages also predict maintaining product lines, suggesting firms' core competencies are tied to vertical linkages. At the level of the industry, both upstream and downstream linkages are important for product adoption, although again upstream linkages are more important.

References

Andreoni, A. (2014): "Structural learning: Embedding discoveries and the dynamics of production," *Structural Change and Economic Dynamics*, 29, 58–74.

Aw, B. Y. and Y. Lee (2009): "Product Choice and Market Competition: The Case of Multiproduct Electronic Plants in Taiwan," *The Scandinavian Journal of Economics*, 111, 711–740.

Bernard, A. B., J. B. Jensen, S. J. Redding, and P. K. Schott (2007): "Firms in International Trade," *The Journal of Economic Perspectives*, 21, 105–130.

Bernard, A. B., S. Redding, and P. K. Schott (2005): "Multi-Product Firms and the Dynamics of Product Mix," *Yale School of Management*.

Bernard, A. B., S. J. Redding, and P. K. Schott (2010a): "Multiple-Product Firms and Product Switching," *The American Economic Review*, 100, 70–97.

——— (2010b): "Multiple-Product Firms and Product Switching," *American Economic Review*, 100, 70–97.

——— (2011): "Multiproduct Firms and Trade Liberalization," *The Quarterly Journal of Economics*, 126, 1271–1318.

Bowen, H. P. and M. F. Wiersema (2005): "Foreign-based Competition and Corporate Diversification Strategy," *Strategic Management Journal*, 26, 1153–1171.

Bryce, D. J. and S. G. Winter (2009): "A General Interindustry Relatedness Index," *Management Science*, 55, 1570–1585.

Combes, P.-P. and L. Gobillon (2014): "The Empirics of Agglomeration Economies," *Working Paper.*

Eckel, C. and P. Neary (2010): "Multi-product firms and flexible manufacturing in the global economy," *Review of Economic Studies*, 77, 188–217.

Fan, J. P. and L. H. Lang (2000): "The Measurement of Relatedness: An Application to Corporate Diversification," *The Journal of Business*, 73, 629–660.

Gaure, S. (2013): "lfe: Linear Group Fixed Effects," *The R Journal*, 5, 104–117.

Goldberg, P. K., A. K. Khandelwal, N. Pavcnik, and P. Topalova (2009): "Multi-product firms and product turnover in the developing world: Evidence from India," *The Review of Economics and Statistics.*

Harrison, A. E., L. A. Martin, and S. Nataraj (2013): "Learning versus Stealing: How Important Are Market-Share Reallocations to India's Productivity Growth?" *The World Bank Economic Review*, 27, 202–228.

Iacovone, L. and B. S. Javorcik (2010): "Multi-Product Exporters: Product Churning, Uncertainty and Export Discoveries*," *The Economic Journal*, 120, 481–499.

Lall, S. V., Z. Shalizi, and U. Deichmann (2004): "Agglomeration economies and productivity in Indian industry," *Journal of Development Economics*, 73, 643–673.

Liu, R. (2010): "Import competition and firm refocusing," *Canadian Journal of Economics/Revue canadienne d'économique*, 43, 440–466.

Lopez, R. A. and J. Sudekum (2009): "Vertical Industry Relations, Spillovers, and Productivity: Evidence from Chilean Plants," *Journal of Regional Science*, 49, 721–747.

Mayer, T., M. J. Melitz, and G. I. P. Ottaviano (2009): "Market size, Competition, and the Product Mix of Exporters," *Working Paper.*

Robins, J. and M. F. Wiersema (1995): "A Resource-based Approach to the Multibusiness Firm: Empirical Analysis of Portfolio Interrelationships and Corporate Financial Performance," *Strategic Management Journal*, 16, 277–299.

Rondi, L. and D. Vannoni (2005): "Are EU leading firms returning to core business? Evidence on refocusing and relatedness in a period of market integration," *Review of Industrial Organization*, 27, 125–145.

Scherer, F. M. (1982): "Inter-industry technology flows and productivity growth," *The review of economics and statistics*, 627–634.

Schoar, A. (2002): "Effects of corporate diversification on productivity," *The Journal of Finance*, 57, 2379–2403.

