5-12-2024

Essays on Foreign Direct Investment, Trade, and Export Quality

Yinhan Zhang

Syracuse University

Abstract

This dissertation studies the impact of foreign direct investment (FDI) on the exports of developing countries, drawing on evidence from China. It is composed of two chapters.

In Chapter 1, I analyze the effects of R&D FDI on export quality, where R&D FDI is investments aimed at establishing offshore research facilities. I construct a novel dataset on China's outbound FDI using supervised machine learning. I analyze over $26,000$ pieces of textual information on the primary business activities of Chinese overseas subsidiaries collected by the Ministry of Commerce and identify the objective of each outbound FDI project. I find a positive correlation between R&D FDI and the export quality of Chinese firms. This correlation is especially pronounced in industries with a large scope for product differentiation. Conversely, firms engaging in other forms of FDI do not experience quality improvements post-investment. I develop a partial equilibrium model featuring heterogeneous firms with endogenous quality and production fragmentation to theorize a mechanism for quality upgrading: hiring offshore experts with cutting-edge innovation capabilities. The results in this chapter have significant policy implications for China and the US, where the former promotes outbound FDI as a development strategy, but the latter views it as a threat to national interests.

In Chapter 2, I examine the impact of R&D FDI on the extensive margins of exports. Using firm-level data from China, I find that for multi-product exporters engaging in R&D FDI, their number of exported products increases by 35.4%, and their number of countries served increases by 13.3% post-investment. Other forms of FDI, such as marketing and

distribution FDI, cannot fully explain this increase among Chinese firms. Additionally, the quality of new products and country entrants is the highest. Continuing exports rank second, and exits rank last.

ESSAYS ON FOREIGN DIRECT INVESTMENT, TRADE, AND EXPORT QUALITY

by

Yinhan Zhang

B.S., Pennsylvania Western University, 2009

M.B.A., The College of William and Mary, 2011

Dissertation

Submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in *Economics*.

Syracuse University

June 2024

Acknowledgements

I would like to express my deepest gratitude to my advisor, Devashish Mitra, for his continuous support, guidance, and patience throughout my Ph.D. studies. His profound expertise and calming wisdom have been a beacon, guiding me through the most challenging times and made completing this dissertation possible.

I am also deeply indebted to the members of my dissertation committee: Gary V. Engelhardt, Mengxiao Liu, and Thomas Pearson. Their invaluable comments and suggestions have been instrumental in shaping my research and fostering my intellectual growth, and their unwavering support and optimism have been a constant source of motivation, especially during the job market phase.

I especially thank my former advisor, Mary E. Lovely, whose early guidance laid a solid foundation for my research. I have also benefited greatly from discussions with other faculty in the Economics Department, including Kristy Buzard, Ryan Monarch, Alex Rothenberg, and Stuart Rosenthal. I would also like to thank Yoonseok Lee, the Director of Graduate Studies, for his help and support.

I could not have persevered without the help and encouragement of my fellow graduate students. I extend my warmest appreciation to Dahae Choo, Stephanie Coffey, Jingxuan Du, Brian Green, Ross Jestrab, Shuyuan Li, Guanyu Liu, Sam Saltmarsh, Zeyuan "Victor" Xiong, Yi Yang, Yimin Yi, and Tianyun Zhu. I will always cherish the memories of our engaging conversations and fun times spent together. Special thanks to Jingxuan Du, whose friendship was a source of immense strength during the last year of my Ph.D.

To my friends and whole extended family, words cannot even begin to express my gratitude. Your love and understanding have illuminated the darkest days of my life. I would like to thank my lifelong friend, Sophy Ma, for walking with me through the ups and downs of my Ph.D. journey. Last but not least, I must thank my parents, whose unconditional love and support gave me the courage to pursue an education on the other side of the world. I dedicate my work to them.

Contents

List of Figures

List of Tables

Chapter 1

Catching Up by Going Abroad: R&D FDI and Export Quality Improvement

1.1 Introduction

In the last two decades, China, traditionally known for producing cheap, low-value goods for exports, has progressively moved up the export quality ladder. Meanwhile, China's outbound foreign direct investment (FDI) has grown substantially, from a modest \$195 million in 2000 to \$145.6 billion in 2015, an impressive 158-fold increase. Figure 1.1 depicts this growth. Unlike the classic case of FDI, where capital flows from developed countries to developing ones—for example, Apple investing in China to take advantage of the lower cost of unskilled labor—Chinese FDI frequently targets developed countries as it seeks to acquire advanced technology and know-how from the global North (Deng, 2007; Luo & Tung, 2018). In some notable instances, Chinese firms have established R&D centers in these developed countries and hired local experts to carry out R&D (Di Minin et al., 2012; Schaefer, 2020; Schaefer & Liefner, 2017). More recently, however, political leaders in the United States and Europe have become increasingly concerned that China may be gaining an unfair competitive edge through these investments, leading to some proposals to ban Chinese FDI in certain sensitive industries.

