AN ABSTRACT OF THE THESIS OF

Jangho Choi for the degree of Doctor of Philosophy in Applied Economics presented on May 15, 2014.

Title: Multi-Product Firms' Productivity and Export Behavior

Abstract approved: ____

Munisamy Gopinath

The objective of this dissertation is to examine multi-product firms' productivity and export behavior. More specifically, this dissertation estimates productivity of firms that produce and sell multiple products, and the role of productivity in such firms' export behavior. In doing so, this dissertation develops a firm-level gravity approach to test whether multi-product firms self select to export or learn from exporting.

In the first of part of this dissertation, I examine whether intra-firm resource reallocation of a multi-product firm affects its Total Factor Productivity (TFP). By extending earlier approaches, I estimate unbiased and consistent TFP of multi-product firms using a revenue-based production function. I find that TFP is more likely to be overestimated when multi-product firms' internalized demand linkage is not taken into account. I also find that multi-product firms' TFP decreases as it expands the number of products produced, but specialization of production does not play a role in TFP.

In the second part, I present a theoretical framework to derive a firm-level gravity equation. By equating the total demand and total production of multi-product firms, I

derive a firm-level gravity equation where export flows from firms to consumers is proportional to the product of economic size of firms, consuming power of a representative consumer, and trade resistance between origin and destination. Using the firm-level gravity equation, I test the hypothesis that high productivity firms self select to export and that the size of export flows is determined by productivity. I find that the economic size of exporting firms and the consuming power of a representative consumer have a positive and statistically significant effect on exports, while trade resistance such as tariff and distance have the opposite effect. I also find that the estimated coefficients of the firm-level gravity equation tend to be smaller than those of the traditional country-level gravity equation.

In the final part of the dissertation, I test whether or not previous export experience improves the productivity of firms. Again, the estimable equation is derived from the equilibrium condition presented in the second part. My result confirms that previous export experience indeed improves productivity of exporting firms, but tariffs have the opposite effect.

The results of this dissertation reveal the economic behavior of multi-product firms, which usually account for a large of economic activity and output in many countries. Understanding such firms' productivity and export behavior can offer strategies for economic growth and development. Empirical findings of this dissertation suggest policy options including lowering tariffs, and improving infrastructure that can lower transportation costs. Further examination of product range and specialization of production can offer strategies to source exports from small and midsize firms.

©Copyright by Jangho Choi
May 15, 2014
All Rights Reserved

Multi-Product Firms' Productivity and Export Behavior

by
Jangho Choi

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented May 15, 2014
Commencement June 2014

Doctor of Philosophy thesis of Jangho Choi presented on May 15, 2014.

APPROVED:

Major Professor, representing Applied Economics

Head of the Department of Applied Economics

Dean of the Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

Jangho Choi, Author

TABLE OF CONTENTS

Page

TABLE OF CONTENTS (Continued)

Page

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

LIST OF TABLES

Multi-Product Firms' Productivity and Export Behavior

1. Introduction

The relationship between trade and productivity has received major research attention in the past few decades. Both motivation and consequences of trade liberalization are important to policy makers, consumers and exporting firms. The traditional Heckscher-Ohlin model and more recent monopolistic competition approaches, e.g. Krugman (1980), have been employed to understand the productivity-trade relationship, but mostly with aggregate or industry data. Such a broad analysis well suits the assumptions underlying such models: perfectly specialization, single-product, homogeneous firms and symmetric market shares.

The assumption of single-product, homogeneous firms equally sharing a market has not held out well in observed data, characterized by firms producing multiple products, and differing in levels of productivity, export destinations and volumes. Recent heterogeneous firms models (Melitz, 2003) address the variation in size, productivity and market share among firms in an industry. Here, openness to trade leads to intra-industry dynamics where high productivity firms expand exports, low productivity firms reduce exports or stops to export, and least productivity firms exit from the market. These various responses are referred to as across-firm resource reallocation. However, the heterogeneous firm model still maintains the single-product firm assumption.

As firm-level data became available, the role of multi-product firms has received increasing research attention in the past few years. A firm with multiple products can internally adjust their resources among products in response to changing economic circumstances, while single product firms can only scale their output up or down (to possibly zero). Here, multi-product firms' resource adjustment is referred to as intra-firm resource reallocation. The multi-product firms are very different from single-product firms in practice (Bernard, Redding and Schott 2010). In most economies, they are small in number but account for a large share of their respective markets. Bernard, Redding, and Schott (2010) report that multi-product firms account for 87% of the total output in U.S. manufacturing but only 39% of the total number of firms during 1987 to 1997.

Emerging literature points out to a new channel of productivity improvement through intra-firm resource reallocation. Intra-firm resource reallocation brings additional theoretical and empirical issues in analyzing the relationship between trade and productivity. Emerging literature proposes theoretical modifications and empirical evidence to account for multi-product firms in the trade and productivity relationship. More specifically, product heterogeneity in addition to firm heterogeneity is introduced to analyze the *intra-firm* resource reallocation in international trade theory and applications (Eckel and Neary (2003), Bernard, Redding and Schott (2011) and Mayer, Melitz and Ottaviano (2011)). Despite increased firm-level data availability, little is known about multi-product firms' structure and behavior in many economies.

A better understanding of multi-product firms' behavior is required in developing and developed countries to examine consequences of trade and openness. In this dissertation, I examine the trade and productivity relationship in the presence of multi-product firms, and present empirical evidence on their export behavior. In the first part of this dissertation, I propose an approach to estimate productivity in the context of a multi-product firm, and test the impact of intra-firm resource reallocation on productivity. In the second part, I extend the traditional gravity-type equation to a firm-level context to analyze the relationship between trade and trade productivity. In the final part, I analyze the relationship between export experience and productivity.

First, I estimate multi-product firms' Total Factor Productivity (TFP) and examine whether intra-firm resource reallocation affects TFP of multi-product firms. When TFP is estimated as a residual in most production function based micro studies, revenue is used as a proxy for output. Although revenue is the product of price and output, these studies do not control for price in the revenue-based production function. Loecker (2011) proposes an empirical method to control for price effects and estimate unbiased productivity by merging production and demand functions. I extend the Loecker's method to account for multi-product firms' internalized demand linkage and price effects. I find that TFP is more likely to be overestimated if multi-product firms' internalized demand linkage is not considered. I also test the impact of intra-firm resource relocation on TFP and find that multi-product firms' productivity improves as a firm reduces the number of products.

In the second part of this dissertation, I suggest a gravity-type framework to analyze firm-level data. The traditional gravity-type equation usually analyzes country- or industry- trade flows. For firm-level data, I develop a framework for the gravity-type equation by equating the total demand and total production of multi-product firms.

By rearranging the equilibrium condition, I show a relationship between productivity and trade. Assuming that high productivity firms self select to export which is referred to as self-selection hypothesis. I empirically show that the firm-level gravity equation has the distinct feature that the export flow of multi-product firms is proportional to the economic size of a firm and consuming power of a representative consumer in a destination. Estimating alongside, a traditional country-level gravity equation allows me to compare firm-level outcomes with aggregate behavior. Rearranging the equilibrium condition leads to the alternative learning-by-exporting hypothesis, where export experience of firms enriches their knowledge and enhance TFP.

For the dissertation, the production accounts of the PROWESS database on Indian firms (31,100 firms with 213,134 observations; 3,844 products with 637,998 observations) are merged with the TIPS database on product-level exports for 1997-2002 (134,133 firms with 19,104,370 observations). This unique database allows the study of the relationship between multi-product firms' structure, trade and productivity.

The results of this dissertation reveal the structure and behavior of multi-product firms, which usually account for a large of economic activity and output in many countries. Understanding such firms' productivity and export behavior can offer strategies for economic growth and development. Empirical findings of this dissertation suggest policy options such as reducing the number of products a firm produces, reducing tariffs, and export encouragement for small and middle size firms.

2. Structure and Behavior of Multi-Product Firms : Evidence from India

2.1 Introduction

A central theme of international trade research has been the impact of trade liberalization on productivity. Work by Prescott (1998), Pavnick (2003), Amity and Konings (2007), and Feenstra and Kee (2008) argue that differences in productivity mostly explain income differences across countries, that trade liberalization improves productivity, and that all economic agents share the gains from productivity following trade liberalization.

Early literature on this theme points out that trade liberalization brings resource/organizational adjustment across industries and this adjustment enhances productivity. A traditional comparative advantage or monopolistic competition model examines responses at the average, i.e. homogeneous firms. In recent years, heterogeneous firm models with a general equilibrium framework expand the debate to include organizational adjustment across firms. Following trade liberalization, more efficient industries or firms expand their production and exporting status, while inefficient industries or firms shrink or even leave the market. The heterogeneous firm models argue that there is organizational adjustment across firms even within an efficient industry. Contributions by Melitz (2003), Bernard and Jensen (2004), Melitz and Ottaviano (2008) and Bernard, Redding and Schott (2007) argue that only firms with productivity levels higher than a certain cutoff self select to serve domestic and foreign markets.[1] The productivity improvement in the heterogeneous-firms

[1] The cutoff arises from a long-run zero-profit condition equating cumulative profits to post-entry fixed production costs.

framework arises through organizational adjustments of industries or firms following trade liberalization. The exit of less efficient industries or firms and the transfer of their resources to more efficient industries or firms lead to improvements in industry or national productivity.

A new strand of the heterogeneous firm literature is now considering explanations of productivity change arising from intra-firm resource reallocation in the presence of product heterogeneity (or product differentiation). Under firm heterogeneity, a firm's technology uses determine its productivity; technology usages include adoption and efficient use of adopted technologies. However, recent literature points out that there is a possibility that intra-firm resource reallocation affect a firm's productivity in addition to technology usages. Restuccia and Rogerson (2013) summarize direct and indirect approaches explaining the impact of intra-firm resource reallocation on a firm's total factor productivity (TFP). The direct approach directly employs multiple factors affecting TFP such as inefficient input use caused by employment protection policies, and measures the magnitude of each factor's impact on TFP. In contrast, the indirect approach does not explicitly select factors, but it measures the aggregated effect of all causal factors on TFP. Work by Eckel and Neary (2010), Bernard, Redding and Schott (2011) and Mayer, Melitz and Ottaviano (2011) show that product heterogeneity drives intra-firm resource allocation. Product heterogeneity can consist of either productivity differences across products (Eckel and Neary, 2010; Mayer, Melitz and Ottaviano, 2011) or differences in attributes across products (Bernard, Redding and Schott, 2011). Attribute differences refer to that products are symmetric in terms of productivity, but differ in terms of characteristics such as brand and quality. These studies claim that as firms change their number of products, they also change firm-level productivity because changes in the number of products cause inefficient management or lead allocation of inputs to less efficient products. In addition to technology adoption and usage, intra-firm resource reallocation can determine attained productivity.

The primary purpose of this chapter is to show whether intra-firm resource reallocation affects multi-product firms' TFP. Multi-product firms' intra-firm resource reallocation is measured by two terms: the number of products a firm produces (product range), and the way a firm allocates input resource across products (specialization of production).

However, I also show that multi-product firms' internalized demand linkage between the products within a firm should be considered when TFP is measured using the revenue production function in which deflated revenue is used as proxy for output to estimate production function. Revenue is the product of price and output, but on the right hand side of the revenue production function, there is no control for the price in revenue. Uncontrolled price effect in revenue causes omitted variable problem and yield biased TFP. Loecker (2011) suggests an alternative method to handle the price effects in revenue. By merging a production function with a demand function, sectors' aggregated demand and dummies for products is newly added to the revenue production function as a control for the price effects in revenue. I extend Loecker's approach to include multi-product firms' internalized demand linkage.

The multi-product firms' internalized demand linkage which is referred as a cannibalization effect implies a reduction in output or revenue of existing products as a result of the introduction of a new product by the same firm. Eckel and Neary (2010) use the cannibalization effect as a distinct feature of multi-product firms. Products' demand within a firm are connected to each others, and so changing the number of products a firm produces could change the revenue of existing products within a firm. At product-level, the cannibalization effect is inherently negative to existing products, but at firm-level, it can be positive by ultimately raising firm-level revenue which is sum of all products' revenue. In addition to the price effect in revenue, I also test whether the cannibalization plays a role in revenue of multi-product firms and a consequence result of the cannibalization in estimation of TFP. I find that TFP is more likely to be over-estimated because TFP contains the cannibalization effects in it when the revenue production function is estimated without considering the cannibalization effects. Note that TFP is measured using Loecker's (2011) approach adopting the Cobb-Douglas production and CES utility functions with the presence of multi-product firms and product-heterogeneity.

Next, multi-product firms' behavior is examined through the testing the following hypothesis: [Hypothesis I: A concentrated production system with a large production share for one or two products increases multi-product firms' TFP, while diversifying it decreases TFP]. This hypothesis implies that Discontinuing a product and (/or) skewing production toward a particular product increases TFP while adding a product and (/or) equalizing the production of all products decreases TFP.

I find empirical evidence supporting that only changes in product range have an impact on multi-product firms' TFP. This has three implications. First, multi-product firms indirectly optimize their productivity to maximize profit. Expanding product range decreases TFP and this leads to cost increases to firms. However, firms can take additional demand and so revenue from the newly added product. If the additional revenue is larger than cost increases, firms add a new product and sacrifice productivity. Otherwise, firms reduce their product range and increase their productivity. Second, aggregate TFP depends not only on organizational adjustment across industries or firms, but also on organizational adjustment within a firm. Multi-product firms' TFP depends not only on technology usages, but also on intra-firm resource reallocation. Product range and the way to allocate input resource across heterogeneous products also affect a firm's TFP. Finally, getting exporter status could change multi-product firms' productivity if a firm reduces non-exporting products and concentrates their production on exporting products.

For the empirical analysis, the production and finance accounts of the PROWESS database on Indian firms (31,100 firms with 213,134 observations; 3,844 products with 213,134 observations) are used. This unique database allows the chapter to focus on multi-product firms' structure, productivity, product range and the specialization of production.

The remainder of this chapter is organized as follows. The next section describes the empirical models. Section 3 contains estimation strategies, while section 4 introduces the data and Section 5 discusses hypotheses and results. The last section concludes.

2.2 Empirical Model

As noted earlier, multi-product firms' productivity is estimated using the Loecker's (2011) approach. However, this chapter considers an additional effect in the TFP estimation: cannibalization. If a firm changes its product range or (and) the specialization of production, physical outputs of products within a firm and prices are affected. In addition, following the assumptions in Bernard, Redding and Schott (2011), this chapter uses a model where all products within a firm possess identical productivity, but they are heterogeneous in terms of attributes such as quality and brand names. Due to product heterogeneity, firms can take additional demand and profit from the newly added product in addition to those from existing products. This approach is analytically tractable and can be readily tested using firm- and product-level data.

2.2.1. PRODUCTION

A product' production is a function of technology and a firm's productivity. Homogeneous technology and common input prices across products are assumed. Autarky is assumed as well. Product *i* of firm *j* has a standard Cobb-Douglas production function, given by

$$(1)\ Q_{it}^{j} = \left(a_{it}^{j}K_{t}^{j}\right)^{\alpha_k}\left(a_{it}^{j}L_{t}^{j}\right)^{\alpha_l}\left(a_{it}^{j}M_{t}^{j}\right)^{\alpha_m} exp\left(\varphi_{t}^{j} + u_{t}^{j}\right)$$

$$= \left(a_{it}^{j}\right)^{\alpha_k+\alpha_l+\alpha_m}\left(K_{t}^{j}\right)^{\alpha_k}\left(L_{t}^{j}\right)^{\alpha_l}\left(M_{t}^{j}\right)^{\alpha_m} exp\left(\varphi_{t}^{j} + u_{t}^{j}\right)$$

$$= \left(a_{it}^{j}\right)^{\gamma}\left(K_{t}^{j}\right)^{\alpha_k}\left(L_{t}^{j}\right)^{\alpha_l}\left(M_{t}^{j}\right)^{\alpha_m} exp\left(\varphi_{t}^{j} + u_{t}^{j}\right)$$

where $i \in I_j(i)$ and $I_j(i)$ is a set of products in firm *j*; Q_{it}^{j} denotes physical output of product *i* in firm *j* at time *t*; Q_{t}^{j} denotes firm *j*'s aggregated physical output; K_{t}^{j}, L_{t}^{j} and M_{t}^{j} are the firm's capital, labor, and material, respectively; φ_{t}^{j} denotes multi-product firms' unobservable productivity; a_{it}^{j} denotes input share of product *i* in firm

j's input use and shows how much inputs are allocated to product i. $\sum_{i \in I_j(i)} a_{it}^j = 1$ and $a_{hit}^j = a_{it}^j$ for h $= \{K_t^j, L_t^j, M_t^j\}$; and u_t^j is measurement error and idiosyncratic shocks to production. The assumptions for input share imply that the production function satisfies input proportionality and that there is no input cost synergy. Returns to Scale is γ, $\alpha_k + \alpha_l + \alpha_m = \gamma$.

A firm's aggregated physical output is the sum of the multiple products' output, given by,

$$(2)\ Q_t^j = \sum_i Q_{it}^j$$

$$= \sum_i \left(a_{it}^j\right)^{\gamma} \left(K_t^j\right)^{\alpha_k} \left(L_t^j\right)^{\alpha_l} \left(M_t^j\right)^{\alpha_m} exp\left(\varphi_t^j + u_t^j\right)$$

$$= \left(K_t^j\right)^{\alpha_k} \left(L_t^j\right)^{\alpha_l} \left(M_t^j\right)^{\alpha_m} exp\left(\varphi_t^j + u_t^j\right) \sum_i \left(a_{it}^j\right)^{\gamma}$$

Physical output is not available due to data constraints. In general, deflated revenue by the producer price index is used with the assumption that unobserved price effects in revenue are eliminated by deflating it. However, uncontrolled price effects in revenue cause omitted price variable problem if omitted prices are correlated with input uses. Alternatively, Loecker (2011) suggest that by combining Cobb-Douglas production function with CES preference, unobserved price effects in revenue can be controlled for.