Teece, D. J. (1982): "Towards an Economic Theory of the Multiproduct Firm," *Journal of Economic Behavior & Organization*, 3, 39–63.

The World Bank (2013): *Global Economic Prospects*, vol. 7, World Bank Publications.

Vandenbussche, H. and C. Viegelahn (2014): "Trade protection and input switching: Firm-level evidence from Indian importers," *Working Paper*.

Appendix A. Product Turnover: ASI vs BRS vs GKPT

One fact that emerges is that product turnover in the ASI data is broadly similar to BRS. Looking at the comparable 4-digit ASIC category, we find that 65% of all firms make some change in their product range in a 5-year horizon, compared to 54% of firms in BRS. For multi-product firms, this difference is smaller: 85% in the ASI data compared with 80% in BRS. The main difference is a higher percentage of multiproduct firms drop products in the ASI than BRS, but this difference is small when the prevalence is weighted by firm sales. Compared to BRS, we find that fewer firms add and drop porducts, leading to higher levels of no activity firms when weighted by firm sales.

Another fact that emerges is that product turnover in the ASI data is higher than in GKPT. Even looking at the highly aggregate 2-digit ASIC category (which has 64 product categories), we find that 26% of all firms make some change in their product range. GKPT find instead that only 10% of firms engage in product range changes where the product is the finest level of aggregation which has 1,500 product categories. For multi-product firms, this difference is even wider: 59% in the ASI data compared with 14% in GKPT. These differences are also present for both the subset of sample firms of the ASI and the subset of census ASI firms.

Compared to GKPT, we also find higher levels of product dropping. In our sample, 7 percent of all firms drop products (4-digit) without adding new ones in the same year. The figure is higher over a three-year horizon (9%) and five-year horizon (10%). In GKPT's sample, only 2% of firms drop products without adding new ones (3% and 5% over a three-year and five-year horizon). This suggests that product dropping is indeed prevalent among Indian manufacturing firms. The right panel of Table 18 weighs the fractions of product-changing firms by their sales revenue. Twelve percent of sales revenue gets dropped at an annual frequency without being replaced by a new product in the same period (in GKPT's sample, the corresponding fraction is three percent). Hence, product rationalization is not only prevalent, but also quantitatively important.[12]

(Table 7: Turnover-BRS-GKPT approximately here)

Yet another difference between our results and GKPT's is the prevalence of product droppings by multi-product firms. In Table 8 we decompose product scope changes by studying the transition matrix of the number of products. Element (n, k) is the number (top panel) or percentage (lower panel) of firm-year observations with n products that have k products the next time we observe them. The matrix in the top panel looks fairly symmetric, suggesting that there is no salient trend in the average number of products in our sample. Since the distribution is skewed towards firms with fewer products, the transition probabilities are higher for a reduction in the number of products, which also explains the high probability of product droppings among multi-product firms in Table 18.

(Table 8: Turnover-BRS-GKPT approximately here)

Appendix B. Additional tables

(Table 9: SP/MP comparison approximately here)

(Table 10: summary statistics approximately here)

(Table 12: sales growth decomp approximately here)

(Table 13: Robustness anticipation approximately here)

[12]The fact that many firms seem to be replacing existing products by new ones raises concerns about the quality of the reported product codes. If plant managers are inconsistent over time in their reporting of product codes, the true fraction of firms that is either adding or dropping products would be higher than the observed fraction of firms. Hence, our estimates of the prevalence of product additions or droppings are lower bounds for the true number. Note also that misreporting of product codes is likely to be washed out as we aggregate products to three-digit industries and one- or two-digit sectors.