Does engaging in R&D FDI enhance the export quality of Chinese firms? Despite considerable debate in the political sphere, the economics literature has yet to address this question.[1] Given the well-established positive association between firms' product quality and exporting success (Bastos & Silva, 2010; Brooks, 2006; Crozet et al., 2012; Manova & Zhang, 2012), as well as countries' long-run economic prosperity (Khandelwal et al., 2013), understanding how developing country firms could move up the export quality ladder is significant.

Why would R&D FDI affect export quality? To formalize the mechanism, I develop a simple two-country trade model in partial equilibrium with Melitz (2003)-type firms. Firms "produce" quality by employing high-skill workers to conduct R&D. Crucially, there is a disparity in the R&D efficiency between the high-skill workers located in the North and those in the South, with the North's high-skill workers being decidedly more efficient—a fact that is well documented in the business literature (Section 1.2.1). After the South liberalizes its outbound FDI, the most productive Southern firms, who are also exporters, can choose to relocate their R&D activities to the North through FDI while still retaining their manufacturing capacity in the South. Conducting R&D FDI will incur an additional fixed cost. However, it also gives those firms who do so access to the superior high-skill labor pool in the North, thereby reducing their cost of quality upgrading. Thus, from this model setup, I hypothesize that firms' export quality increases post-R&D FDI.

A central empirical challenge in studying the impact of Chinese outbound FDI is the lack of publicly available data on the objectives of these FDI projects. However, distinguishing between these objectives is crucial, as different types of FDI are motivated by different

[1] A small body of literature empirically examines the impact of Chinese FDI in general on various dimensions of firm performance, including productivity, trade, and the size of operation (Chen & Tang, 2014; Cozza et al., 2015; Huang & Zhang, 2017; Yan et al., 2023). Among these papers, only Yan et al. (2023) focus on export quality as the primary outcome of interest. Chen and Tang (2014) use export unit value to proxy for quality, while this paper employs Khandelwal et al. (2013)'s quality measure. Notably, with the exception of Huang and Zhang (2017), all of these studies treat FDI as a single variable that includes all types of FDI, thus making it difficult to cleanly identify the underlying factors driving their findings.

considerations, which could lead to different outcomes. For instance, while projects that establish distribution networks may impact trade, projects that drill for natural gas would not have the same effect. Thus, studies that treat all types of FDI as one homogeneous variable, even when they yield significant findings, would inevitably invite this question: What precisely is responsible for the observed outcomes?

In this paper, I overcome this challenge by constructing a new dataset on Chinese outbound FDI. This dataset includes critical information about the objective or type of each FDI project. The data is based on the Directory of Overseas Investment Enterprises, a comprehensive list of Chinese FDI projects compiled by the Ministry of Commerce (MOFCOM). Until 2014, Chinese firms that wished to invest overseas had to undergo a stringent and burdensome application process (Section 1.2). As part of this process, firms reported their reasons for doing FDI to MOFCOM, usually in brief 30 to 80 Chinese characters. By utilizing this information, a researcher can extrapolate the economic motivations behind each FDI project.

However, because firms write these texts in the natural language, conventional methods such as regular expression in Python cannot fully capture the complexity of the texts to generate precise categorizations. Therefore, I use text classification, a supervised machine learning technique, to automate the categorization of the MOFCOM FDI projects. I first drop FDIs destined for tax havens from the raw data, leaving over $26,000$ observations to analyze and categorize (Section 1.3.1). I then construct a training sample by selecting $2,000$ observations from the $26,000$ pool and carefully assigning each project to an FDI category (Section 1.3.1). Finally, I use this sample to train the machine learning model. The final model achieves a remarkable 95% accuracy rate when tested against the training sample, a significant improvement over the 70% accuracy rate I previously achieved using the regular expression method.