2.2.2 DEMAND

All consumers have a common utility function satisfying the Constant Elasticity Substitution (CES) preference. A representative consumer's demand on product *i* from firm *j* is given by,

(3) $Q_{it}^{j} = Q_{st}\left(\frac{P_{it}^{j}}{P_{st}}\right)^{\sigma_s} exp\left(\xi_t^j\right)$.

where σ_s denotes elasticity of substitution and it is allowed to vary by sector s, $\sigma_s < -1$; Q_{st} denotes total demand of sector *s*; P_{it}^{j} denotes price of product *i* from firm *j*; P_{st} denotes aggregate price index of sector *s*; ξ_t^j denotes firm specific unobserved demand shocks. Note that the CES preference assumption implies constant markups over marginal costs.

2.2.3 AGGREGATION

Firm *j*'s aggregated revenue is given by $R_t^j = \sum_i P_{it}^j Q_{it}^j$. By combing the production and the demand functions, firm *j*'s revenue function is given by,

(4) $R_t^j = \sum_i P_{it}^j Q_{it}^j$

$$= \sum_i Q_{st}^{-\frac{1}{\sigma_s}} \left(exp\left(\xi_t^j + u_t^j\right)\right)^{-\frac{1}{\sigma_s}} P_{st} \left(Q_{it}^j\right)^{\frac{\sigma_s+1}{\sigma_s}}$$

$$= Q_{st}^{-\frac{1}{\sigma_s}} \left(exp\left(\xi_t^j + u_t^j\right)\right)^{-\frac{1}{\sigma_s}} P_{st} \sum_i \left(\left(a_{it}^j\right)^{\gamma} Q_t^j\right)^{\frac{\sigma_s+1}{\sigma_s}}$$

$$= Q_{st}^{-\frac{1}{\sigma_s}} \left(exp\left(\xi_t^j + u_t^j\right)\right)^{-\frac{1}{\sigma_s}} P_{st} \left(Q_t^j\right)^{\frac{\sigma_s+1}{\sigma_s}} \sum_i \left(a_{it}^j\right)^{\gamma\left(\frac{\sigma_s+1}{\sigma_s}\right)}$$

$$= Q_{st}^{-\frac{1}{\sigma_s}} \left(exp\left(\xi_t^j + u_t^j\right)\right)^{-\frac{1}{\sigma_s}} P_{st} \left(Q_t^j\right)^{\frac{\sigma_s+1}{\sigma_s}} A_t^j$$

$$= Q_{st}^{-\frac{1}{\sigma_s}} \left(exp\left(\xi_t^j + u_t^j\right)\right)^{-\frac{1}{\sigma_s}} P_{st} \left(\left(K_t^j\right)^{\alpha_k} \left(L_t^j\right)^{\alpha_l} \left(M_t^j\right)^{\alpha_m} exp\left(\varphi_t^j\right)\right)^{\frac{\sigma_s+1}{\sigma_s}} A_t^j$$

$$= P_{st}\left(K_t^j\right)^{\left(\frac{\sigma_s+1}{\sigma_s}\right)\alpha_k} \left(L_t^j\right)^{\left(\frac{\sigma_s+1}{\sigma_s}\right)\alpha_l} \left(M_t^j\right)^{\left(\frac{\sigma_s+1}{\sigma_s}\right)\alpha_m} A_t^j Q_{st}^{-\frac{1}{\sigma_s}} \left(exp\left(\xi_t^j\right)\right)^{-\frac{1}{\sigma_s}}$$

$$\left(exp\left(\varphi_t^j + u_t^j\right)\right)^{\frac{\sigma_s+1}{\sigma_s}}$$

where A_t^j denotes firm *j*'s the specialization of production which is the sum of a product *i*'s input share powered by the elasticity of substitution and the returns to scale, $A_t^j = \sum_i \left(a_{it}^j\right)^{\gamma\left(\frac{\sigma_s+1}{\sigma_s}\right)}$. Note that the specialization of production reveals how a firm allocates its inputs across products. The specialization of production has the largest value when a firm produces a single product or a firm allocates most of its input resource toward a particular product. As a firm reallocates its inputs toward other products, the specialization of production decreases and it has the lowest value if all inputs are equally distributed over all products within a firm. This feature is preserved regardless of sign or size of elasticity of substitution. Now the specialization of production plays a role in estimating the production function. Unlike Loecker's (2011) approach the specialization of production and product range are not constant over time in this chapter.

By taking logs, equation (4) can be written as the following,

$$(5)\ r_t^j = p_{st} + \left(\frac{\sigma_s+1}{\sigma_s}\right)\alpha_k\, k_t^j + \left(\frac{\sigma_s+1}{\sigma_s}\right)\alpha_l\, l_t^j + \left(\frac{\sigma_s+1}{\sigma_s}\right)\alpha_m\, m_t^j + a_t^j - \frac{1}{\sigma_s} q_{st} - \frac{1}{\sigma_s}\xi_t^j + \frac{\sigma_s+1}{\sigma_s}\varphi_t^j + \frac{\sigma_s+1}{\sigma_s} u_t^j$$

where lower cases denote logs. Including sectors' total demand for multi-sector firms, estimable equation is given by,

$$(6)\ \tilde{r}_t^j = r_t^j - p_{st}$$
$$= \beta_k k_t^j + \beta_l l_t^j + \beta_m m_t^j + \beta_a a_t^j + \sum_{s=0}^{S} \beta_q s_s q_{st} + \varphi_t^{j^*} + \xi_t^{j^*} + u_t^{j^*}$$

where $\tilde{r}_t^j$ is deflated revenue by producer price index in sector *s*, $\tilde{r}_t^j \equiv r_t^j - p_{st}$; s_s denotes multi-sector firms' revenue share in sector s and it can be between 0 and 1; $\beta_h = \left(\frac{\sigma_s+1}{\sigma_s}\right)\alpha_h$ for $h = \{k, l, m\}$; $\beta_q = \frac{1}{|\sigma_s|}$; $\varphi_t^{j^*} \equiv \varphi_t^j \left(\frac{\sigma_s}{\sigma_s+1}\right)$; $\xi_t^{j^*} \equiv \xi_t^j \frac{1}{|\sigma_s|}$; and

$$u_t^{j^*} \equiv u_t^j \left(\frac{\sigma_s}{\sigma_s + 1}\right).$$

2.3 Estimation Strategy

2.3.1 Estimation Strategy for production function

Unobserved demand shocks, ξ_t^j, in equation (6) are decomposed into observable nesting structures of the product data and product range, n_j^t, and the stochastic component, $\tilde{\xi}_t^j$, as in the following[2],

$$(7)\ \xi_t^j = \sum_{p \in G(i)} D_{pt}^j + n_t^j + \tilde{\xi}_t^j$$

where D_{pt}^j denotes fixed effects for product-group and $G(i)$ denotes the set of product-group; $\tilde{\xi}_t^j$ denotes stochastic components and *i.i.d.* across firms and time. Product range of firms has an impact on demand shocks due to the cannibalization effect, where output or market share of one product can be changed as a result of introduction of a new product by an identical firm. The cannibalization effect captures internalized demand linkages of multi-product firms (Eckel and Neary, 2010)[3]. Decomposing demand shocks by nesting the structures of products is examined by Goldberg (1995) and Loecker (2011). Note also that if the assumption that production function is common across products is not satisfied, then dummies for products could capture the difference in production technology.

---

[2] Nesting structures of the product data includes product (5 digit SITC), product-group (3 digit SITC), industry (2 digit SITC) and sector (1 digit SITC). All these nesting structures could be included in estimation, but it brings a practical implementation problem due to limitations of estimation procedures. This chapter only considers product-group fixed effect as a nesting structure of product.

[3] Loecker (2011) decomposes demand shocks by the fixed effects for product, protection rate in trade, and a stochastic term.

A multi-product firm's productivity determination process follows a first-order Markov process and so, it is a function of previous productivity, previous product range, previous the specialization of production and stochastic components, η_t^j, as given by,

$$(8)\ \varphi_t^j = g_t(\varphi_{t-1}^j,\ a_{t-1}^j,\ n_{t-1}^j) + \eta_t^j$$

where stochastic components, η_t^j, are interpreted as innovation or productivity shocks and *i.i.d* across producers and time[4]. Product range and the specialization of production in equation (8) capture the relationship between intra-firm resource reallocation and TFP.

The productivity determination process reflects the assumption that firms determine the specialization of production and product range based on current productivity, but current the specialization of production and product range affect next period productivity because firms needs some time to reorganize management and production systems, and eliminate inefficiency across products without affecting input use.

Shocks in demand and productivity lead to the main estimable equation as the following,

$$(9)\ \tilde{r}_t^j = \beta_k k_t^j + \beta_l l_t^j + \beta_m m_t^j + \beta_a a_t^j + \beta_n n_t^j + \textstyle\sum_s s_s^j\, \beta_q q_{st} + \sum_{i \in G(i)} D_{pt}^j$$

$$+ \textstyle\sum_{y \in J(y)} D_y + \varphi_t^{j^*} + \tilde{\xi}_t^j + u_t^{j^*}$$

$$= \beta_k k_t^j + \beta_l l_t^j + \beta_m m_t^j + \beta_a a_t^j + \beta_n n_t^j + \beta_{sq} q_{st} + \delta_p D_{pt} + \delta_y D_y + \varphi_t^{j^*} + \upsilon_t^{j^*}$$

[4] $\eta_t^j = \varphi_t^j - g_t(\varphi_{t-1}^j,\ n_{t-1}^j,\ a_{t-1}^j)$ and it is interpreted as innovation.

where D_{pgt}, and D_y denote dummies for product-groups and year, respectively; $v_t^{j^*}$ denotes stochastic demand shocks and residuals, $v_t^{j^*} \equiv \tilde{\xi}_t^j + u_t^{j^*}$; β_{sq} denotes coefficients on $s_s^j \beta_q$. However, equation (9) is not directly estimated due to the correlations between unobserved productivity and freely adjustable inputs.

The specialization of production, a_t^j, is not directly observable because the data showing a product's input use is not available. As a proxy for a product's input use, a product' deflated revenue share is used. If a firm allocates large portions of inputs to a particular product, the output of that product would be larger than any other product within a firm. Hence, revenue falls under the assumption since input prices are common across products, and that mark-ups are constant over marginal costs as in CES preference. As an instrument for the specialization of production, the Herfindhal Index of product-level deflated revenue, $h_t^j = \sum_i \left(\frac{\tilde{r}_{it}^j}{\sum_i \tilde{r}_{it}^j} \right)^2$, is used to capture the degree of concentration in production over products within a firm. The Herfindhal index has the lowest value if a firm allocates its inputs equally across products and the index increases as a firm skews its inputs toward a particular product. Revenue-based Herfindhal index yields a similar measure of specialization of production.

Estimating equation (9) through Least Squares would yield biased results because of correlation between unobserved productivity and freely adjustable inputs. Firms tend to use more freely adjustable input when they have positive productivity shocks, and those freely adjustable inputs include labor and material. This correlation causes endogeneity problem and lead biased results.

Basically, this chapter estimates equation (9) through a two-stage-procedure to deal with the correlation problem which is introduced by Levinsohn and Petrin (2003) and Loecker (2011). Material is used as proxy for unobserved productivity and is given by,[5]

(10) $m_t^j = m_t\left(k_t^j, a_t^j, n_t^j, q_{st}, D_{pt}, D_{yt}, \varphi_t^j\right)$.

Material demand is monotonically correlated with unobserved productivity so it can be used as a proxy for unobserved productivity (Levinsohn and Petrin (2003); Loecker (2011)). By inverting the material demand function, a function, $h_t(\cdot)$, is defined as,

(11) $\varphi_t^j = h_t^j\left(k_t^j, l_t^j, m_t^j, a_t^j, n_t^j, q_{st}, D_{pt}, D_{yt}\right)$.

Now unobserved productivity is a function of inputs, the specialization of production, the number of production, total demand of sector and dummies for product and year. If I use material as proxy for unobserved productivity, revenue does not have a linear relationship with inputs as in equation (9) due to the correlation between material and other variables as in equation (10). I rewrite the equation (9) as the following,

(12) $\tilde{r}_t^j = \Lambda_t^j(\,X_1, X_2) + {v_t^j}^*$

where

(13) $\Lambda_t^j(X_1, X_2) = \sum_{a=0}^{4}\sum_{b=0}^{4-a}\omega_{ab}\left(X_1\left(k_t^j, l_t^j, m_t^j, h_t^j(\cdot)\right)\right)^a\left(X_2\left(\,a_t^j, n_t^j, q_{st}, D_{pt}, \delta_y D_y\right)\right)^b$

where $X_1(\cdot)$ includes all inputs in production and $X_2(\cdot)$ includes other variables. Note that interaction terms with dummies for product-group and year are not considered due to a practical implementation problem. The non-linear term, $\Lambda_t^j(\,X_1, X_2)$, is

[5] Investment can be considered as a proxy for unobserved productivity as well. However, data constraints restrict the use of investment as the proxy for productivity. Data has a significant number of firms report zero investment.

approximated by fourth-order polynomials and estimated by OLS. Then, fitted value of the non-linear term, $\widehat{\Lambda_t^j(\cdot)}$, is calculated. Note that the fitted value does not include the residual, υ_t^{j*}.

In the first stage, unobserved productivity is calculated using the fitted value of the non-linear term as the following,

$$(14)\ \widehat{\varphi_t^{j}}^{*}(\beta_k^*,\ \beta_l^*, \beta_m^*, \beta_a^*,\ \beta_n^*,\ \beta_{sq}^*, \delta_p^*, \delta_y^*)$$

$$= \widehat{\Lambda_t^j(\cdot)} - \beta_k^* k_t^j - \beta_l^* l_t^j - \beta_m^* m_t^j - \beta_a^* a_t^j - \beta_n^* n_t^j - \beta_{sq}^* q_{st} - \delta_p^* D_{pt} - \delta_y^* D_{yt}\ .$$

where $\beta_k^*,\ \beta_l^*,\ \beta_m^*,\ \beta_a^*,\ \beta_n^*, \delta_p^*$ and δ_y^* are unbiased estimates of inputs and demand variables. To calculate the initial value of the calculated productivity, I use the coefficients of inputs and other variables by estimating equation (9) via OLS. Using the current and lagged productivity $\left(\widehat{\varphi_t^{j}}^{*},\ \widehat{\varphi_{t-1}^{j}}^{*}\right)$ from equation (14), the productivity determination process is estimated using the following,

$$(15)\ \widehat{\varphi_t^{j}}^{*} = g_t\left(\widehat{\varphi_{t-1}^{j}}^{*},\ a_{t-1}^j,\ n_{t-1}^j\right) = \beta_1\ \widehat{\varphi_{t-1}^{j}}^{*} + \beta_2 a_{t-1}^j + \beta_3 n_{t-1}^j + \eta_t^j\ .$$

Let unbiased estimates from equation (15) be β_1^*, β_2^* and β_3^*. The stochastic *i.i.d.* component which is interpreted as innovation or productivity shocks, $(\widehat{\eta_t^j})$, is calculated using the estimates from equation (15),

$$(16)\ \eta_t^{j*} = \widehat{\varphi_t^{j}}^{*} - g_t\left(\widehat{\varphi_{t-1}^{j}}^{*},\ a_{t-1}^j,\ n_{t-1}^j\right) = \widehat{\varphi_t^{j}}^{*} - \beta_1^*\ \widehat{\varphi_{t-1}^{j}}^{*} - \beta_2^* n_{t-1}^j - \beta_3^* a_{t-1}^j$$

In the second stage, unbiased coefficients of inputs and demands are estimated. As mentioned above, equation (9) is not estimable due to the endogeneity problem in equation (9). The independent condition between independent variables and error terms are needed. I define a new moment condition using the independency between a stochastic term in the productivity determination process in equation (8) and inputs.

The moment condition assumes that some inputs and the specialization of production could be correlated unobserved productivity, but those variables are weakly correlated with unobserved productivity shocks, since firms are not able to anticipate productivity shocks, and hence, they are not able to change their inputs uses or the specialization of production.[6] However, to avoid possible correlation, I take one period time lag. The moment equation is used as such,

$$(17)\quad E\left\{\left(\eta_t^j\left(\beta_k,\ \beta_l,\ \beta_m,\ \beta_a,\ \beta_n,\ \beta_{sq},\delta_p,\delta_y\right)\right)\begin{pmatrix} k_t^j \\ l_{t-1}^j \\ m_{t-1}^j \\ a_{t-1}^j \\ n_{t-1}^j \\ q_{st-1} \\ D_{pt} \\ D_{yt} \end{pmatrix}\right\} = 0 \quad ,$$

where variables for freely adjustable inputs and the specialization of production are lagged due to correlation between productivity and those inputs. The moment condition estimated through Generalized Method of Moments (GMM). The two-stage-procedure is iterated until the moment condition is converged to zero.