Table 1. Numbers of Distinct Products, Industries and Sectors

	BRS	GKPT	ASI
Products	1,440	1,886 CMIE	5,204 5-digit/ 1,108 4-digit
Industries	455	108	262
Sectors	20	22	64 2-digit/9 1-digit
Classification	5,4,2 digit SIC	CMIE/ 4,2 digit NIC	5,4,3,2,1 digit ASIC

Table 2. Frequency and Sales Shares of Single- and Multi-product Firms

# of Products/Industries	5-digit			4-digit			3-digit		
	Obs	% Firms	% Sales	Obs	% Firms	% Sales	Obs	% Firms	% Sales
1	152,946	58.6	28.7	159,873	61.2	30.4	176,882	67.8	37.8
2	53,859	20.6	20.4	56,503	21.6	21.5	54,777	21.0	24.1
3	26,864	10.3	12.4	24,460	9.4	14.4	19,430	7.4	13.3
4	14,477	5.5	8.6	11,413	4.4	9.7	5,869	2.2	8.0
5	6,183	2.4	7.4	4,585	1.8	5.7	2,415	0.9	5.3
6	3,028	1.2	3.7	2,134	0.8	4.3	1,030	0.4	5.8
7	1,678	0.6	3.7	1,085	0.4	5.6	441	0.2	2.2
8	1,050	0.4	3.3	599	0.2	3.6	139	0.1	1.1
9	641	0.2	4.9	299	0.1	2.0	51	0.0	0.6
10+	331	0.1	7.1	106	0.0	2.7	23	0.0	1.8

# of Products/Industries	2-digit			1-digit		
	Obs	% Firms	% Sales	Obs	% Firms	% Sales
1	212,420	81.4	50.7	239,970	91.9	68.3
2	36,568	14.0	28.4	19,219	7.4	27.3
3	8,608	3.3	12.2	1,683	0.6	4.1
4	2,523	1.0	5.0	168	0.1	0.3
5	717	0.3	2.0	15	0.0	0.0
6	180	0.1	1.6	2	0.0	0.0
7	34	0.0	0.0			
8	5	0.0	0.0			
9	2	0.0	0.0			
10+						

Source: Author's calculations from ASI data.

Table 3. Average Sales Shares by Product Rank

4-digit Products										
Rank	1	2	3	4	5	6	7	8	9	10
1	100	87	78	72	62	57	54	51	46	44
2		13	17	18	21	22	22	20	20	20
3			5	7	10	11	11	12	12	11
4				2	5	6	64	7	8	8
5					2	3	4	5	5	6
6						1	2	3	4	4
7							1	2	3	3
8								1	2	2
9									1	1
10										1

2-digit Sectors									
Rank	1	2	3	4	5	6	7	8	9
1	100	86	75	68	63	59	52	56	53
2		14	19	21	21	21	21	17	21
3			5	8	10	10	12	12	11
4				3	5	5	7	7	9
5					2	3	4	5	3
6						1	2	2	1
7							1	1	1
8								0	1
9									0
10									

Note: Columns indicate the number of products/sectors, rows indicate the rank of the product/sector.

Table 4. Moving Upstream and Downstream

	Dependent variable: Product addition dummy			
	(1)	(2)	(3)	(4)
$top5upstreamSector^k_{n(j)}$	0.005***	0.005***	0.005***	0.004***
	(0.0001)	(0.0001)	(0.0001)	(0.001)
$top5downstreamSector^k_{n(j)}$	0.006***	0.006***	0.005***	0.007***
	(0.0001)	(0.0001)	(0.0001)	(0.001)
$sameSector^k_{jt}$	0.218***	0.217***	0.216***	−0.123***
	(0.0001)	(0.001)	(0.001)	(0.003)
Firm-year FE α_{jt}		yes	yes	yes
Sector-year FE α^k_t			yes	yes
Firm-sector FE α^k_j				yes
Clustering	firm	firm	firm	firm
N	18,577,140	18,577,140	18,577,140	18,577,140
R^2	0.161	0.178	0.179	0.571

$^{*}p < 0.05$, $^{**}p < 0.01$. The sample consists of all open establishments. Dependent variable is a dummy that is one if firm j adds a new four-digit product in sector k between time t and $t + 1$.

TABLE 5. Product Additions Controlling for Sector Product Adoption Rates

Dependent variable: Product addition dummy	(1)
inputSimilarity$^{k}_{jt}$	0.012**
	(0.000)
outputSimilarity$^{k}_{jt}$	0.004
	(0.002)
Firm-year FE α_{jt}	yes
Sector-year FE α_{kt}	yes
Sector-pair FE $\alpha_{ks(j)}$	yes
Clustering	firm
N	18,254,820
R^2	0.042

*$p < 0.05$, **$p < 0.01$. The sample consists of all open establishments and excludes observations where the firm is already producing a good in sector k (i.e. same sector dummy = 1 in the above regression).

Table 6. Input capabilities: Product Additions

	Dependent variable: Product addition dummy				
	(1)	(2)	(3)	(4)	(5)
$\text{inputSimilarity}_{jt}^{k}$		0.012***	0.012***	0.012***	0.012***
		(0.0001)	(0.0002)	(0.0002)	(0.0003)
$\text{outputSimilarity}_{jt}^{k}$		0.029***	0.028***	0.029***	0.001
		(0.001)	(0.002)	(0.002)	(0.002)
$\text{Downstream}_{jt}^{k}$	0.038***	0.029***	0.028***	0.026***	0.040***
	(0.0002)	(0.0002)	(0.001)	(0.001)	(0.003)
Upstream_{jt}^{k}	0.062***	0.052***	0.051***	0.049***	0.046***
	(0.0002)	(0.0002)	(0.001)	(0.001)	(0.003)
Firm-year FE α_{jt}			yes	yes	yes
Sector-year FE α_{t}^{k}				yes	yes
Sector-pair FE $\alpha_{n(j)}^{k}$					yes
Clustering	no	no	firm	firm	firm
N	18,254,820	18,254,820	18,254,820	18,254,820	18,254,820
R^2	0.007	0.010	0.031	0.033	0.042

*$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. Sample consists of all firm-year observations where the factory is reported as open. Observations where firm j already has positive sales in sector k at time t are excluded. Dependent variable is a dummy that is one if firm j adds a new four-digit product in sector k between time t and $t+1$.

Table 7. Product Turnover

		% of Firms, over 5-year horizon							
		Unweighted				Sales-Weighted			
		no activity	add only	drop only	add and drop	no activity	add only	drop only	add and drop
4-digit	single	52	11		37	77	7		16
	multi	15	6	21	58	15	6	20	59
	all	35	9	10	47	32	6	15	47
GKPT	single	80	19		1	76	24		0
	multi	63	26	8	3	53	29	3	15
	all	72	22	4	2	57	28	2	12
BRS	single								
	multi	20	32	12	36	6	12	8	75
	all	46	14	15	25	11	10	10	68

Note: Number in the table is the (sales-weighted) percentage of firm-year observations that fall in the respective category. Product additions and drops are defined forward-looking, i.e. if a firm has one product in year 2001, and sells the same product plus an additional one in year 2002, this would count as one observation in the "add only" category in 2001 (also, it would count as a single-product firm). Hence, by definition, single-product firms cannot only drop a product. Rows “BRS” are reproduced from Table 3 in Bernard et al. (2010b). Rows “GKPT” are reproduced from Table 4 in Goldberg et al. (2009).

Table 8. Transition Matrix for the Number of Products

A. Number of Observations

# products, current period	1	2	3	4	5	6	7	8	9	10	11+
	# products, next time we observe the firm										
1	76023	7817	1688	528	209	92	35	18	8	3	0
2	7810	19402	3766	897	260	78	31	10	5	2	1
3	1767	3597	6790	1720	451	133	51	17	12	3	1
4	597	865	1762	2890	697	206	63	26	8	5	0
5	257	244	441	696	917	334	114	51	13	8	0
6	107	77	142	224	312	338	142	55	21	4	1
7	46	43	49	75	107	143	162	83	27	6	4
8	41	16	19	22	38	62	91	99	40	7	1
9	13	3	6	8	14	26	33	44	53	13	3
10	2	1	2	8	3	4	7	8	11	8	2
11+	0	0	0	1	0	2	4	0	1	1	8

B. Transition Probabilities

# products, current period	1	2	3	4	5	6	7	8	9	10	11+
	# products, next time we observe the firm										
1	88	9	2	1	0	0	0	0	0	0	0
2	24	60	12	3	1	0	0	0	0	0	0
3	12	25	47	12	3	1	0	0	0	0	0
4	8	12	25	41	10	3	1	0	0	0	0
5	8	8	14	23	30	11	4	2	0	0	0
6	8	5	10	16	22	24	10	4	1	0	0
7	6	6	7	10	14	19	22	11	4	1	1
8	9	4	4	5	9	14	21	23	9	2	0
9	6	1	3	4	6	12	15	20	25	6	1
10	4	2	4	14	5	7	13	14	20	14	4
11+	0	0	0	6	0	12	24	0	6	6	47

Note: Products are defined at the four-digit level. Sample is conditional on the firm having positive sales in both periods (no entry and exit).