2.3.2 Estimation Strategy for Multi-product Firms' TFP

The original unobserved productivity shocks can be recovered by,

[6] Freely adjustable inputs, the specialization of production, and sectors' total demand are lagged due to the correlation between unobserved productivity shocks, $\eta_t^j(\cdot)$, and corresponding variables. Freely adjustable inputs (labor and material) have the correlation with productivity shocks because firms with high productivity shocks tend to use more freely adjustable inputs than those with low productivity shocks. The specialization of production correlate with productivity shocks since productivity shocks might change the shares of each product even though the physical output of all products within a firm increase at the exact same amount. Sectors' total demand is correlated is a weighted sum of firm-level production, and hence it is correlated firm-level productivity shocks.

$$(18)\quad \widehat{\varphi_t^j} = (\tilde{r}_t^j - \beta_k{}^* k_t^j - \beta_l{}^* l_t^j - \beta_m{}^* m_t^j - \widehat{\beta_a}\, a_t^j - \widehat{\beta_n}\, n_t^j - \widehat{\beta_{sq}}\, q_{st}$$

$$-\hat{\delta}_p\, D_{pt} - \hat{\delta}_y\, D_y)\left(\frac{\sigma_s}{1+\sigma_s}\right)$$

where σ_s is obtained by $\sigma_s = -\frac{1}{|\beta_{sq}|}$; as for multi-product firms, the shares of physical output in sector s are used as a weigh $\sum_s \left(\frac{\sigma_s}{1+\sigma_s}\right)\frac{r_{st}^j}{\sum_s r_{st}^j}$, instead of $\left(\frac{\sigma_s}{1+\sigma_s}\right)$.

2.4 DATA

2.4.1 Firm- and Product-level Data

This chapter uses the PROWESS database (PROWESS) on Indian firms over the 1989 – 2009 periods. These data have been collected by the Centre for Monitoring the Indian Economy (CMIE), and show the production and financial performance of companies in India from 1988 to present. As per the Companies Act in 1956, all Indian business entities are required to report production and financial information. CMIE has been collecting this information using its own classifying system for product and industry (PROWESS code). Goldberg et al. (2009, 2010a, 2010b), Topalova and Khandelwal (2011), and Loecker et al. (2012) use the same database as well, but their data period is 1989 - 2003.[7]

---

[7] Revenue, Capital, and Material are deflated in this chapter. Product-level revenue is deflated by the corresponding sector's Indian Whole Sale Price Index (WPI). Capital stock is net fixed assets deflated by corresponding sectors' WPI. Materials are expenses of raw or intermediate goods and of water and electricity. Materials are deflated by the corresponding sector's primary article index in WPI. As for labor, I calculate the number of employees using reported information. As for multi-sector firms, firms' market share in each sector is used as a weight. About 30 percent of firms in data report both the number of employees and total wage, while the other 70 percent of firms report total wage only. First, the average wage rate is calculated using the information from the 30% percent of firms. By dividing total wage by the number of employees, average wage rate is calculated. Then using the average wage rate, the actual number of employees is calculated. As for WPI, it has been collected by the Office of the Economic Advisor in the Department of Industrial Policy and Promotion, Ministry of Commerce and Industry. 435 representative products in all industry except service are reported by 13-digit Indian National Industrial Classification code (NIC). NIC code is converted to SITC by taking a weighted

PROWESS defines product by a 20-digit PROWESS code and firms by name. This chapter converts PROWESS code to 5 digit Standard International Trade Classification (SITC) code. Note that other existing literature using PROWESS converted it to Indian National Industrial Specification Code. The hierarchical structure of industrial classification is as follows: product is defined by 5-digits in SITC coding. Product-group is defined by 3- digits in SITC coding. Product-group includes most products using identical raw materials. Industry is defined by 2-digit SITC codes and Sector is defined by 1-digit SITC codes. The classification of intermediate goods is made by the use of raw materials, but that of final goods made by purpose and function of goods. As per the SITC code, there are 2,970 products, 1,023 product-groups, 67 industries and 10 sectors.

The advantage of PROWESS is the availability of product-level revenue by firm. The history of product range and the specialization of production can be tracked from 1989 to 2009 periods. The disadvantage is that PROWESS is not a comprehensive database and covers mid- and large- size companies in the organized sector. However, this database includes firms and products accounting for 60-70% of total economic activity in India. Although small firms are not included in the PROWESS, this chapter focuses on intra-firm resource reallocation on productivity. Most small firms are likely to be single-product firms and thus, intra-firm resource reallocation across products may not be a relevant issue.

The original PROWESS covers all industries including the service industry, but this chapter does not include service industry. Data from 1988 and 2010 are dropped as well because of few observations, which were not consistent with data for other

average over all products where weights are based on market shares.

years in the sample. In addition, all firms having the inconsistent accounting year problem are dropped with the exception of firms with financial accounting years. In PROWESS, 80% of firms in PROWESS follow the financial year from April of a calendar year to March of the following year, 15% of firms follow the calendar year from January to December of the accounting year. Most foreign firms use the calendar year. After aggregating the original database by the hierarchy structure of industrial classification, a whole firm's observations having missing years or missing variables are dropped. There is no systematic structure generating these missing variables problems.

2.4.2 Measuring Sector's Total Demand

Following Klettte and Griliche (2005) and Loecker (2011), each Sector's total demand, Q_{st}, is measured by a weighted average of deflated revenue.

(19) $q_{st} = \sum_{j=1}^{J_s} Market\ Share_{st}^{j} \times \widetilde{r_{st}^{J}}$

where J_s denotes the number of firms in sector *s*; $Market\ Share_{st}^{j}$ is the market share of firm *j* in sector *s*; $\widetilde{r_{st}^{J}}$ denotes logged deflated revenue of firm *j* in sector *s*. The market share of firm *j* in sector *s* is measured as,

(20) $r_{st}^{j} = log\left(R_{st}^{j}\right) = log\left(\sum_{i \in I_s} R_{ist}^{j}\right)$,

(21) $Market\ Share_{st}^{j} = \frac{R_{st}^{j}}{\sum_j R_{st}^{j}}$,

where I_s denotes the set of products in sector s; R_{ist}^{j} denotes deflated revenue of product *i* by firm *j* in sector *s*; R_{st}^{j} denotes deflated revenue of firm *j* in sector *s*.

2.4.3 Measuring Intra-firm Resource Reallocation

In order to measure intra-firm resource reallocation, I adopt two variables; the number of products and the specialization of production. The number of products is measured by counting the number of products after aggregating the data by 5-digit SITC code. The definition of products is different from that of brand names. A 5-digit product category could include multiple brand names. Even though a firm frequently changes a brand name, the product that belonged to a brand is continually produced without any change in its status (Baldwin and Harrigan, 2011).

The specialization of production is measured by the Herfindhal Index which is the sum of squares of the product-level deflated revenues by sector-specific producer price index. For a single product firm, the specialization of production has unity. If all products within a firm are evenly produced, it has the lowest value which is the inverse value of the number of products within a firm. Using the product-level deflated revenue enables us to avoid quantity aggregation problem because it is quite complicated to aggregate the physical output of multiple products because the units of products are different. Some products use weight as their unit, but the other products use count as their unit. However, I do not observe the unit of each product in the CMIE database.

2.4.4 Data Description

In this section, I present the basic statistics of data. Table 2.1 illustrates the variation in revenue, capital, labor, material, the specialization of production and product range across 35,423 observations, 4,926 firms and an average of 7.19 producing years per firm from 1989 to 2009 periods.

Considerable temporal variation is visible in these data: between-year standard deviations of all variables are greater than 1.45, except for the specialization of production. However, the specialization of production ranges from 0 to 1, so, it varies considerably as well. This table highlights the dramatic growth of Indian firms. Within-year standard deviations show the extent of firm heterogeneity in the data.

TABLE 2.1 The Variation in Revenue, Capital, Labor, Material, Specialization of Production and Product Range

Variable		Mean	Std. Dev.	Min	Max	Observations
Revenue	overall	5.30	1.90	-2.84	13.77	N =35,423
	between		1.92	-2.43	13.07	n = 4,926
	within		0.64	-3.02	8.56	T-bar = 7.19
Capital	overall	3.59	1.95	-5.94	12.20	N =35,423
	between		1.93	-4.35	11.86	n = 4,926
	within		0.78	-3.22	9.10	T-bar = 7.19
Labor	overall	4.25	1.88	-1.73	12.32	N =35,423
	between		1.83	-1.69	12.19	n = 4,926
	within		0.50	-0.07	7.57	T-bar = 7.19
Material	overall	4.23	2.08	-3.31	13.22	N =35,423
	Between		2.14	-3.31	12.00	n = 4,926
	Within		0.75	-4.16	8.74	T-bar = 7.19
Specialization of Production	Overall	0.85	0.22	0.10	1.00	N =35,423
	Between		0.18	0.16	1.00	n = 4,926
	Within		0.10	0.18	1.53	T-bar = 7.19
Product Range	Overall	2.12	1.93	1.00	42.00	N =35,423
	Between		1.45	1.00	35.05	n = 4,926
	Within		0.73	1.00	17.36	T-bar = 7.19

Notes. All variables are logged values except the specialization of production and product range. Between Std. Dev. indicates the variation of average values over years and Within-Std. Dev. indicates the variation from the overall average at a given year. There are 35,423 observations, 4,926 firms and average of 7.19 producing years production per firm over the 1989 – 2009 periods. Revenue is deflated by sector’s Producer Price Index.

Table 2.2 reports the number of firms, average product range and average the specialization of production by time. The total number of firms has increased from 334 to 2,060 during the sample period, and 144 firms were present throughout the sample period.

TABLE 2.2 Firms, Product Range and Specialization of Production

Year		1989	1991	1993	1995	1997	1999	2001	2003	2005	2007	2009
1. The number of Firms												
Firms	Single	107	252	517	767	755	811	885	1,235	1,305	1,206	1,005
	Multi	227	383	645	785	818	925	908	1,159	1,211	1,233	1,055
	Total	334	635	1,162	1,552	1,573	1,736	1,793	2,394	2,516	2,439	2,060
2. Average Product Range with and without Single Product Firms												
With Single Product Firms		3.21	2.59	2.29	2.13	2.17	2.15	2.09	1.97	1.99	2.07	2.13
Without Single Product Firms		4.26	3.63	3.33	3.24	3.26	3.16	3.15	3.00	3.06	3.11	3.21
3. Average Specialization of Production with and without Single Product Firms												
With Single Product Firms		0.79	0.82	0.83	0.84	0.84	0.85	0.85	0.87	0.86	0.86	0.85
Without Single Product Firms		0.69	0.71	0.69	0.69	0.69	0.72	0.71	0.72	0.72	0.72	0.71

Notes. This table shows the average statistics of firms, specialization of production and product range. Row 1: summaries the number of firms by year. Single indicates single-product firms and Multi indicates multi-product firms. Row 2: summaries the average number of products per firm. Row 3: shows the degree of concentration in production over products. The specialization of production ranges from 0 to 1.

The share of single-product firms has also increased from 32.0% in 1989 to 48.8% in 2009. Note that the share of single-product firms is relatively smaller than that in the United States. For example, Bernard, Redding and Schott reported that the share of single-product firms is 61% during 1987 to 1997 in U.S. In India, the average product range varies between 1.97 and 3.21, but the average specialization of production varies between 0.79 and 0.87. Regardless of the inclusion of single-product firms, the trend in both variables is identical. This highlights specialization of Indian firms’ production toward particular products. Work by Melitz (2003), Mayer, Melitz and Ottaviano (2011) and Bernard, Redding and Schott (2011) pointed out that firms tend to reallocate their resource toward particular products in response to intensifying competition following trade liberalization.

Table 2.3 and 2.4 summarize the share of firms by changes in product range and the specialization of production. On average, 15.2% firms change their product range during the sample period. This means that firms are more likely to change their specialization of production than product range.

TABLE 2.3 Changes in Product Range

	1990	1992	1994	1996	1998	2000	2002	2004	2006	2008
1. The Share of Firms by Changes in Product Range, %										
-5+	0.2	0.2	0	0.1	0.1	0.1	0.2	0	0.1	0.1
-4	0	0.1	0.1	0.1	0.1	0	0.1	0	0.1	0
-3	0.4	0.5	0.6	0.4	0.3	0.3	0.2	0.3	0.2	0.1
-2	1.0	1.0	0.6	0.8	0.8	1.2	0.9	0.5	1.0	0.9
-1	3.5	5.1	4.6	6.0	6.0	6.7	6.1	5.1	4.7	5.7
0	47.8	55.0	57.7	72.9	69.7	76.5	63.0	74.6	78.7	79.1
1	5.4	7.3	7.4	7.6	7.8	5.5	5.9	6.8	6.9	7.1
2	0.6	1.1	1.7	1.8	1.5	0.7	1.1	0.9	1.6	1.2
3	0.2	0.4	0.6	0.4	0.3	0.2	0.2	0.2	0.4	0.1
4	0	0.2	0.1	0.1	0	0.1	0.1	0.1	0.1	0.1
5+	0.2	0.1	0.1	0.1	0.1	0	0.1	0.1	0	0.1
2. The Share of New Born Firms, %										
New Firms	40.7	29.0	26.5	9.7	13.3	8.7	22.1	11.4	6.2	5.5
3. Total Number of Firms										
Total Firms	518	826	1,443	1,572	1,683	1,726	2,132	2,494	2,494	2,341

Notes. Row 1: summarizes the share of firms by changes in product range. The first column shows the changes in product range, and the other columns show corresponding shares of firms by %. The value -1 in the first column indicates a firm reduces a product range by -1 compared to previous year product range. Row 2: summaries the share of the newly establishing firms. The sum of Row 1 and Row 2 is 100 %. Row 3: shows the total number of firms by year.

TABLE 2.4 Changes in the Specialization of Production

	1990	1992	1994	1996	1998	2000	2002	2004	2006	2008
1. The Share of Firms by the Specialization of Production, %										
-0.8 ~ -0.7	0.2	0	0	0.1	0	0	0	0	0	0.1
-0.7 ~ -0.6	0	0.1	0.1	0.2	0.1	0	0.1	0.2	0	0
-0.6 ~ -0.5	0	0.2	0.1	0.2	0.2	0	0.1	0.2	0.1	0
-0.5 ~ -0.4	0.2	0.7	1.0	0.6	0.8	0.5	0.3	0.7	0.8	0.4
-0.4 ~ -0.3	0.2	1.2	0.8	0.9	0.8	0.9	0.7	0.7	0.7	0.6
-0.3 ~ -0.2	1.0	0.7	2.3	1.7	1.4	1.3	1.7	0.9	1.7	0.9
-0.2 ~ -0.1	2.3	3.3	4.2	4.1	3.6	2.7	2.6	3.0	3.0	3.7
-0.1 ~ 0	16.0	18.4	16.3	19.3	18.4	20.1	15.6	17.5	19.5	20.7
0 ~ 0.1	35.1	43.1	43.3	57.8	54.8	60.5	50.9	60.0	63.2	62.7
0.1 ~ 0.2	2.7	2.2	2.9	2.5	3.5	3.0	3.1	3.0	2.5	2.6
0.2 ~ 0.3	0.6	0.5	1.2	1.5	2.0	1.1	1.3	1.1	1.1	1.4
0.3 ~ 0.4	0.4	0.2	0.6	0.6	0.5	0.6	0.8	0.6	0.6	0.6
0.4 ~ 0.5	0.6	0.3	0.5	0.6	0.5	0.3	0.5	0.5	0.4	0.6
0.5 ~ 0.6	0	0	0.2	0.2	0.1	0.2	0.1	0.2	0.1	0.1
0.6 ~ 0.7	0	0.1	0	0	0	0.1	0.1	0	0.1	0.1
2. The Share of New Born Firms, %										
New Firms	40.7	29.0	26.5	9.7	13.3	8.7	22.1	11.4	6.2	5.5
3. Total Number of Firms										
Total Firms	518	826	1,443	1,572	1,683	1,726	2,132	2,494	2,494	2,341

Notes. Row 1: summarizes the share of firms by changes in the specialization of production. The first column shows the range of changes in the specialization of production, and the other columns show corresponding shares of firms by percentage. The range of -0.8~-0.7 in the first column indicates a firm reduces its specialization of production by the value between -0.8 and -0.7 as compared to the previous specialization of production. Row 2: summaries share of the newly established firms. The sum of Row 1 and Row 2 is 100 %. Row 3: shows the total number of firms by year.

Table 2.5 shows the correlation between product range and the specialization of production. Mayer, Melitz and Ottaviano (2011), and Bernard, Redding and Schott (2011) did not distinguish these two variables explicitly because they show identical information. The VIF between product range and the specialization of production is 1.86 which is smaller than 5; the R-square is 0.46, and the correlation coefficient is -0.68. These two variables are negatively correlated, but one variable only explains the half of the other variable.

TABLE 2.5 Correlation Diagnostic between Product Range and the Specialization of Production

Correlation Coefficient	Tolerance	R-Squared
-0.68	0.54	0.46

2.5 Results

This section documents the estimation results of production function and the testing results of two hypotheses about the behavior and structure of multi-product firms. First, the estimation and results of the Cobb-Douglas production with CES preference is outlined. Then, calculated (unobserved) TFP from estimating the production function is discussed. Next, this section examines the relationship between TFP and revenue, the specialization of production and product range. Finally, the link between intra-firm resource reallocation and multi-product firms’ TFP is estimated. Those hypotheses, as noted earlier, are drawn from work by Bernard, Redding and Schott (2011) and Mayer, Melitz and Ottaviano (2011).

2.5.1 Production Function with Sectors’ Total Demand and Intra-firm Resource Reallocation

The Cobb-Douglas production function with CES preference in equation (9) is specified with the following variables: deflated revenue, capital, material, sectors’ total demand, and intra-firm resource reallocation (the number of products and the specialization of production). Fixed effects for product-group and for year are applied and all variables are in logs except fixed effect terms.