Table 9. Comparison of Multi-Category Firms in GKPT and ASI

	Share of Firms		Share of Output		Mean Categories	
Type of Firm	Our Sample	GKPT	Our Sample	GKPT	Our Sample	GKPT
Single Product	0.61	0.53	0.30	0.20	1.00	1.00
Multiple Product	0.39	0.47	0.70	0.80	2.81	3.06
Multiple Industry	0.22	0.33	0.62	0.62	2.55	2.01
Multiple Sector	0.19	0.24	0.49	0.54	2.34	1.68

Note: 'Mean categories' refers to the average number of products, industries, or sectors in the respective subsample. Moreover, 'product' refers to 4-digit products, 'industry' refers to 3-digit industries, and 'sector' refers to two-digit sectors.

Table 10. Summary Statistics for the Variables in Product Turnover Regressions

Description	Variable	Obs	Mean	Std. Dev.
Product addition dummy	$addProduct_{jt}^k$	18,577,140	0.005	0.071
Product drop dummy	$dropProduct_{jt}^k$	18,577,140	0.005	0.072
Input Similarity Index	$inputSimilarity_{jt}^k$	18,577,140	0.064	0.181
Output Similarity Index	$outputSimilarity_{jt}^k$	18,577,140	0.016	0.117
Downstream Linkages	$Downstream_{jt}^k$	18,577,140	0.016	0.090
Upstream Linkages	$Upstream_{jt}^k$	18,577,140	0.016	0.084

Note: Table refers to the joint sample (all observations that are present for all variables). Product add and drop dummies refer to additions (droppings) of a four-digit product within a two-digit sector.

Table 11. Correlation table for vertical similarity and input similarity

	$inputSimilarity_{jt}^k$	$outputSimilarity_{jt}^k$
$top5upstreamSector_{n(j)}^k$	0.18	0.23
$Upstream_{jt}^k$	0.21	0.21
$top5downstreamSector_{n(j)}^k$	0.29	0.22
$Downstream_{jt}^k$	0.22	0.16

Notes: Table shows the correlation coefficients of row and column variables. Sample is all firm-year-sector triples (j, t, k) where the firm is open and does not have positive sales in sector k (as in the product addition regressions).

Table 12. Fraction of net sales growth due to within-product/industry growth

	1-digit	2-digit	3-digit	4-digit	5-digit
2003	0.93	0.89	0.91	0.81	0.78
2004	0.97	0.97	0.83	0.65	0.63
2005	1.02	0.95	0.95	0.92	0.77
2006	0.98	1.00	0.93	0.77	0.69
2007	0.88	0.81	0.74	0.58	0.56
2008	0.99	0.89	0.72	0.60	0.46
average	0.96	0.92	0.85	0.72	0.65

Notes: This table shows how much of the net sales growth is within products/industries. Some details on the construction: (1) we only look at firm-year observations in subsequent years (i.e. if there is a gap, then we don't take these observations into account. (2) we calculate the net sales growth of firms in subsequent years (3) the table then shows the net growth within x-digit products as a fraction of total net growth in (2). Hence, a fraction of 1 would mean that switching of products/industries did not contribute anything to *net* growth (but could still be large). A fraction of greater than one would mean that product switching led on average to lower sales values.