Table 2.6 reports estimation results of the production function with sectors' total demand and intra-firm resource reallocation. I compare this estimation results with a few baseline specifications. Approach I denotes the general revenue production estimation [column (1)]. Approach II denotes Loecker's (2011) specification [column (2)], Approach III denotes our specification [column (3)].

Table 2.6 Estimates of Production Function with Sectors' Total Demand and Intra-firm Resource Reallocation

Coefficients On	(1) Approach I	(2) Approach II	(3) Approach III
Capital	0.1008*** (0.0020)	0.0704*** (0.0019)	0.0669*** (0.0021)
Labor	0.3308*** (0.0023)	0.3312*** (0.0023)	0.3394*** (0.0024)
Material	0.5871*** (0.0021)	0.5783*** (0.0020)	0.5751*** (0.0021)
Specialization of Production	Not included	0.070*** (0.0088)	0.1228*** (0.0135)
Product Range	Not included	Not included	0.1471*** (0.0082)
Sector's Total demand (9 sectors)	Not included	All positive***	All positive***
Fixed effects of Products and Year	Included	Included	Included
No. of obs	35,423	35,423	35,423

Notes. The dependent variable is the deflated revenue. The dependent variable is the deflated revenue. Capital is deflated by sector-specific Whole Sale Price Index (WPI). Labor is the number of employees. Material is deflated by primary article index in WPI. Approach I in column (1) is the results from the general revenue production function where only capital, labor and material are consider. Approach II in column (2) is the results from Loecker's (2011) specification where the revenue production function is estimated with inputs, sectors' total demand and the specialization of production. Instead product range, the specialization of production is used because I relax the assumption that inputs are evenly spread across products by the inverse of the number of products within a firm. Approach III in column (3) is the results from our specification where the revenue production function is estimated with sectors' total demand, the specialization of production and product range. Standard Errors are in parentheses. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Columns (1)-(2) show the estimation results from other specifications. Column (1) shows the estimation results from the general revenue production function where the specialization of production, product range, and sectors' total demand are not considered. All estimates in column (1) are biased due to price effects in revenue are not controlled. Note that coefficients in column (1) are not weighted by the elasticity of substitution.

Column (2) shows the estimation results from Loecker's (2011) specification where only the specialization of production is considered as a measure of inputs allocation to each product within a firm. In his paper, Loecker assumes that inputs are equally allocated to each product at the ratio of the inverse of product range, and so he uses product range as a measure for inputs allocation to each product. However, I relax this assumption and use the specialization of production as a measure of inputs allocation to each product.

Column (3) shows the estimation results from our specification, where both product range and the specialization of production are considered. The estimation result shows that revenue is affected by technology (three inputs), allocation of inputs to products (the specialization of production) and internalized demand linkage (product range). Both skewing production toward a particular product and expanding product range increase revenue. The specialization of production represents the way to allocate inputs across products. As a multi-product firm skews inputs toward a particular product, the firm's aggregated production increases due to returns to scale. At the same time, however, price of the particular product decreases as its supply increases. Positive sign on the specialization of production implies that revenue increase from the supply enlargement is bigger than revenue decrease from the price reduction, and as a result, a multi-product firm's aggregated revenue increases as the firm increases the specialization of production. Expanding product range could take demand away from the existing product(s) within a firm due to cannibalization effects. However, positive sign on product range implies that cannibalization plays a significant role for multi-product firms, and that expanding product range ultimately leads increase in aggregate revenue because consumers' recognition to the firm is sufficiently favorable.

The coefficients of inputs and sectors' total demand in column (2) are not biased, even though Loecker's (2011) method does not consider the internalized demand linkage of multi-product firms, because all coefficients are estimated by the moment condition in equation (17), where all variables are independent from productivity shocks. In this case, however, the effects of product range on revenue would be included in TFP because TFP is calculated by subtracting the explanatory part by considered independent variables from revenue.

Table 2.7 reports that detailed estimation results of sectors' total demand, β_{sq}, implied elasticity of substitutions and calculated markups. I allow the elasticity of substitution vary by sector. Note that the implied elasticity is calculated by $\boldsymbol{\sigma_s} = -\frac{1}{\boldsymbol{\beta_{sq}}}$. The estimation result shows that the elasticity substitutions vary between -9.4162 and -4.3459 in Approach III, and that agricultural sectors (sector 0 and sector 4) are relatively inelastic. The returns to scale can be calculated based on the results in Table 2.6 and 2.7. The returns to scale vary by sector. In Approach III, Mineral Fuels sector has $\gamma = 1.1166$ which is the lowest returns to scale, and Animal and vegetable oils sector has $\gamma = 1.2747$ which is the largest returns to scale.

Table 2.7 Estimates of Sector's Total Demand Parameters

Demand Controls	(1) Approach II			(2) Approach III		
	Coefficient $(\boldsymbol{\beta_{sq}})$	Implied Elasticity $(\boldsymbol{\sigma_s})$	Implied Markup $\left(\frac{\boldsymbol{\sigma_s}}{\mathbf{1}+\boldsymbol{\sigma_s}}\right)$	Coefficient $(\boldsymbol{\beta_{sq}})$	Implied Elasticity $(\boldsymbol{\sigma_s})$	Implied Markup $\left(\frac{\boldsymbol{\sigma_s}}{\mathbf{1}+\boldsymbol{\sigma_s}}\right)$
Food and Live Animals (s=0)	0.1684*** (0.0018)	-5.9382	1.2025	0.1817*** (0.0018)	-5.5036	1.2220
Beverage and Tobacco (s=1)	0.1504*** (0.0035)	-6.6489	1.1770	0.1463*** (0.0034)	-6.8353	1.1714
Crude Material (s=2)	0.1079*** (0.0018)	-9.2678	1.1210	0.1105*** (0.0018)	-9.0498	1.1242
Mineral Fuels (s=3)	0.0940*** (0.0021)	-10.6383	1.1038	0.1062*** (0.0021)	-9.4162	1.1188
Animal and Vegetable Oils (s=4)	0.2348*** (0.0039)	-4.2589	1.3068	0.2301*** (0.0039)	-4.3459	1.2989
Chemicals and Related Products (s=5)	0.1310*** (0.0011)	-7.6336	1.1507	0.1351*** (0.0011)	-7.4019	1.1562
Manufactured goods (s=6)	0.1296*** (0.0009)	-7.7160	1.1489	0.1355*** (0.0009)	-7.3801	1.1567
Machinery and Transport Equipment (s=7)	0.1305*** (0.0010)	-7.6628	1.1501	0.1346*** (0.0010)	-7.4294	1.1555
Miscellaneous Manufactured Articles (s=8)	0.1340*** (0.0015)	-7.4627	1.1547	0.1400*** (0.0015)	-7.1429	1.1628
Other Trans Actions and Goods (s=9)	0.1684*** (0.0018)	-5.9382	1.2025	0.1772 (0.0268)	-5.6433	1.2154

Notes. This table shows the detailed estimation results of sectors' total demand in the revenue production function. All variables are logged. Approach II in column (1) is the results from Loecker's (2011) specification where the revenue production function is estimated with inputs, sectors' demand and the specialization of production. Instead product range, the specialization of production is used because I relax the assumption that inputs are evenly spread across products by the inverse of the number of products within a firm. Approach III in column (2) is the results from our specification where the revenue production function is estimated with inputs, sectors' demand, the specialization of production, and product range. Standard Errors are in parentheses. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

There is no significant difference between Approach II and Approach III since both method use the moment condition. Implied markups vary between 1.1188 and 1.2989. As discussed above, implied markups play a role in calculating TFP because all coefficients in Table 2.6 are weighted coefficients by the markups. In order to recover the original TFP, calculated TFP is weighted by markups as in equation (18).

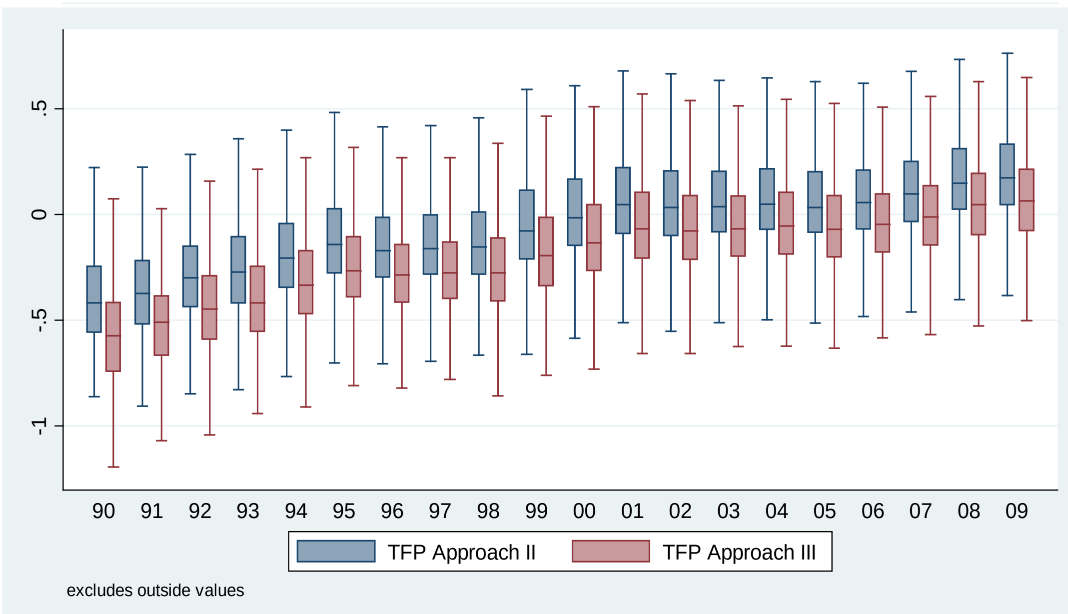
2.5.2 Calculation of Productivity

TFP is calculated by equation (18). Instead of including fixed effects for product-group and times in the revenue production function, those are included in the productivity determination process due to practical implementation problems. After estimating the revenue production function, fixed effects terms are subtracted from fitted revenue to recover TFP.

Figure 2.1 shows the distribution and trends of logged TFPs from different approaches. The distribution is represented by the shape of the box and the length of two lines. It looks that the lengths of upper and lower lines are quite similar, but this means that TFP is widely spread on upside because TFP is in logged values. Figure 2.1 also shows TFP from Approach II is larger than that from Approach III. The gaps between these two TFPs vary by year. The gap was 0.1919 in 1990 and it was decreased to 0.1196 in 2002 which is the lowest value throughout the data period. Then, it rose to 0.1226 in 2009.

TFP from Approach II is overestimated due to the omission of product range in the revenue production function. Product range is added to the revenue production function to account the cannibalization effect of multi-produce firms. Table 2.1 shows that the coefficient of product range is positive and statistically significant. If the product range is not considered in the revenue production function, then the effect of the product range would be included in TFP because TFP is calculated as residuals after estimating the production function. Approach II yields larger TFP because it does not account the cannibalization effect in the revenue production function. The size of overestimation varies as product range changes.

Figure 2.1 Distribution of Total Factor Productivity (1989-2009)



2.5.3 Analyzing the Effect on Intra-firm Resource Reallocation on Productivity

Next, this chapter explores the impact of intra-firm resource reallocation on TFP as in Hypothesis I,

Hypothesis I. Concentrating production system increases multi-product firms' productivity, while diversifying it decreases productivity, i.e. Discontinuing a product and (/or) skewing production toward a particular product increase a firm's productivity, while adding a product and (/or) equalizing the production of all products decreases productivity.

Hypothesis I reflects the research questions in emerging literature: Eckel and Neary (2003), Bernard, Redding and Schott (2011) and Mayer, Melitz, and Ottaviano (2011). However, I do not have to implement an additional estimation to test Hypothesis I, since the productivity determination process as in equation (8) already embeds it.

Table 2.8 reports the estimation results of the productivity determination process. It has the log-linearized form, and so the coefficients are interpreted as elasticity, Coefficients are interpreted as percentage changes because the productivity determination process assumes a first-order Markov process. Approach III shows that previous TFP accounts about 86 percent of current TFP, that TFP increases at a constant rate by 1% every year, and that changes product range has account about -1% of TFP. The specialization of production has insignificant effect on TFP, and its support covers negative and positive ranges as well. However, Approach II shows the different results that both the specialization of production and product range have insignificant on productivity. This is due to that Approach II does not single out the effects of product range which is multi-product firms' internalized demand linkage, when it measures TFP using the revenue production function. The results in Table 2.8 also demonstrate that controlling for the multi-product firms' internalized demand linkage is crucial to measure unbiased TFP. Otherwise, TFP from Approach II is more likely to be overestimated and the effect of product range on TFP is likely to be distorted.

The results in Table 2.8 partially confirm the Hypothesis I. I find that only changing product range has an impact on productivity and skewing production does not. As in Table 2.8, skewing production has an impact only on revenue. Our findings corroborate the debates in the recent literature: work by Eckel and Neary (2003) and Mayer, Melitz and Ottaviano (2011). Both studies suggest that expanding product range has negative effects on productivity. However, our finding is partially confirms the finding in Bernard, Redding and Schott (2011). Bernard, Redding and Schott (2011) suggest that both the specialization of production and product range have significant effects on productivity.

Table 2.8 Impact of Intra-firm Resource Reallocation on Productivity

Coefficient on	(1) Approach II	(2) Approach III
Previous TFP	0.8729*** (0.0026)	0.8688*** (0.0027)
Previous Specialization	-0.0034 (0.0038)	-0.0058 (0.0037)
Previous Product Range	-0.0007 (0.0022)	-0.0107*** (0.0023)
Constant	0.0297*** (0.0011)	0.0174 (0.0012)
No. of obs	30413	30413
Adjust R^2	0.79	0.78

Notes. This table examines the productivity determination process. It exploits calculated TFP as a residual after generating the Cobb-Douglas production function with sectors' total demand, the specialization of production and product range. Approach II in column (1) is the results from Loecker's (2011) specification where the revenue production function is estimated with inputs, sectors' total demand and the specialization of production. Instead product range, the specialization of production is included because I relax the assumption in Loecker (2011) that inputs are evenly spread across products by the inverse of the number of products within a firm. Approach III in column (2) is the results from our specification where the revenue production function is estimated with inputs, sectors' total demand, the specialization of production and product range. All variables are logged. Standard Errors are in parentheses. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

2.6 Conclusion

I propose an approach to estimate unbiased TFP of multi-product firms and examine the behavior and structure of multi-product firms by establishing two hypotheses using a rich database on Indian firms. I find that TFP is likely to be over-estimated when the revenue production function is estimated without considering multi-product firms' internalized demand linkage because TFP contains the effects of multi-product firms' internalized demand linkage in it. I also find that firms' productivity does not depend only on a firm's technology usages, but also intra-firm resource reallocation. TFP increases with reducing product range. I confirm that TFP is the major source explaining different revenue, profit, product range and production combination across firms. More successful firms produce more output

These findings have three main implications. First, firms indirectly optimize productivity through intra-firm resource reallocation to maximize profit. Second, productivity is the major source explaining multi-product firms' production and export status. Finally, multi-product firms' productivity could increase soon after receiving export status because they reallocate resources toward a newly exporting product.

3 Multi-Product Firm, Firm-level Gravity and Self-Selection

3.1 Introduction

A substantial body of literature in international trade economics has analyzed the determinants of trade flows between countries. One of the prominent methods in international trade economics is the gravity equation, which predicts that the bilateral trade between two countries is proportional to their economic size and trade resistance. The strength of the gravity equation is in its simplicity and goodness of fit to the empirical data. The gravity equation is relatively easy to derive, and fits well with most trade flows in empirical analysis. The gravity equation has estimated border effects, the home market effects, and the impact of trade resistance including distance, tariffs and language on trade flows.

The first theoretical derivation of the gravity equation is from Anderson (1979). Helpman (1987) then shows that trade flows between two countries of similar size are larger than those between different country sizes. McCallumn (1995) compares the trade flows between Canadian Provinces to those between Canadian provinces and U.S. states. Anderon and Wincoop (2003, 2004) use fixed effects to account for the price effects in the gravity equation, and provide the standard model in empirical analysis. Helpman, Melitz and Rubinstein (2008) propose the estimation strategy to analyze the trade flows considering zero trade flows. However, most empirical works with the gravity equation analyze country-level bilateral or multilateral trade flows using aggregated trade data at country-level or at most industry-level (I refer to this as the country-level gravity equation).

Presently, as firm-level data becomes available, studies in international trade dominantly adopts the Melitz-type heterogeneous firm model to analyze the impact of trade liberalization on the economy, and the heterogeneous response of firms to trade openness. The strength of firm-level data is the rich information compared with country- or industry-level aggregated data. With firm-level data, the Melitz-type heterogeneous firm model gives plenty of economic insights such as the threshold productivity to export, and the shift of a threshold in response to trade liberalization. However, the heterogeneous firm model is too mathematically complicated to derive the model. In contrast, the gravity equation is relatively simple to derive the model, and it describes the trade flows well in empirical analysis.

The objective of this chapter is to first develop a new framework of a firm-level gravity equation with the presence of multi-product firms. Drawing on the theory of multi-product firms employed in chapter 2, I derive an equilibrium condition relating total demand and total production of multi-product firms. The equilibrium then allows me to examine export behavior of multi-product firms using a firm-level gravity equation. Alongside, I also estimate an aggregate gravity equation to compare results with those from the firm-level analysis.

For empirical analysis, I use data on Indian firms' export by exporter and destination from 1997 to 2002. I match this export data with Indian firms' balance sheet (production) information. Then, I combine these with physical distance, tariffs, and the average value of and the variance of the exchange rate to construct the firm-level gravity equation. Helpman, Melitz and Rubinstein (2008) emphasize the importance of controlling zero export flows. I estimate the firm-level gravity equation via Instrumental Variable Tobit (IV Tobit).

The remainder of this chapter is organized as follows. The next section models the firm-level gravity equation from CES preference and a Cobb-Douglas production function in the presence of multi-product firms. Section 3 provides the estimation strategy and specification issues in estimation. Section 4 describes the data. Section 5 provides results. Section 6 concludes.

3.2 Model

I assume that firms are heterogeneous in terms of productivity and that varieties (or products) are horizontally differentiated by attributes of products such as a brand name, age of a brand, quality and the way to utilize preference (or purpose of a good). Firms can produce multiple products, but they can supply only one product to a corresponding variety. Within a firm, products are identical in terms of productivity which is a firm' productivity, but they have different attributes.

3.2.1 DEMAND

Consumers in each country have identical preference. As mentioned above, all products are horizontally differentiated in terms of attributes. Consumers in country *f* at time *t* have a homogeneous utility function satisfying CES preference on a continuum of variety *v*,

(1) $U_t^f = \left(\int_{v \in V^f} q_t^f(v)^{\frac{\sigma+1}{\sigma}} dv\right)^{\frac{\sigma}{\sigma+1}}$

where V^f denotes the set of varieties available in country *f*; $q_t^f(v)$ denotes demand on variety *v* at time *t*; and σ denotes the elasticity of substitution. The elasticity of substitution is assumed to be identical across countries and $\sigma < -1$. The budget constraint in country *f* at time *t* is given by,

(2) $E_t^f = \int_{v \in V^f} p_t^f(v)\, q_t^f(v) dv$

where E_t^f denotes total expenditure in country *f* at time *t*; $p_t^f(v)$ denotes price of variety *v* at time *t*. Utility maximization gives the demand function on variety *v*,

(3) $q_t^f(v) = \left(\frac{p_t^f(v)}{P_t^f}\right)^{\sigma} \frac{E_t^f}{P_t^f}$,

where the aggregated price index in country *f* at time *t* is given as

$$P_t^f = \left\{\int_{v \in V^f} \left(p_t^f(v)\right)^{1+\sigma} dv\right\}^{\frac{1}{1+\sigma}}$$

3.2.2 Trade Resistance

I assume iceberg-type trade resistance between countries. The price of variety *v* in each country is equalized as the following,

(4) $p_t^f(v) = \tau_t^{hf} p_t^h(v)$

where $p_t^h(v)$ denotes the price of variety *v* produced in origin *h*; $p_t^f(v)$ denotes the price of the product in destination *f*; τ_t^{hf} denotes iceberg-type trade resistance from origin *h* to destination *f* at time *t*. Note that domestic trade resistance is $\tau_t^{hh} = 1$. In the presence of trade resistance ($\tau_t^{hf} \geq 1$), the aggregated price index in destination *f* is rewritten as,

$$(5)\ P_t^f = \left\{\int_{v \in V^f} \left(p_t^f(v)\right)^{1+\sigma} dv\right\}^{\frac{1}{1+\sigma}} = \left\{\int_{v \in V^f} \left(\tau_t^{hf} p_t^h(v)\right)^{1+\sigma} dv\right\}^{\frac{1}{1+\sigma}}$$

3.2.3 Firm-level Demand

In this section, I show the way to calculate a multi-product firm's total demand. Firstly, I introduce a product's demand from destination *f*. Then I calculate a multi-product firm's demand from destination *f* by aggregating a product's demand from destination *f*. Using the multi-product firm's demand from destination *f*, I calculate a multi-product firm's total demand from the all the destinations.

From equation (3), product *v* of firm *j* in country h faces export demand from country *f* given by,

$$(6)\ q_{jt}^{hf}(v) = \left(\frac{\tau_t^{hf} p_t^h(v)}{P_t^f}\right)^{\sigma} \frac{E_t^f}{P_t^f} = \left(\frac{\tau_t^{hf} p_t^h(v)}{P_t^f}\right)^{\sigma} Q_t^f \quad .$$

where Q_t^f denotes the aggregated market demand in destination *f* at time *t*, $Q_t^f = \frac{E_t^f}{P_t^f}$. I assume that trade resistance is identical regardless of type of varieties, so, $\tau_t^{hf}(v) = \tau_t^{hf}$ for $v \in V_{jt}^f$ where V_{jt}^f denotes the set of products produced by multi-product firm *j*. A multi-product firm's total demand from destination *f* is given as the following,

$$(7)\ q_{jt}^{hf} = \sum_{v \in V_{jt}^f} q_{jt}^{hf}(v) = \sum_{v \in V_{jt}^f} \left(\frac{\tau_t^{hf} p_t^h(v)}{P_t^f}\right)^{\sigma} Q_t^f = \left(\frac{\tau_t^{hf}}{P_t^f}\right)^{\sigma} Q_t^f \left(\sum_{v \in V_{jt}^f} \left(p_t^h(v)\right)^{\sigma}\right).$$

Equation (7) shows that the size of export demand from destination *f* on multi-product firm *j* is determined by trade resistance to destination *f*, aggregated price index in destination *f*, aggregated market demand in destination *f*, and the sum of products' prices to the power sigma at the origin (I define the sum of products' prices to the power sigma as multi-product firm's price index). However, the aggregation of a product's demand into a multi-product firm's demand does not require the product's symmetric export assumption meaning that product within a firm is exported by the same amount. In fact, the share of the individual products' export is determined by the ratio of prices at the origin which is the ratio of equation (6) and (7), $\frac{q_{jt}^{hf}(v)}{q_{jt}^{hf}} = \frac{\left(p_t^h(v)\right)^{\sigma}}{\sum_{v \in V_{jt}^f}\left(p_t^h(v)\right)^{\sigma}}$. Solving equation (7) for $\sum_{v \in V_{jt}^f}\left(p_t^h(v)\right)^{\sigma}$ gives,

(8) $\sum_{v \in V_{jt}^f}\left(p_t^h(v)\right)^{\sigma} = q_{jt}^{hf}\left(\frac{\tau_t^{hf}}{P_t^f}\right)^{-\sigma}\frac{1}{Q_t^f}$.

Here, I make an additional assumption that the set of exported products is identical across destinations, $V_{jt}^f = V_{jt}$ for $f = \{1, \ldots, F\}$. Then, the multi-product firm *j*'s total demand from all destinations is given as,

(9) $q_{jt}^h = \sum_{f=1}^F\left\{q_{jt}^{hf}\right\}$

$$= \sum_{f=1}^F\left\{\left(\frac{\tau_t^{hf}}{P_t^f}\right)^{\sigma} Q_t^f \sum_{v \in V_{jt}}\left(p_t^h(v)\right)^{\sigma}\right\}$$

$$= \left(\sum_{v \in V_{jt}}\left(p_t^h(v)\right)^{\sigma}\right)\sum_{f=1}^F\left\{\left(\frac{\tau_t^{hf}}{P_t^f}\right)^{\sigma} Q_t^f\right\} .$$

Equation (9) implies that a multi-product firm's total demand is the function of multi-product firm's price index, trade resistance, aggregated price index and aggregated market demand. In equation (8), I know that the multi-product firm's price index can be expressed as the function of variables in trade. By substituting equation (8) into equation (9), I can decompose the multi-product firm's total demand in equation (9) into bilateral and multilateral relationships as the following,

$$(9)'\ q_{jt}^{h} = \left(\sum_{v\in V_{jt}} \left(p_t^h(v)\right)^{\sigma}\right) \sum_{f=1}^{F} \left\{ \left(\frac{\tau_t^{hf}}{P_t^f}\right)^{\sigma} Q_t^f \right\}$$

$$= q_{jt}^{hf} \left(\frac{\tau_t^{hf}}{P_t^f}\right)^{-\sigma} \frac{1}{Q_t^f} \sum_{f=1}^{F} \left\{ \left(\frac{\tau_t^{hf}}{P_t^f}\right)^{\sigma} Q_t^f \right\}$$

$$= q_{jt}^{hf} \left(\frac{\tau_t^{hf}}{P_t^f}\right)^{-\sigma} \frac{1}{Q_t^f} \left(\Pi_t^h\right)^{\sigma} \qquad ;\ \left(\Pi_t^h\right)^{\sigma} = \left\{ \sum_{f=1}^{F} \left(\frac{\tau_t^{hf}}{P_t^f}\right)^{\sigma} Q_t^f \right\}$$

$$= q_{jt}^{hf} \frac{1}{Q_t^f} \left(\frac{\tau_t^{hf}}{P_t^f \Pi_t^h}\right)^{-\sigma} \qquad ;\ for\ \forall f\ .$$

For convenience, I replace the sum of multilateral resistance to the power sigma, $\sum_{f=1}^{F} \left(\frac{\tau_t^{hf}}{P_t^f}\right)^{\sigma} Q_t^f$, with $\left(\Pi_t^h\right)^{\sigma}$, which is referred to as outward multilateral resistance in the traditional country-level gravity equation. Note that P_t^f is referred to as inward multilateral trade resistance.

3.2.4 Firm-level Production

Each firm can produce multiple products, V_{jt}, and corresponding physical outputs is supplied as much as demanded from all the destinations in the world. As mentioned above, multi-product firms can produce multiple products, but they can supply only one product to each corresponding variety. Homogeneous technology and common input prices across products are assumed. As mentioned above, I also assume that firms are heterogeneous in terms of productivity, but products are differentiated in terms of attributes such as brand name, age of brand, quality and a purpose of a product. Within a multi-product firm, products are homogeneous in terms of productivity which is identical with a firm's productivity, but are heterogeneous in terms of attributes. These assumptions imply that a firm's productivity determines a product' the physical output.

Multi-product firm *j*'s has a standard Cobb-Douglas production function given by,

(10) $q_{ijt} = f\left(K_{ijt}, L_{ijt}, M_{ijt}\right) exp\left(\varphi_{jt}\right) = \left(K_{ijt}\right)^{\alpha_k}\left(L_{ijt}\right)^{\alpha_l}\left(M_{ijt}\right)^{\alpha_m} exp\left(\varphi_{jt}\right)$,

where q_{ijt} denotes product *i*'s physical output in multi-product firm *j* at time *t*; $f\left(K_{ijt}, L_{ijt}, M_{ijt}\right)$ denotes production from the homogeneous technology; K_{ijt}, L_{ijt}, and M_{ijt}, are product *i*'s input use of capital, labor and material, respectively; and φ_{jt} denotes firm *j*'s productivity of firm *j* and all products within a firm share the firm's productivity.

I adopt the concept of "a product's input share," which is introduced by Loecker (2011) to aggregate the product-level production function to firm-level production function. A product's input share is the ratio of product *i*'s input usage to firm *j*'s total input usage, where $a_{ijt} = \frac{K_{ijt}}{K_{jt}} = \frac{L_{ijt}}{L_{jt}} = \frac{M_{ijt}}{M_{jt}}$ and $\sum_{i \in V_{jt}} a_{ijt} = 1$. This input share assumption implies that there is no input synergy effect between products; this is referred to as input separability.

With the product's input share, product *i*' production function by firm *j* is rewritten as,

$$
\begin{aligned}
(11)\quad q_{ijt} &= f\left(K_{ijt}, L_{ijt}, M_{ijt}\right) exp\left(\varphi_{jt}\right) \\
&= \left(K_{ijt}\right)^{\alpha_k}\left(L_{ijt}\right)^{\alpha_l}\left(M_{ijt}\right)^{\alpha_m} exp\left(\varphi_{jt}\right) \\
&= \left(a_{ijt}K_{jt}\right)^{\alpha_k}\left(a_{ijt}L_{jt}\right)^{\alpha_l}\left(a_{ijt}M_{jt}\right)^{\alpha_m} exp\left(\varphi_{jt}\right) \\
&= a_{ijt}^{\gamma}\left(K_{jt}\right)^{\alpha_k}\left(L_{jt}\right)^{\alpha_l}\left(M_{jt}\right)^{\alpha_m} exp\left(\varphi_{jt}\right)
\end{aligned}
$$

where $\gamma = \alpha_k + \alpha_l + \alpha_m$ denotes the returns to scale of technology.

The multi-product firm' total production is given as the following,

(12) $q_{jt} = \sum_i q_{ijt}$

$$= \sum_i a_{ijt}^{\gamma} (K_{jt})^{\alpha_k} (L_{jt})^{\alpha_l} (M_{jt})^{\alpha_m} exp(\varphi_{jt})$$

$$= (K_{jt})^{\alpha_k} (L_{jt})^{\alpha_l} (M_{jt})^{\alpha_m} exp(\varphi_{jt}) \sum_i a_{ijt}^{\gamma} \;,$$

where the specialization of production, $\sum_i a_{ijt}^{\gamma}$, which reveals how a multi-product firms allocates its inputs across products and it also shows multi-product firms' concentration level in production (I define this term as specialization of production). With increasing or decreasing returns to scale technology, $\gamma \neq 1$, the specialization of production has the distinct feature that it has the largest value when one product within a firm is dominantly produced over all other products, and it has the lowest value when all products within a firm are evenly produced.[8] With increasing or decreasing returns to scale technology, the term of the specialization of production implies that multi-product firms lose its efficiency as they produce multiple products. However, with constant returns to scale or for a single product firm, ($\gamma = 1$ or $a_{ijt} = 1$), the specialization of production does not play any role in production function.

There are multiple features that the production function of multi-product firm should consider: economies of scale, economies of scope, and the effect of within firm resource reallocation (which I define this term as intra-firm resource reallocation). Economies of scale is captured by returns to scale in the production function, γ. Economies of scope is not considered in this chapter due to mathematical convenience and the input severability assumption in the production function. The

[8] It has been known that returns to scale is little bit greater than one, in general. See (Loecker, 2011;)

effect of within firm resource reallocation (or intra-firm resource reallocation) is captured by the specialization of production.

3.2.5 Equilibrium

At equilibrium (under the market clearing condition), a multi-product firm's total demand is equal with a multi-product firm's production. By combining the multi-product firms' total demand and supply, the equilibrium condition is given by,

$$(13)\quad q_{jt}^{hf}\,\frac{1}{Q_t^f}\left(\frac{\tau_t^{hf}}{P_t^f\Pi_t^h}\right)^{-\sigma} = \left(K_{jt}\right)^{\alpha_k}\left(L_{jt}\right)^{\alpha_l}\left(M_{jt}\right)^{\alpha_m} exp\left(\varphi_{jt}\right)\sum_i a_{ijt}^{\gamma}\ .$$

The equilibrium condition, equation (13), shows the relationship between trade and production variables, and how to rearrange it depends on economic theory and econometric efficiency. In international trade economics, there are two hypotheses explaining the relationship between trade liberalization and productivity: self-selection and learning-by-Exporting. The self-selection hypothesis implies that high productivity firms self select to export, and there is no feedback effect from export to productivity. The learning-by-Exporting hypothesis, however, implies that firms learn from exporting, and the empirical data shows that a firm's productivity is dramatically improved after attaining exporter status.

In this chapter, I consider the self-selection hypothesis only, and the learning-by-exporting hypothesis is tested in the following chapter. As mentioned above, the self-selection hypothesis points out that high productivity firms self-select to export and the size of export is determined by the magnitude of an exporting firm's productivity. The self-selection hypothesis leads us to solve equation (13) for export flows, q_{jt}^{hf}, and it gives the firm-level gravity equation,

$$(14)\ \ q_{jt}^{hf} = f(K_{jt}, L_{jt}, M_{jt}) exp(\varphi_{jt}^{h}) \left(\sum_i a_{ijt}^{\gamma}\right) Q_t^f \left(\frac{\tau_t^{hf}}{\Pi_t^h P_t^f}\right)^{\sigma} = q_{jt}^{h} Q_t^f \left(\frac{\tau_t^{hf}}{\Pi_t^h P_t^f}\right)^{\sigma}.$$

where $f(K_{jt}, L_{jt}, M_{jt}) exp(\varphi_{jt}^{h}) \left(\sum_i a_{ijt}^{\gamma}\right) = q_{jt}^{h}$ from equation (12).

The functional form of the firm-level gravity equation, equation (14), is almost identical to the country-level gravity equation where multi-product firm *j*'s export to destination *f* is determined by the economic size of trading partners, bilateral trade resistance and multilateral trade resistance. Multilateral trade resistance, Π_{jt}^{h} and P_t^f, capture outward and inward multilateral resistance, and these two terms reflect the fact that relative trade costs across countries matter in trade.

Equation (14) shows the relationship between multi-product firms' exports and other variables. Differentiating equation (14) in terms of technology gives the relationship between technology and export volume,

$$(15)\ \ \frac{\partial q_{jt}^{hf}}{\partial f(\cdot)} = exp(\varphi_{jt}^{h}) \left(\sum_i a_{ijt}^{\gamma}\right) Q_t^f \left(\frac{\tau_t^{hf}}{\Pi_t^h P_t^f}\right)^{\sigma} \quad > 0.$$

Equation (15) shows that technology improvement increases the export of firms.

$$(16)\ \ \frac{\partial q_{jt}^{hf}}{\partial exp(\varphi_{jt}^{h})} = f(\cdot) \left(\sum_i a_{ijt}^{\gamma}\right) Q_t^f \left(\frac{\tau_t^{hf}}{\Pi_t^h P_t^f}\right)^{\sigma} \quad > 0.$$

Equation (16) confirms the self-selection hypothesis by which productivity is the major source determining firms' export status and the volume of the export. It also shows that high productivity firms export more than low productivity firms.