Table 13. Robustness towards Anticipation

	Dependent variable: Product addition dummy				
	(1)	(2)	(3)	(4)	(5)
similarityInputFirst	0.008***				
	(0.0002)				
similarityOutputFirst	0.107***				
	(0.002)				
similarityInputL1		0.009***			
		(0.0002)			
similarityOutputL1		0.123***			
		(0.002)			
similarityInputLY1			0.010***		
			(0.0003)		
similarityOutputLY1			0.163***		
			(0.003)		
similarityInputL2				0.009***	
				(0.0003)	
similarityOutputL2				0.131***	
				(0.003)	
similarityInputLY2					0.008***
					(0.0003)
similarityOutputLY2					0.135***
					(0.003)
$\text{Downstream}_{jt}^{k}$	0.024***	0.027***	0.033***	0.034***	0.033***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Upstream_{jt}^{k}	0.045***	0.048***	0.058***	0.057***	0.057***
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Firm-year FE α_{jt}	yes	yes	yes	yes	yes
Sector-year FE α_{kt}	yes	yes	yes	yes	yes
Clustering	firm	firm	firm	firm	firm
N	18,254,820	10,350,400	7,890,727	6,109,020	6,092,687
R^2	0.044	0.054	0.062	0.056	0.057

*$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. The independent variables are lagged observations of the input and output similarity indexes. $similarityInputFirst_j^k$ and $similarityOutputFirst_j^k$ are the input and output similarity indices at the time of the first observation of firm j. Suffixes LY1 (LY2) are variables lagged by one and two years, respectively; suffixes L1 and L2 denote the preceding observation even if it was more than one year ago. Regressions exclude observations where the firm already has positive sales.

Table 14. Identification from Time-Variation

	Dependent variable: Product addition dummy			
	(1)	(2)	(3)	(4)
$\text{inputSimilarity}_{jt}^{k}$	0.012***	0.002***	0.012***	0.002***
	(0.0002)	(0.0004)	(0.0003)	(0.0004)
$\text{outputSimilarity}_{jt}^{k}$	0.029***	0.001	0.001	0.0002
	(0.002)	(0.004)	(0.002)	(0.004)
Vertical linkage controls	yes	yes	yes	yes
Firm-year FE α_{jt}	yes	yes	yes	yes
Sector-year FE α_{t}^{k}	yes	yes	yes	yes
Firm-Sector FE α_{j}^{k}		yes		yes
Sector-pair FE $\alpha_{n(j)}^{k}$			yes	yes
Clustering	firm	firm	firm	firm
N	18,254,820	18,254,820	18,254,820	18,254,820
R^2	0.033	0.529	0.042	0.529

*$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. Sample consists of all firm-year observations where the factory is reported as open. Sectors where firm j already has positive sales are excluded. Dependent variable is a dummy that is one if firm j adds a new four-digit product in sector k between time t and $t + 1$.

Table 15. Vertical Linkages and Profitability

Dependent variable: Log Profitability Growth	(1)	(2)	(3)	(4)	(5)
add_{jt-1}	0.0225***	0.0329***	0.0175***	0.0294***	0.0243**
	(0.00264)	(0.00432)	(0.00505)	(0.00753)	(0.00755)
$addInOldSector_{jt-1}$		-0.0110**	-0.000716	0.000159	0.00357
		(0.00422)	(0.00460)	(0.00708)	(0.00708)
$newAddInTopUpstreamSector_{jt-1}$			0.00558	0.00992	0.00860
			(0.00516)	(0.00706)	(0.00732)
$newAddInTopDownstreamSector_{jt-1}$			0.0192***	0.0312***	0.0316***
			(0.00524)	(0.00733)	(0.00751)
Year FE α_t			yes	yes	yes
Firm trend α_j				yes	yes
Sector-time dummies $\alpha_{n(j),t}$					yes
N	109,822	109,138	109,110	109,110	108,831
R^2	0.001	0.001	0.002	0.339	0.338

* p<0.05, ** p<0.01, *** p<0.001. Sample consists of all firm-year observations where the factory is reported as open.

Table 16. Input Similarity and Profitability

Dependent variable: Log Profitability Growth	(1)	(2)	(3)	(4)	(5)
$\sum_k add^k_{jt-1}$	0.0245***	0.0414***	0.0344***	0.0536***	0.0279***
	(0.00172)	(0.00259)	(0.00383)	(0.00546)	(0.00660)
$\sum_{k \in S(j,t-1)} add^k_{jt-1}$		-0.0292***	-0.0220***	-0.0293***	-0.00958
		(0.00334)	(0.00443)	(0.00635)	(0.00702)
$\sum_{k \notin S(j,t-1)} add^k_{jt-1} inputSimilarity^k_{jt-1}$			0.0183*	0.0218*	0.0282**
			(0.00746)	(0.0103)	(0.0107)
$\sum_{k \notin S(j,t-1)} add^k_{jt-1} outputSimilarity^k_{jt-1}$			0.0318	0.00963	-0.0357
			(0.109)	(0.145)	(0.154)
Year FE α_t			yes	yes	yes
Firm FE α_j				yes	yes
Sector-time dummies α^k_t					yes
N	109822	109822	109822	109822	109278
R^2	0.002	0.003	0.003	0.338	0.345