With increasing returns to scale technology,

$$(17)\ \ \frac{\partial q_{jt}^{hf}}{\partial \sum_i a_{ijt}^{\gamma}} = exp(\varphi_{jt}^{h})\, Q_t^f \left(\frac{\tau_t^{hf}}{\Pi_t^h P_t^f}\right)^{\sigma} \quad > 0.$$

Equation (17) shows that specialization in a particular product increase the total export of multi-product firms.

$$(18)\quad \frac{\partial q_{jt}^{hf}}{\partial Q_t^f} = f\left(K_{jt}, L_{jt}, M_{jt}\right) exp\left(\varphi_{jt}^h\right)\left(\sum_i a_{ijt}^\gamma\right)\left(\frac{\tau_t^{hf}}{\Pi_t^h P_t^f}\right)^\sigma \quad > 0 .$$

Equation (18) shows that firms export more to richer destination.

$$(19)\quad \frac{\partial q_{jt}^{hf}}{\partial \tau_t^{hf}} = \sigma f\left(K_{jt}, L_{jt}, M_{jt}\right) exp\left(\varphi_{jt}^h\right)\left(\sum_i a_{ijt}^\gamma\right) Q_t^f \left(\tau_t^{hf}\right)^{\sigma-1}\left(\frac{1}{\Pi_t^h P_t^f}\right)^\sigma \quad < 0$$

where $\sigma < -1$. Equation (19) shows that an increase in bilateral trade resistance reduces exports to that destination. All effects from equation (15) to (19) are exactly identical as in the country-level gravity equation, but the effect in equation (17) is the new finding of the firm-level gravity equation. In the first chapter of this thesis, I find that multi-product firms in India gradually specialize their production and become a single product firm. However, the reason of specialization is not known. Goldberg, Khandelwal, Pavcnik and Topalova (2010) claims that trade liberalization does not have a impact on the number of products a firm produces. In this chapter, equation (12) shows that as a firm expands the number of products, inefficiency in production increases and the firm-level gravity equation, equation (17), shows that as multi-product firms specializes their production, they can export more. Note again that multi-product firms' specialization plays role under IRS and DRS technology.

The quantity (or volume) based firm-level gravity equation in equation (14) can be rewritten as the value based firm-level gravity equation. The firm-level gravity equation by value is more appropriate for the empirical data because firms report their export by value, in general, or there is the unit problem among products when aggregating quantity by firm even though export quantity is available.[9]

[9] For instance, in the empirical data in this chapter, counts, weight, volume, length, pairs, and sets are used as the unit (there are approximately 20 units in our data). However, export value has no aggregation problem as long as it uses the same currency.

By substituting equation (6) into $r_{jt}^{hf}(v) = p_{jt}^{hf}(v)q_{jt}^{hf}(v)$, the firm-level gravity equation by value is the following[10],

(20) $r_{jt}^{hf} = r_{jt}E_t^f\left(\frac{\tau_t^{hf}}{P_t^f \Pi_t^h}\right)^{1+\sigma}$

3.3 Data

3.3.1 Export and Production Data

This chapter uses Indian firms' export transactions, matched to corresponding firms' production dataset. As for export transactions, I use Tips Software Service's database (TIPS). Tips software service collects export transaction records from Indian Customs. It classifies export data by 8-digit HS code with exporter name, date, product description, export value, export volume, unit of volume, destination port and destination from April 1997 to March 2002 (6 years). I can identify the exporting firm with exporter name.[11] There is another study that uses Tips Software Service's database. Goldberg et al. (2010) use the same database, but their study is about the import side. Tips Software Service's database is not a comprehensive database. It covers 8-12 major seaports and air ports in India accounting for 50-60% of total exports in India.

As for a firm's total revenue, I use the PROWESS database (PROWESS) on Indian firms over the 1997-2002 periods. By the Companies Act 1956, all Indian companies are required to report their production information on capacity, unit-price, output, input use, the number of products and revenue in their annual report. The

[10] See Appendix I.

[11] Exporter name is not available after November 2004 by Indian Customs Notification No.128/2004. Some exporter names are still available until March 2005 even though it is prohibited by the law.

Centre for Monitoring the Indian Economy has collected firms' annual report information. PROWESS is not a comprehensive database as well and usually covers mid- and large-size companies accounting 60-70% of overall firms in India. Work by Goldberg et al. (2009, 2010a, 2010b), Topalova and Khandelwal (2011) and Loecker et al. (2012) use the same database.

The match of production and export databases allow for the constructing production and export relationship. However, there is an issue to match two databases. Two databases are not perfectly matched because both databases are incomprehensive, and the set of firms belongs to both databases is part of the databases (for which the match rates are approximately 23.1% for TIPS and 27.0% for PROWESS)[12]. In addition, about 80% of firms in PROWESS follow the fiscal accounting year from April of one calendar year to March of the following year, while 15% of firms follow the calendar accounting year from January to December of the same accounting year.[13] So I firstly convert the firm following the fiscal accounting year to the calendar accounting year, and then merge them with the firms following the calendar accounting year. The conversion from the fiscal to the calendar accounting year is implemented as in the following,

$$(21)\quad \textit{Calendar accounting year } t = \frac{9}{12}\,\textit{Fiscal accoutnting year } t + \frac{3}{12}\,\textit{Fiscal accoutnting year } t+1\ .$$

Two databases use the different classification code. TIPS uses 8-digit HS code to classify export transactions and PROWESS uses 20-digit PROWESS code to define

[12] This means that 76.3% of firms in TIPS do not have corresponding production information in PROWESS and 73% of firms in PROWESS are either non-exporting firms, or exporting firms without export transaction records in TIPS.

[13] 5% of firms are switching accounting years from the fiscal to the calendar or from the calendar to the fiscal.

sector a firm belongs to. I convert both codes to the Standard International Trade Classification (SITC) to identify the sector a firm belongs to. Note that SITC defines sector by 1-digit SITC code and contains 9 sectors.

Table 3.1 shows the evolution of the Indian export flows from 1997 to 2002 used in this analysis. Two features stand out: first, export flows significantly decrease during 1998-1999 due to economic crisis in East Asia; second, both the average number of destinations by firm and the average export value by firm gradually increase since 1999. As is evident from the table, I conclude that the growth of export is driven not only by the increase in the average export value, but also mainly by the average number of destinations by firm.

Table 3.1 Exporting Firms, Export Value and Destinations

	Year					
	1997	1998	1999	2000	2001	2002
The Total Export Value	1.02	0.88	0.48	1.09	1.46	1.35
The Total Number of Destinations	76	82	82	92	91	98
The Total Number of Exporting Firms	328	342	306	395	386	388
The Average Number of Destinations by Firm	4.71	5.19	3.96	4.61	4.79	4.99
The Average Export Value by Firm	3.12	3.58	1.56	2.75	3.79	3.48
Observations	1,719	1,941	1,882	2,084	2,140	2,181

Note. Export data used in this table is incomprehensive data. If export transactions do not have corresponding production information, then they are dropped. The unit of the Total Export Value is Billion Indian Rupee and of the Average Export is Million Indian Rupee.

3.3.2 Trade Resistance Data

A destination's tariff data comes from the World Trade Organization's Tariff Download Facility. It provides the tariff rate through a 2-digit HS code (chapter-wide average tariff rates) of most countries in the world (approximately 130 countries) from 1996 to the present. Each tariff rates of a destination is against all origin countries in the world.

Table 3.2 shows sector-wide tariff rates from 1997 to 2002. The tariff rates in all sectors gradually have decreased. Agriculture related sectors (Food and Live animals, Beverage and Tobacco and Animal and Vegetable Oils) have relatively higher tariff rates than other sectors.

Table 3.2 Sector-wide Tariff

Sector	**Year**					
	1997	1998	1999	2000	2001	2002
Food and Live Animals (s=0)	0.165	0.152	0.158	0.137	0.134	0.138
Beverage and Tobacco (s=1)	0.434	0.398	0.390	0.351	0.343	0.338
Crude Material (s=2)	0.075	0.079	0.073	0.070	0.059	0.062
Mineral Fuels (s=3)	0.062	0.065	0.058	0.052	0.047	0.048
Animal and Vegetable Oils (s=4)	0.136	0.130	0.126	0.117	0.107	0.104
Chemicals and Related Products (s=5)	0.095	0.091	0.081	0.074	0.067	0.066
Manufactured Goods (s=6)	0.110	0.113	0.104	0.093	0.089	0.088
Machinery and Transport Equipment (s=7)	0.101	0.096	0.087	0.081	0.075	0.069
Miscellaneous Manufactured Articles (s=8)	0.147	0.130	0.132	0.122	0.118	0.117
The Number of Countries included	76	82	82	92	91	98

Note. The tariff rate is average tariff rate against countries having export relationship with India.

As in Table 3.2, I observe that tariff rates significantly vary by sector, and construct a sector-wide tariff rates by taking the weighted average of the chapter-wide tariff rates:

$$(22)\quad Tariff_s^f = \sum_{c\in S(c)} Tariff_c^f \times \frac{import_c^f}{\sum_{c\in S(c)} import_c^f} \quad ,$$

where $Tariff_s^f$ denotes the sector-wide tariff rate of destination *f* by 2-digit SITC code; $Tariff_c^f$ denotes the chapter-wide tariff rate of destination *f* by 2-digit HS code; $import_c^f$ denotes total import value of destination *f* in chapter c; and $S(c)$ denotes the set of sector.[14]

I use a destination' GDP per capita from the World Bank's World Development Indicator.[15] For distance, a destination' population, border share, colonial link and common language, I use the CEPII's database on country specific distance which is developed by Mayer and Zignago (2006). To measure consuming power of a representative consumer in a destination, I use the destination's GDP per capita, and to measure the destination's overall market size, I use the destination's population.

For exchange rate, comprehensive data is not available. I make three groups depending on the availability of the exchange rates. The exchange rate comes from the Border of Governors of the Federal Reserve System.[16] Exchange Rate I includes 35 countries of which the bilateral exchange rate are available (unit: destination's local currency/Indian Rupee (DLC/INR)). Exchange Rate II shows the average and the variance of USD. Exchange rate information of other 65 countries in the data are not available, and I assume that those countries use USD in export transactions. Goldberg and Tille (2008) state that that exporting firms in developing countries can invoice the export transaction in USD as an invoicing currency. Exchange Rate III includes both Exchange Rate I and II. In empirical analysis, I mainly use exchange rate III.

---

[14] There are 96 chapters in HS code, and 10 sectors in SITC code.
[15] GDP per capita is measured by purchasing power parity (constant international dollars).
[16] I use the nominal exchange rate

Table 3.3 summarizes the movements in the GDP per capita and the exchange rate. In 1999, the exchange rate in Exchange Rate I and III dramatically drops because 10 countries in Europe started to use EURO as their major local currency.

Table 3.3 The Movements in GDP and Exchange Rate

	Year					
	1997	1998	1999	2000	2001	2002
(1) The Average GDP per Capita	10,158.39	8,872.11	9,734.20	8,804.80	9,133.46	9,195.19
(2) Exchange Rate I						
Average	352.35	367.83	176.54	127.42	110.35	105.17
Variance	5,723.36	9,176.20	257.36	171.78	82.44	369.43
Countries	35	35	35	35	35	35
(3) Exchange Rate II						
Average	2.75	2.42	2.32	2.22	2.12	2.23
Variance	0.006	0.008	0.000	0.003	0.006	0.001
Countries	65	65	65	65	65	65
(4) Exchange Rate III						
Average	149.95	140.56	70.30	45.77	40.18	36.19
Variance	2,409.84	3,469.06	100.43	59.75	28.99	121.88
Countries	100	100	100	100	100	100

Note. (1) The unit of the Average GDP per Capita is constant price US Dollars. (2) The exchange rate does not come from comprehensive database, and some country's exchange rate is not available. I make three groups depending on the availability of the exchange rate. Exchange Rate I include the bilateral exchange rate of the countries reporting their exchange rate (35 countries).(3) Exchange Rate II includes the countries not reporting the exchange rate for which USD is used as the bilateral exchange rate of those countries (65 countries). (4) Exchange Rate III includes the both countries using their own bilateral exchange rate and USD as the vehicle exchange rate (100 countries). The value in parenthesis denotes the number of countries using the corresponding exchange rate. The unit of the Exchange Rate I is (local currency/INR*100), of the Exchange Rate II is (USD/local currency * 100), and of Exchange Rate III is the both of Exchange Rate I and II. In the empirical analysis, I use the Exchange Rate III. In 1999, 10 countries in Europe started to use EURO as their major local currency. As a result, the exchange rate in 1999 dramatically drops in 1999.

3.4 Estimation Strategies

Our theoretical framework provides a justification for firm-level export flows. The unit of empirical analysis is firm's export by sector and by destination. By taking the log on the right hand side of equation (20), the firm-level gravity equation is

(23) $r_{sjt}^{hf} = \log r_{jt}^{h} + log\, E_t^f + (1+\sigma_s)\, log\, \tau_{st}^{hf} + \sigma\, log\, \Pi_t^h + (1+\sigma)\, log\, P_t^f$

where firm-level export value, r_{jt}^{hf}, is approximated by firm-sector-level export value, r_{sjt}^{hf}[17]. Total expenditure, E_t^f, is approximated by GDP per capita of destination *f*. Trade resistance, τ_t^{hf}, include multiple factors: sector-wide tariffs, distance, and the average and the variance of exchange rate.

The main estimable gravity equation is given by

(24) $r_{sjt}^{hf} = \alpha_r\, log\, r_{jt}^{h} + \alpha_E\, log\, E_t^f + \beta_\tau\, log\, \tau_{st}^{hf} + \delta_f d_f + \delta_s d_s + \delta_t d_t + \varepsilon_{sjt}^{h\,f}$

where α denotes the coefficients of the firm-level gravity equation as opposed to the true parameters having unity as their value; β_τ denotes the estimated coefficients of trade resistance, $\delta_\tau = \frac{\beta_\tau}{|1+\sigma_s|}$; $\varepsilon_t^{h\,f}$ denotes stochastic components and is *i.i.d.* across destination and time; outward multilateral resistance (Π_t^h) is not considered since I implement a one-to-multiple-country model; inward multilateral resistance (P_t^f) is captured by fixed effects for destination (d_f); fixed effects for sector and time (d_f, d_t) are included; and stochastic components $(\varepsilon_{sjt}^{h\,f})$ are *i.i.d.* across firms, sectors, destination and time.

There are two estimation issues in this equation: endogeneity and zero export flows. The correlation between export value and a firm' total revenue occurs because a firm's total revenue increases when the firm's export value increases, creating correlation issues between export revenue and a firm's aggregated revenue. In order to address the correlation, I use capital as the instrumental variable for a firm's total

---

[17] Firm-sector-level export value, and multi-product firm's total revenue is deflated by the Indian Whole Sale Price Index.

revenue. I assume that capital is pre-determined at the beginning of the year and that it is not adjustable in the short run due to high adjustment cost of the capital. However, as a firm's capital increases, firms' revenue increases.[18] The second issue is the zero export flows. Literature analyzing export flows using the traditional gravity equation encounters zero export flows problem because the gravity equation does not consider the zero export flows, but there are lots of zero export flows in the empirical data. Ignoring zero export flows could cause biased results in the estimation, and it is not consistent with the empirical data (Helpman, Melitz and Rubinstein, 2008). Alternatively, Tobit is used to handle the zero export flows problem. However, taking the log of zero export value is not defended. I only take the log of the right hand side of the equation (20), and the main estimable equation (24) becomes the level-log equation, instead of the log-log equation.

Finally, I adopt Instrumental Variable Tobit (IV Tobit) to handle the endogeneity and the zero export flows problems.

$$(25)\quad {r_{sjt}^{hf}}^{*} = \alpha_r \, log\, r_{jt}^{h} + \alpha_E \, log\, E_t^f + \alpha_W log W_t + \delta_\tau (1+\sigma_s)\, log\, \tau_{st}^{hf} + \delta_f d_f + \delta_t d_t + \varepsilon_t^{hf}$$

$$\text{where } {r_{sjt}^{hf}}^{*} = \begin{cases} {r_{sjt}^{hf}}^{*} & if \quad r_{sjt}^{hf} > 0 \\ 0 & if \quad r_{sjt}^{hf} \le 0 \end{cases} .$$

The coefficients of the firm level gravity equation are denoted by α as opposed to the true parameters having unity as their value. The null hypothesis, $\alpha_q = 1$ and $\alpha_g = 1$. As for trade resistance, I can estimate $\delta_\tau(1+\sigma_s)\, log\, \tau_t^{hf}$. I then recover the

[18] In empirical analysis, I use net fixed assets for capital and capital is deflated by a corresponding sector's the Indian Whole Sale Price Index. As for multi-sector firms, I use weighted average of sectors' Indian Whole Price Index.

true parameters, δ_τ, by making an assumption about the elasticity of substitution $(1+\sigma_s)$ as $\sigma_s = -7$.

3.5 Estimation Results

I begin by estimating the firm-level gravity equation estimated via Tobit. Then I compare the estimation results from the firm-level gravity equation with those from the country-level gravity equation. The country- level gravity equations are estimated by GLS.

Table 3.4 provides the estimates of both the country- and firm-level gravity equations using the data on Indian firms during 1997-2002 from a set of 714 Indian firms exporting to 106 destinations. Column (1) in Table 3.4 shows the estimation results of the firm-level gravity equation via IV Tobit, Column (2) shows the marginal effects of estimated coefficients, and Column (3) provides the implied coefficients for the country-level gravity equation. The estimated coefficient, β_τ, on trade resistance are the product of the original coefficient and the elasticity of substitution, $(|1+\sigma_s|)$. Note that the firm-level gravity equation is the level-log equation.

In equation (20), I showed that export flows is proportional to the product of economic size of exporting firms, consuming power of a representative consumers and trade resistance between them and this is the distinct feature of the gravity equation. Estimation results in Table 3.4 are consistent with the feature of the gravity equation. In column (1), the economic size of exporting firms has a significant and a positive effect on export value.

Table 3.4 Estimates of the Country- and the Firm-level Gravity Equation

Coefficients On	Firm-level Gravity		
	IV Tobit (1)	Marginal Effects (2)	Implied Coefficients $\left(\delta_\tau = \frac{\beta_\tau}{\lvert 1+\sigma_s \rvert}\right)$ (3)
Total Revenue of a Firm	545.86*** (11.82)	40.12*** (0.87)	-
GDP per Capita of a Destination	110.83 (80.01)	8.15 (5.93)	-
Population of a destination	25.25*** (6.10)	1.86*** (0.45)	-
Tariffs	-2,080.14*** (357.82)	-153.12*** (26.43)	-25.12***
Distance	-9,639.60*** (2,616.56)	-708.63*** (192.61)	-118.11***
Average of Exchange Rate	-53.50 (40.12)	-3.92 (2.98)	-
Variance of Exchange Rate	9.71 (19.22)	0.71 (1.42)	-
Shared Border	-30,131.54*** (8,619.02)	-2,205.63*** (634.96)	-367.61***
Colonial Link	-3,984.78 (2,644.06)	-293.12 (194.64)	-48.85
Common Language	10,294.55*** (2,867.11)	753.56*** (210.79)	125.59***
Constant	-2,034.15 (1432.15)	149.51 (104.90)	-
Fixed Effect for Destination, Sector and Year	Included	-	-
No. of obs	1,348,391	-	-

Notes. All variables are logged. Unit of export value and total revenue of a firm is 10,000 INR. Exchange Rate III is used for the analysis. Standard Errors are in parentheses. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

High level of total revenue of a firm implies that the firm has higher productivity. The positive sign on instrumented total revenue of a firm is consistent with the self-selection hypothesis in the heterogeneous firm model such as Melitz (2003) and Bernard, Eaton, Jensen and Kortum (2003). Note that the positive coefficient in column (3) interpreted as that one percent increase in the independent variable increases the export value by coefficients. For example, one percent increase in total revenue of a firm increases the export value by 40.12 (unit: 10,000 INR).

For trade resistance, I include tariff rate, distance, shared border, colonial link and common language. I do not regard the average and the variance of exchange rate as trade resistance because exchange rate is multiplicative to export flows, $\mathrm{r}^{\mathrm{hf}}_{\mathrm{jt\_INR}} = \mathrm{e}^{\mathrm{hf}}_{\mathrm{t}}\mathrm{r}^{\mathrm{hf}}_{\mathrm{jt\_local}}$ where $\mathrm{r}^{\mathrm{hf}}_{\mathrm{jt\_INR}}$ denotes export revenue by INR, and $\mathrm{r}^{\mathrm{hf}}_{\mathrm{jt\_local}}$ denotes export revenue by local currency. In addition, I assume that the elasticity of substitution is identical across sectors, $\sigma = -7$, and divide the estimated coefficients in column (2) by 6 to calculate the implied coefficients in column (3). If this assumption is relaxed, then the effect of tariff and distance becomes bigger as the elasticity of substitution becomes inelastic (smaller). Both tariff and distance have negative and significant effect on export value, but the effect of distance is much larger than that of tariff. This is consistent with the work by Berthelon and Freund (2008), Bleaney and Neaves, (2013), Brun, Carrére, Guillaumont and Melo (2005) and Chaney (2013).

Shared border has a negative and significant sign, but the data of this chapter is not enough for the analysis of shared border effect because the data does not include the export transactions through roads in land. Moreover, political conflict limits India’s trade with its close neighbors. Colonial link has a negative but insignificant effect. Common language has a positive and significant effect.

The exchange rate appears to not have an effect on export flows. The average of exchange rate has negative but statistically insignificant effect on export value (the unit of exchange rate: Local Currency/ INR). There are several explanations for this insignificant effect of exchange rate. To encourage export and reduce exchange rate risk, some countries adopt fixed exchange rate system for export firms. For example, the Indian Customs announces the fixed exchange rate and keeps it constant for two or three weeks. The Korean Customs also announce the fixed exchange rate every two weeks as well. In these countries, export firm can change local currency into foreign current at fixed exchange rate. In addition, exporting firms make a contract in a financial market to avoid exchange rate risk. In this case, firms export revenue remains stable regardless of changes in exchange rate.

In Table 3.5, I compare the estimates from the firm-level gravity equation with those from the country-level gravity equation. Note that country-level export data for the country-level gravity equation is constructed by aggregating firm-level export data. The coefficients are not directly comparable because the firm-level gravity equation has the level-log form due to the zero export flows problem, while the country-level gravity equation has the log-log form. However, if I convert the coefficients of the firm-level gravity equation from the level changes to the percentage changes using the mean value of the independent variable, I find that estimated coefficients of the country-level gravity equation tend to be larger than those from the firm-level gravity equation because variations in the dependent variable of the aggregated data (country-level data) is larger than those in the disaggregated data (firm-level data), even thought the independent variables are identical for the aggregated and the disaggregated data.

The firm-level gravity equation yields the identical results as in the country-level gravity equation in terms of the sign of estimated coefficients and the significance of those. The statistical significance of estimated coefficients in Column (1) is identical with those in Column (2). This implies that both gravity equations give identical information, but the firm-level gravity equation gives more because it reveals the effects of the economic size of exporting firms and a representative consumer. The country-level gravity equation only shows the effect of the product of both countries' GDP per capita when year fixed effects are accounted in the empirical estimation, while firm-level gravity equation provide the effects of the economic size of a firm and a representative consumers.

Table 3.5 Recovering the Original Parameters of the Firm-level Gravity Equation

Coefficients on	Firm-level Gravity Equation	Country-level Gravity Equation
GDP of India X GDP of a destination		0.31** (0.14)
Total Revenue of a Firm	40.12*** (0.87)	-
GDP per capita of a Destination	8.15 (5.93)	-
Population of a destination	1.86*** (0.45)	0.01 (0.02)
Tariffs	-25.12*** (4.41)	-1.97*** (0.68)
Distance	-118. 11 (32.10)	-184.13* (98.81)
Average of Exchange Rate	-3.92 (2.98)	-0.18** (0.09)
Variance of Exchange Rate	0.71 (1.42)	0.05 (0.04)
Shared Border	-367.61 (105.83)	-430.63 (257.98)
Colonial Link	-48.85 (32.44)	-8.47 (7.10)
Common Language	125.59 (35.13)	59.41 (33.84)
Constant	149.51 (104.90)	1575.73 (941.05)
Fixed Effect for Destination, Sector and Year	Included	Included
No. of obs	1,348,391	2,700

Notes. This table shows the original coefficients of trade resistance. The estimated coefficients of trade resistance is the product of original coefficients and the elasticity of substitution, $(|1+\sigma_s|)$. For other variables, the estimated coefficients are identical with the original coefficients.

3.6 Conclusion

Analyzing the impact of trade liberalization on export flows has a long tradition, but most of the literature has concentrated on country-level trade flows using the country-level gravity equation. In this chapter, I develop a firm-level gravity equation which is the expansion of the country-level gravity equation. On the theoretical side, I newly find that specialization of production has an impact on export flows with increasing or decreasing returns to scale. On the empirical side, I show that the firm-level gravity equation has the same characteristics of the country-level trade flows, but it reveals the role of productivity, specialization of production and a firm's size.

The firm-level gravity equation has large potential for applications because the firm-level gravity equation enables the ability to analyze the behavior of exporting firms. In the next chapter of Ph.D. thesis, I provide the example of the application of the firm-level gravity equation.

3.7 Appendix

3.7.1 Firm-level Demand by Value

Equation (6) is given as the following,

$$\text{(A.1) } q_{jt}^{hf}(v) = \left(\frac{\tau_t^{hf} p_t^h(v)}{P_t^f}\right)^{\sigma} \frac{E_t^f}{P_t^f} = \left(\frac{\tau_t^{hf} p_t^h(v)}{P_t^f}\right)^{\sigma} Q_t^f$$

where $\sigma < -1$.

Export demand by value is given as,

$$\text{(A.2) } r_{jt}^{hf}(v) = p_{jt}^{hf}(v) q_{jt}^{hf}(v)$$

$$= p_{jt}^{hf}(v) \left(\frac{\tau_t^{hf} p_t^h(v)}{P_t^f}\right)^{\sigma} \frac{E_t^f}{P_t^f}$$

$$= \tau_t^{hf} p_t^h(v) \left(\frac{\tau_t^{hf} p_t^h(v)}{P_t^f}\right)^{\sigma} \frac{E_t^f}{P_t^f}$$

$$= \left(\frac{\tau_t^{hf} p_t^h(v)}{P_t^f}\right)^{1+\sigma} E_t^f$$

Firm-level export demand by value is given as,

$$\text{(A.3) } r_{jt}^{hf} = \sum_{v \in V_{jt}} r_{jt}^{hf}(v)$$

$$= \sum_{v \in V_{jt}} \left(\frac{\tau_t^{hf} p_t^h(v)}{P_t^f}\right)^{1+\sigma} E_t^f$$

$$= \left(\frac{\tau_t^{hf}}{P_t^f}\right)^{1+\sigma} E_t^f \left(\sum_{v \in V_{jt}} p_t^h(v)^{1+\sigma}\right)$$

where I assume that $V_{jt}^f = V_{jt}$ for $f = \{1, \dots, F\}$.

Solve equation (A.3) for $\sum_{v \in V_{jt}} p_t^h(v)^{1+\sigma}$,

$$\text{(A.4) } \sum_{v \in V_{jt}} p_t^h(v)^{1+\sigma} = \frac{r_{jt}^{hf}}{E_t^f} \left(\frac{\tau_t^{hf}}{P_t^f}\right)^{-(1+\sigma)}$$

A multi-product firm's total demand by value is given as,

(A.5) $r_{jt}^{h} = \sum_{f=1}^{F} r_{jt}^{hf}$

$$= \sum_{f=1}^{F} \left(\frac{\tau_t^{hf}}{P_t^f}\right)^{1+\sigma} E_t^f \left(\sum_{v\in V_{jt}} p_t^h(v)^{1+\sigma}\right)$$

$$= \left(\sum_{v\in V_{jt}} p_t^h(v)^{1+\sigma}\right)\left(\sum_{f=1}^{F} \left(\frac{\tau_t^{hf}}{P_t^f}\right)^{1+\sigma} E_t^f\right)$$

$$= \frac{r_{jt}^{hf}}{E_t^f}\left(\frac{\tau_t^{hf}}{P_t^f}\right)^{-(1+\sigma)} \left(\sum_{f=1}^{F} \left(\frac{\tau_t^{hf}}{P_t^f}\right)^{1+\sigma} E_t^f\right)$$

$$= \frac{r_{jt}^{hf}}{E_t^f}\left(\frac{\tau_t^{hf} p_t^h(v)}{P_t^f}\right)^{-(1+\sigma)} \left(\Pi_t^h\right)^{1+\sigma}$$

$$= \frac{r_{jt}^{hf}}{E_t^f}\left(\frac{\tau_t^{hf}}{P_t^f \Pi_t^h}\right)^{-(1+\sigma)}$$

where $\sum_{v\in V_{jt}} p_t^h(v)^{1-\sigma} = \frac{r_{jt}^{hf}}{E_t^f}\left(\frac{\tau_t^{hf}}{P_t^f}\right)^{-(1+\sigma)}$ as in (A.4); and $\left(\Pi_t^h\right)^{1-\sigma} =$

$\sum_{f=1}^{F} \left(\frac{\tau_t^{hf}}{P_t^f}\right)^{1-\sigma} E_t^f$ for W_t is total world output.

3.7.2 Firm-level Production by Value

Firm-level total output is given by,

(A.6) $q_{ijt} = a_{ijt}^{\gamma}\left(K_{jt}\right)^{\alpha_k}\left(L_{jt}\right)^{\alpha_l}\left(M_{jt}\right)^{\alpha_m} exp\left(\varphi_{jt}\right)$

A product's output by value is given by,

(A.7) $r_{ijt} = p_{ijt}q_{ijt} = p_{ijt}a_{ijt}^{\gamma}\left(K_{jt}\right)^{\alpha_k}\left(L_{jt}\right)^{\alpha_l}\left(M_{jt}\right)^{\alpha_m} exp\left(\varphi_{jt}\right)$

A firm's total output by value is given by,

(A.8) $r_{jt} = \sum_i r_{ijt} = \sum_i p_{ijt}a_{ijt}^{\gamma}\left(K_{jt}\right)^{\alpha_k}\left(L_{jt}\right)^{\alpha_l}\left(M_{jt}\right)^{\alpha_m} exp\left(\varphi_{jt}\right)$

3.7.3 Equilibrium by Value

At equilibrium, firm-level total demand is equal to firm-level total production. By equating equation (A.5) with (A.8), the equilibrium condition is given by,

$$\text{(A.9)}\ \frac{r_{jt}^{hf}}{E_t^f}\left(\frac{\tau_t^{hf}}{P_t^f \Pi_t^h}\right)^{-(1+\sigma)} = \sum_i p_{ijt} a_{ijt}^{\gamma} (K_{jt})^{\alpha_k} (L_{jt})^{\alpha_l} (M_{jt})^{\alpha_m} exp(\varphi_{jt}) = r_{jt} \quad .$$

By solving equation (A.9) for r_{jt}^{hf}, I have the firm-level gravity equation by value,

$$\text{(A.10)}\ r_{jt}^{hf} = r_{jt} E_t^f \left(\frac{\tau_t^{hf}}{P_t^f \Pi_t^h}\right)^{1+\sigma} \quad .$$

4 Multi-Product Firms and Learning-by-Exporting

4.1 Introduction

Why exporting firms are more productive than non-exporting firms? Does high productivity of exporting firms comes from learning through exporting or do high productivity firms self-select into export markets? The previous chapter addressed self-selection, while the current one explores learning-by-exporting.

As noted earlier, the learning-by-exporting hypothesis refers to productivity improvement through export participation. This hypothesis explains the mechanism of productivity improvement in various ways, but most of them are referred to as knowledge diffusion from foreign countries. To attract international buyers and to win intense competition in foreign countries, exporters improve product quality and export processes. As a result, exporting firms' productivity can be improved (Wagner, 2007; and Loecker, 2013). Export relationship becomes a channel of productivity improvement because of access to better technology, improved management skills, access to more skilled workers who can manage foreign networks, product innovation, international patenting and R&D efforts (Redding, 2010). Exporting firms tend to invest more than their non-exporting counterparts to increase productivity for higher profit and to compete in the foreign market (Melitz and Costantini, 2008; and Desmet and Parente, 2010). Aw, Roberts and Xu (2008, 2011) find positive feedback effects between R&D and export participation because R&D investment increases both productivity and export profits.

Despite the debate between self-selection and learning-by-exporting, evidence favors the former over the latter. Keller (2004) summarizes case studies supporting the learning-by-exporting hypothesis, but still there is a skeptical view based on econometric studies.

The primary objective of this chapter is to examine the learning-by-exporting hypothesis. I extend on the framework in the previous chapter to derive a mechanism of productivity improvements at firm-level, and test the learning-by-exporting hypothesis. It allows me to decompose the effects of export volume, tariff, distance and exchange rate on productivity.

For empirical analysis, I combine two databases of Indian firms' production and export from 1997 to 2002 for 6 years. The first database includes Indian firm's production and sales records at firm- and product-level, respectively. The second database contains the product-level export records with firm identification. Combining these two databases allow us to track multi-product firm's production and export history. The effect of learning-by-exporting would be bigger as productivity gap between the origin country and the rest of the world is large. India is one of the largest developing countries, and trade liberalization has been slower than that of other developing countries (Haidar, 2012). Thus, India is the case to analyze the role of export experience in productivity.

The reminder of this chapter is organized as follows. The modeling framework, as noted earlier, relies on the equilibrium condition from the previous chapter. Data are also taken from the earlier chapter. Section 4.4 gives the main estimation results. Section 4.5 concludes with a discussion of the effects of export experience on productivity. Appendix gives in-depth information on the derivation of the value based market clearing condition.

4.2 Data

The data of this chapter are exactly same with those in the third chapter of this thesis (3.3 Data), but the following provides a brief summary.

I estimate TFP using the PROWESS database, and then match the PROWESS with the TIPS to analyze the impact of export experience on productivity. However, the match rate between PROWESS and TIPS is not 100% because both PROWESS and TIPS are not comprehensive database. The set of firms in each database are not perfectly identical. The PROWESS covers mid- and large-size companies accounting 60-70% of overall firms in India, and TIPS covers top 9 to 11 sea and air ports in India accounting 50-60% of export transactions in India. In addition, there is an accounting year difference between PROWESS and other variables such as tariff rate and GDP. PROWESS uses fiscal accounting year from April of a current year to March of the following year, while all other variables use calendar accounting year from January to December of the same year. Basically, I estimate TFP based on the fiscal accounting year, then convert them to calendar accounting year TFP with the following equation,

(26) $\mathrm{TFP_t} = \frac{9}{12}\mathrm{TFP_t} + \frac{3}{12}\mathrm{TFP_{t+1}}$

4.3Estimation Strategies

By taking log on equation (14) in Chapter 3, the estimable equation for testing learning-by-exporting is given as,

$$(27)\quad \varphi_{jt}^{h} = \log r_{jt-1}^{hf} - log\, Q_{t-1}^{f} - \sigma\, log\, \tau_{t-1}^{hf} + \sigma\, log\, P_{t-1}^{f} - log g_{t-1}(\cdot) - \log \textstyle\sum_i a_{ijt-1}^{\gamma}$$

Note that I do not consider the outward multi-lateral resistance, Π_{t-1}^{h}, because I analyze one-origin-and-multiple-destinations relationship. In the empirical analysis, productivity, φ_{jt}^{h}, is measured by TFP following the procedure shown in Chapter 2. Firms' lagged export volume, q_{jt-1}^{hf}, is not directly observable and is proxied by lagged deflated export value (Free On Board or F.O.B), r_{jt-1}^{hf}. A destination's lagged aggregated market demand $\left(Q_{t-1}^{f}\right)$ is measured by a destination's lagged GDP per capita. Bilateral trade resistance $\left(\tau_{t}^{hf}\right)$ contains a destination's previous tariff rates, distance, and previous average and variance of exchange rates. Inward multi-lateral resistance $\left(P_{t-1}^{f}\right)$ is captured by dummies for destinations, and fixed effects for time are considered as well[19].