* $p<0.05$, ** $p<0.01$, *** $p<0.001$. Sample consists of all firm-year observations where the factory is reported as open. $S(j,t)$ denotes the set of all two-digit sectors where firm j has positive sales revenue at time t.

TABLE 17. VERTICAL LINKAGES AND PROFITABILITY: INTENSIVE & EXTENSIVE MARGIN

	Dependent variable: Log Profitability				
	(1)	(2)	(3)	(4)	(5)
Sales share in top5 upstream sectors	0.653***	0.634***	0.381***	0.185***	0.140***
	(0.0241)	(0.0241)	(0.0217)	(0.0215)	(0.0214)
Sales share in top5 downstream sectors	0.798***	0.792***	0.471***	0.249***	0.210***
	(0.0234)	(0.0233)	(0.0212)	(0.0211)	(0.0210)
numberOfProducts				0.100***	
				(0.0013)	
Firm controls		Yes	Yes	Yes	Yes
Year FE α_t		Yes	Yes	Yes	Yes
Firm FE α_j			Yes	Yes	Yes
Number Of Product Dummies					Yes
N	321936	321551	321551	321551	321551
R^2	0.019	0.022	0.772	0.779	0.781

* p<0.05, ** p<0.01, *** p<0.001. Sample consists of all firm-year observations where the factory is reported as open.

Table 18. Product Turnover

		Percentage of Firms												Sales-Weighted Percentage of Firms											
		1-year horizon				3-year horizon				5-year horizon				1-year horizon				3-year horizon				5-year horizon			
		no activity	add only	drop only	add and drop	no activity	add only	drop only	add and drop	no activity	add only	drop only	add and drop	no activity	add only	drop only	add and drop	no activity	add only	drop only	add and drop	no activity	add only	drop only	add and drop
1-digit	single	93	4		3	92	5		4	91	5		4	93	6		1	92	7		1	91	7		1
	multi	51	4	38	7	40	4	48	8	34	3	53	9	59	5	25	11	58	5	30	7	56	4	36	4
	all	89	4	4	4	86	4	5	4	85	5	6	5	81	5	9	5	80	6	10	3	80	6	12	2
2-digit	single	84	7		10	81	8		11	79	9		12	89	7		4	89	7		4	89	7		4
	multi	41	7	31	21	30	7	38	24	26	7	42	25	41	9	27	23	35	10	34	22	35	6	40	19
	all	74	7	7	12	69	8	9	14	66	9	11	15	64	8	14	14	61	8	18	13	62	6	20	12
3-digit	single	75	8		17	70	11		19	68	12		20	86	7		7	85	8		7	84	8		8
	multi	36	8	24	33	26	8	29	36	22	8	31	39	29	10	25	37	23	14	26	38	22	9	34	36
	all	62	8	8	22	54	10	11	25	51	10	12	27	48	9	16	27	44	12	17	27	43	9	22	26
4-digit	single	63	7		30	56	10		35	52	11		37	80	5		15	79	6		15	77	7		16
	multi	26	6	16	51	18	7	20	56	15	6	21	58	23	6	17	54	16	11	17	56	15	6	20	59
	all	47	7	7	39	39	8	9	44	35	9	10	47	39	5	12	44	33	10	12	45	32	6	15	47

Note: Number in the table is the (sales-weighted) percentage of firm-year observations that fall in the respective category. Product additions and drops are defined forward-looking, i.e. if a firm has one product in year 2001, and sells the same product plus an additional one in year 2002, this would count as one observation in the "add only" category in 2001 (also, it would count as a single-product firm). Hence, by definition, single-product firms cannot only drop a product.