The production from the homogeneous technology, $g\left(K_{jt-1}, L_{jt-1}, M_{jt-1}\right)$, is calculated after estimating the revenue based production function. Firstly, the Cobb-Douglas production function is estimated. Secondly, TFP is calculated as residuals after generating the production function. Finally, the production from the

[19] Lagged export values are deflated by sector-wide Indian Whole Sale Price Index which is Producer Price Index in India. I use real value for a destination's lagged real GDP per capita, but nominal exchange rate for lagged exchange rate. Note that there are 10 sectors: Foods and Live Animals, Beverage and Tobacco, Crude Material, Mineral Fuels, Animal and Vegetable Oils, Chemicals and Related Products, Manufactured Goods, Machinery and Transport Equipment, Miscellaneous Manufactured Articles, and Other Transactions and Goods.

homogeneous technology is calculated by subtracting TFP from the Cobb-Douglas production function (or total physical output).

The specialization of production is approximated by the Herfindhal index of products' revenue within a firm, $h_{jt} = \sum_i \left(\frac{\tilde{r}_{ijt}}{\sum_i \tilde{r}_{ijt}}\right)^2$ where $\tilde{r}_{ijt}$ denotes product i's deflated revenue in firm j at time t.[20] A product's is deflated by sector-wide producer price index. The feature of the Herfindhal index is similar with the specialization of production that it converges to unity as a particular product is dominantly produced over other products within a firm, and it converges to zero as all products are evenly produced. Note that I use a product' deflated revenue to construct Herfindhal index, instead of output.

Finally, the main estimable equation is given by,

$$(28)\quad \varphi_{jt}^{h} = \beta_q \, log \, r_{jt-1}^{hf} + \beta_d \, log \, Q_{t-1}^{f} + \delta_\tau \, log \, \tau_{t-1}^{hf} + \beta_f \, log \, g_{t-1}(\cdot) + \beta_a log \sum_i a_{ijt-1}^{\gamma} + \delta_P d_f + \delta_t d_t + \varepsilon_t^{h\,f}$$

where d_f, d_t denote the fixed effects for destinations and time; and $\varepsilon_t^{h\,f}$ denotes stochastic components which is *i.i.d.* across destinations and time.

The coefficients of the main estimable equation are denoted by δ_τ for trade resistance, and β_b for $b = \left\{r_{jt-1}^{hf}, Q_{t-1}^{f}, g_{t-1}(\cdot), \sum_i a_{ijt-1}^{\gamma}\right\}$ and δ_c for $c = \left\{d_f, d_t\right\}$ as opposed to the true parameters having σ and unity as its value. The null hypotheses are $\delta_\tau = \sigma$, $\beta_b = 1$ and $\delta_c = 1$. As for the elasticity of substitution, I have two options. The first option is to assume that $\sigma = -7$. The second option is to use the elasticity of substitution from the estimated revenue production function. In this case,

[20] I use a product's deflated revenue as a proxy for physical output. A product' physical output is not available in general.

the elasticity of substitution varies by sector. However, the different elasticity of substitution and various effects of trade resistance on productivity is not my major objective. Thus, in this chapter, I simply assume $\sigma = -7$.

4.4 Learning-by-Exporting

Table 4.1 displays the results of examining the learning-by-exporting hypothesis. Column (1) shows the estimated coefficients from Tobit, Column (2) shows the marginal effects of Column (1), and Column (3) shows the original coefficient divided by $|\sigma = 7|$.

During 1997 – 2002, previous export value has a positive and statistically significant effect on productivity. According to the learning-by-exporting hypothesis building an international network through export transactions provides a channel of productivity improvement because exporting firms have access to better technology, better management skills, more skilled workers who can manage foreign networks, product innovation, international patenting and R&D efforts. The positive coefficient on previous export value supports my hypothesis.

I expect a negative and significant sign on tariff. If so, the negative sign on previous tariff supports the learning-by-exporting hypothesis because it implies that high tariff rate impedes the diffusion of knowledge through trade. Under the low tariff rate exporting circumstance, exporting firms find it relatively easy to construct the international network through export transactions where they can learn from foreign firms and improve their productivity. Although the coefficient on tariff is negative, it is not statistically significant.

Table 4.1 Estimation Results of the Learning-by-Exporting

Coefficients On	**GLS (1)**	**Original Coefficient($\lvert\sigma = 7\rvert$) (3)**
Previous Export Value	0.0047** (0.0023)	-
Destination's Previous GDP per capita	0.0055 (0.0047)	-
Previous Tariff	-0.0400 (0.0751)	-0.0057
Previous Production from the Homogeneous Tech.	0.0016 (0.0029)	-
Previous Specialization of Production	0.0792*** (0.0127)	-
Previous Average Exchange Rate	-0.0003 (0.0104)	-
Previous Variance of Exchange Rate	-0.0017 (0.0048)	-
Distance	-0.0048 (0.0124)	-0.0006
Shared Border	-0.0052 (0.0325)	-0.0008
Colonial Link	-0.0276 (0.0287)	0.0039
Common Language	-0.0013 (0.0124)	-0.0018
Constant	-0.1548 (0.1288)	-
Fixed Effects for Time	Included	-
Fixed Effects for Destination	Not Included	-
Fixed Effects for Sector	Included	-
No. of obs	2,067	-

These results do not contradict the self-selection hypothesis because I take one period time lag to test the learning-by-exporting hypothesis. These estimation results are consistent with Loecker (2013) where he finds significant productivity gains from export entry which is consistent with the self-selection hypothesis. However, these are not consistent with the findings Haidar (2012). He examines the two hypotheses of the self-selection and the learning-by-exporting using Indian firms' balance sheet and income statement information from 1991 to 2004. He confirms the self-selection hypothesis only. For learning-by-exporting hypothesis, he builds the hypothesis whether attaining exporter status significantly improves exporter's productivity. The hypothesis in this chapter is different from that in Haidar, and I do not test the effect of attaining exporter status on productivity, but do test the effect of export relationship on productivity.

The coefficient on destination's previous GDP per capita has a positive sign, but it is statistically insignificant. I assume that a destination's high GDP per capita of implies a destination's more advanced technology, and I expect that the coefficient of destination's previous GDP per capita should have positive sign.

Previous production from homogeneous technology has a positive and statistically significant effect on productivity. The positive sign can be interpreted as learning-by-producing. As a firm produces more output, the firm's productivity improves because it accumulates experience.

The previous specialization of production has a positive and significant effect on productivity. In Chapter 2, I find that the number of a firm's products has an effect on its productivity, but specialization of production does not affect productivity. I also showed that both the number of products and the specialization of production provide different information and one variable explains the half of the other variable. Here the previous specialization of production capture the effect of the number of products on productivity and has a positive effect on a productivity.

Both the average and the variance of exchange rate have a negative effect on productivity, but both are statistically insignificant meaning that there is no evidence that exchange rate has a effect on productivity. Distance also has no effect on productivity.

4.5 Conclusion

The causality between productivity and exports has been a subject of debate. While most literature supports the self-selection hypothesis meaning that high productivity firms self-select to export, the learning-by-exporting hypothesis points out the possibility of productivity improvement through international market participation.

In this chapter, I examine whether past export experience improves productivity of a firm using Indian firms' balance sheets and exporting records from 1997 to 2002. This chapter provides evidence supporting the learning-by-exporting hypothesis where previous export experience improves productivity of firms because export traction provides an access to advanced foreign technology, management, skilled workers. I also find that there could be learning-by-producing besides the learning-by-exporting. Both the previous production from the homogeneous technology and the previous specialization of production improve firm's productivity. However, exchange rate and distance do not have any effect on productivity.

This chapter shows that targeting non-exporting firms for export participation can raise overall productivity of firms, industries and the economy leading to welfare improvements.

5 References

Ackerberg, Daniel, Kevin Caves, and Garth Frazer. 2006. “Structural Identification of Production Functions.” mimeo, R&R Econometrica

Amiti, Mary and Jozef Konings. 2007. “Trade Liberalization, intermediate Inputs, and Productivity: Evidence from Indonesia.” American Economic Review, 97. pp.1611-1638.

Anderson, James E. 1979. “A Theoretical Foundation for the Gravity Equation” American Economic Review 69:106-16.

Anderson, James E. and Eric Van Wincoop. 2003. “Gravity with Gravitas: A Solution to the Border Puzzle” American Economic Review. 93(1), pp.170-192.

Anderson, James E. and Eric Van Wincoop. 2004. “Trade Costs” Journal of Economic Literature, 42(3), pp.691-751.

Aw, Bee Yan, Mark J. Roberts, and Daniel Yi Xu. 2008. “R&D Investments, Exporting, and the Evolution of Firm Productivity” American Economic Review. 98:2. pp.451-456.

Aw, Bee Yan, Mark J. Roberts, and Daniel Yi Xu. 2011. “R&D Investments, Exporting, and Productivity Dynamics” American Economic Review. 101. pp.1312-1344.

Baldwin, Richard and James Harrigan. 2011. “Zeros, Quality, and Space: Trade Theory and Trade Evidence.” American Economic Journal: Microeconomics NO.3, pp.60-88.

Bernard, Andrew B. and J. Bradford Jensen. 2004. “Why Some firms Export.” Review of Economics and Statistics, 86, pp.561-569.

Bernard, Andrew B., Stephen J. Redding, and Peter K. Schott. 2007. “Firms in International Trade.” Journal of Economic Perspectives, Vol.21, No.3, pp.105-130.

Bernard, Andrew B., Stephen J. Redding, and Peter K. Schott. 2010. “Multiple-Product Firms and Product Switching.” American Economic Review 2010.

Bernard, Andrew B., Stephen J. Redding, and Peter K. Schott. 2011. “Multiple-Product Firms and Trade Liberalization.” The Quarterly Journal of Economics, Vol.126(3), pp.1271-1318.

Berthelon, Mathias and Caroline L. Freund. 2008. “On the Conservation of Distance in International Trade” Journal of International Economics, Vol.75(2), pp. 310-320.

Bleaney, Michael and Abelardo Salazar Neaves. 2013. “Declining Distance Effects in international Trade: Some Country-level Evidence” The World Economy, Vol.36, Is.8, pp.1029-1040.

Brun, Jean-François, Céline Carrére, Patrick Guillaumont and Jamiel de Melo. 2005. “Has Distance Died? Evidence from a Panel Gravity Model.” World Bank Economic Review, Vol.54(1), pp.99-120.

Chaney, Thomas. 2008. “Distorted Gravity: The Intensive and Extensive Margins of International Trade” American Economic Review, 98:4, pp.1707-1721.

Chaney, Thomas. 2013. “The Gravity Equation in International Trade: An Explanation” NBER Working Paper No. 19285

Desmet, Klaus, and Stephen L. Parente. 2010. “Bigger is Better: Market Size, Demand Elasticity, and Innovation” International Economic Review, Vol.51, Issue 2. pp.319-333.

Eaton, Jonathan, Samuel Kortum, and Francis Kramarz. 2011. “An Anatomy of International Trade: Evidence from French Firms” Econometrica, Vol.79, No.5, pp.1453-1498.

Eckel, Carsten and J. Peter Neary. 2010. “Multi-product Firms and Flexible Manufacturing in the Global Economy.” Review of Economic Studies, 77, pp.188-217.

Feenstra, Robert and Hiau L. Kee. 2008. “Export Variety and Country Productivity: Estimating the Monopolistic Competition Model with Endogenous Productivity.” Journal of international Economics, 74, pp.500-518.

Fernandes, Ana Margarida, and Alberto Isgut. 2005. “Learning-by-doing, learning-by-exporting, and productivity: evidence from Colombia” The World Bank Policy Research Working Paper No. 3544, Washington, DC.

Goldberg, Linda S., and Cédic Tille. 2008. “Vehicle Currency Use in International Trade.” Journal of International Economics, 76, pp.177-192.

Goldberg, Pinelopi Koujianou. 1995. “Product Differentiation and Oligopoly in International markets: The Case of the U.S. Automobile Industry.” Econometrica, Econometric Society, Vol.63(4), pp.891-951.

Goldberg, Pinelopi Koujianou, Amit Khandelwal, Nina Pavcnik, and Petia Topalova. 2009. “Trade Liberalization and New Imported Inputs.” American Review Papers and Proceedings, 99(2), pp.494-500.

Goldberg, Pinelopi Koujianou, Amit Khandelwal, Nina Pavcnik, and Petia Topalova.

2010. “Multi-product Firms and Product Turnover in the Developing World: Evidence from India.” Review of Economics and Statistics, 92(4), pp.1042-1049.

Goldberg, Pinelopi K., Amit Khandelwal, Nina Pavcnik, and Petia Topalova. 2010. “Imported Intermediate Inputs and Domestic Product Growth: Evidence from India.” Quarterly Journal of Economics, 125(4), pp.1727-1767.

Haidar, Jamal Ibrahim. 2012. “Trade and Productivity: Self-selection or Learning-by-Exporting in India” Economic Modelling 29, pp.1766-1773.

Kraay, Aart. 1999. “Exports and Economic Performance: Evidence from the Panel of Chinese Enterprises (in English)”, “*Expectations et Performances Economiques: Etude d’un Panel d’Enterprises Chinoises*” Revue d’Economie Du Developpement, Vol.1-2/1999, pp.183-207

Keller, Wolfgang. 2004. “International Technology Diffusion” Journal of Economic Literature 42(3): pp.752-782.

Klette, T. J., and Zvi Griliches. 1996. “The Inconsistency of Common Scale Estimators When Output Prices are Unobserved and Endogenous.” Journal of Applied Econometrics, 11, pp.343-361.

Helpman, Elhanan. 1987. “Imperfect Competition and International Trade: Evidence from Fourteen Industrial Countries” Journal of the Japanese and International Economies 1. pp.62-81.

Helpman, Elhanan, Marc Melitz, and Yona Rubinstein. 2008. “Estimating Trade Flows: Trading Partners and Trading volumes” The Quarterly Journal of Economics, Vol.CXXIII Issu 2, pp.441-487.

Levinsohn, James and Amil K. Petrin. 2003. “Estimating Production Functions Using Inputs to Control for Unobservables.” *Review of Economic Studies*, Vol.70(2), No.243, pp.317-342

Loecker, Jan De. 2011. “Product Differentiation, Multi-product Firms and Estimating the Impact of Trade Liberalization on Productivity.” Econometrica, Vol.79, No.5, pp.1407-1451.

Loecker. Jan De. 2013. “Detecting Learning by Exporting” American Economic Journal: Microeconomics 2013, 5(3): pp.1-21

Loecker, Jan De, and Pinelopi K. Goldberg, Amit K. Khandelwal, and Nina Pavcnik. 2012. “Prices, Markups and Trade Reform.” NBER Working Papers 18739.

Mayer, Theierry and Soledad Zignago. 2006. “Notes on CEPII’s distance measures” MPRA paper 31243

Mayer, Thierry, Marc J. Melitz and Gianmarco I. P. Ottaviano. 2011. “Market Size, Competition, and the Product Mix of Exporters.” NBER Working Paper Series 16959/ CEPR DP 8349.

McCallum, John. 1995. “Nathoanl Borders Matter” American Economic Review 85. pp.615-623.

Melitz, Marc J. 2003. “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity.” Econometrica, Vol.71, No.6., pp.1695-1725.

Melitz, Marc J., and Gianmarco I. P. Ottaviano. 2008. “Market Size, Trade, and Productivity.” Review of Economic Studies 75, pp.295-316

Melitz, Marc J. and James A. Costantini. 2008. “The Dynamics of Firm-level Adjustment to Trade Liberalization” In The Organization of Firms in a Global Economy, Helpman, E., D. Marin and T. Verdier. Cambridge: Harvard University Press.

Pavnic, Nina. 2003. “Trade Liberalization, Exit, and Productivity Improvements: Evidence from Chilean Plants.” Review of Economic Studies, 69: pp.245-276.

Prescott, Edward C. 1998. “Needed: A Theory of Total Factor Productivity.” International Economic Review Vol.39, No.3, pp.525-551.

Redding, Stephen J. 2010. “Theories of Heterogeneous Firms and Trade” CEP discussion Paper No. 994

Restuccia, Diego and Richard Rogerson. 2013. “Misallocation and Productivity.” Review of Economic Dynamics 16. pp.1-10.

Topalova, Petia, and Amit Khandelwal. 2011. “Trade Liberalization and Firm Productivity: The Case of India.” The Review of Economics and Statistics, 93(3), pp.995-1009.

Wagner, Joachim. 2007. “Exports and Productivity: A Survey of the Evidence from Firm-level Data” The World Economy, 30(1), pp.60-82