With the FDI dataset in hand, I test the central prediction of my model by employing a two-way fixed effect (TWFE) specification with firm productivity, product-year, country-year, and province-year fixed effects as controls. Data on Chinese exports comes from the General Administration of Customs of China, while information on firm operations, such as revenue, assets, and the number of employees, comes from the Annual Survey of Industrial Firms. Combining these two datasets creates a 14-year unbalanced panel spanning 2000 to 2013, with close to 13 million observations at the firm-product-(export destination) country-year level. Because product quality is not observable in the trade data, I infer it using the approach of Khandelwal et al. (2013), where quality is the residual of an OLS regression on the export demand function. The intuition of this approach is straightforward: conditional on price, higher-quality products sell more units and, therefore, have a higher residual. My main results suggest that firms who conduct R&D FDI experience, on average, a 69.05% increase in export quality post-FDI, a magnitude comparable to the improvements induced by trade liberalization (Bas & Strauss-Kahn, 2015). The results are robust to alternative specifications.

Because firms self-select into FDI, the baseline TWFE model may not be sufficient in mitigating this bias. Therefore, I perform a placebo test by estimating a version of the baseline specification using two other types of FDI—production FDI and distribution FDI—as the treatment variable. This choice of alternative FDIs is rooted in the understanding that, in theory, neither FDI should influence export quality. The ensuing estimates reveal a coefficient close to zero for distribution FDI and a negative, insignificant coefficient for production FDI, suggesting that confounding factors are unlikely to be responsible for the positive baseline result. Additionally, heterogeneity analyses find that the impact of R&D FDI is only significant in sectors where the products are differentiated and where the scope for such differentiation is large. In contrast, the impact is insignificant in sectors where the products are homogeneous. This differentiated outcome aligns well with the quality-upgrading mechanism proposed in this paper: better inputs (high-skill workers) lead to better outputs (quality), but only in sectors where R&D is likely to make any difference.

In addition to the papers already cited, this paper also relates to two other strands of literature. First, it contributes to research on the determinants of export quality upgrading at the firm level. Verhoogen (2023) provides an excellent literature review on this topic, where he groups the determinants into three classes: the supply-side factors, like better material inputs; the demand-side factors, like increased demand from wealthier markets; and changes in firm capacity. While the literature has paid much attention to both the supply-side factors (Bas & Strauss-Kahn, 2015; Fan, Li, & Yeaple, 2018; Fan et al., 2015; Feng et al., 2016; Xu & Mao, 2018) and the demand-side factors (Verhoogen, 2008), relatively little has been done looking at changes in firm capacity. This paper contributes to this last category by arguing that firms may enhance their R&D capacity through FDI.

The second literature this paper contributes to concerns theories on what motivates Chinese firms to undertake outbound FDI. For instance, Fan, Lin, and Tang (2018) show that rising labor costs in China increase the likelihood of Chinese firms conducting production FDI. Tian and Yu (2020) find that high communication costs between international buyers and sellers can induce Chinese firms to pursue marketing and distribution FDI. On the other hand, this paper demonstrates that the desire to seek out high-skill workers could drive Chinese firms to engage in R&D FDI.

The remainder of this paper is organized in the following order: Section 1.2 describes the background of China's outbound FDI liberalization and presents three case studies to illustrate the quality upgrading process. Section 1.3 presents the data and the data cleaning procedures, including the use of supervised machine learning to automatically categorize FDI projects. Section 1.4 develops the theoretical model. Sections 1.5 and 1.6 discuss the empirical strategy and various results, respectively.

1.2 Institutional Background

Between 2000 and 2015, Chinese FDI outflows rose 158 times to $145.6 billion annually, making China the world's third largest contributor of capital after only the United States and Japan.[2] This surge in outbound FDI was certainly a result of China's rapid economic development around the same time; however, it was also profoundly affected by a series of policy reforms implemented by the central government in the early 2000s that liberalized and promoted such investments (Luo et al., 2010). Figure 1.1 shows the growth of outbound FDI closely tracking China's policy reforms.[3]

China's outbound FDI policies underwent three stages. In the first stage, before 1990, Chinese outbound FDI was exceedingly rare. The government, driven by concerns over potential capital flight, imposed stringent controls on all outgoing FDI projects. These controls included caps on the investment dollar amount, typically at \$10 million, and mandates for repatriating all profits earned overseas (Luo et al., 2010). Then, in the 1990s, the government further tightened the FDI approval process. Specifically, Chinese companies intending to invest overseas were required to submit a detailed feasibility report, along with the relevant contracts and by-laws, to the government for approval. Projects exceeding \$1 million were required to seek approval from the State Planning Commission, a former ministry of the national government, and those under \$1 million were reviewed and approved by the provincial governments. Lastly, projects involving state-owned assets or whose investment amount is greater than \$30 million had to gain approval from the State Council, the top administrative body in China.

It was not until 2004 that any meaningful FDI liberalization occurred when new regulations significantly raised the capital threshold requiring central government approval. Under the updated rules, provincial governments had the authority to approve resource development projects of less than \$30 million[4] and other non-resource investments less than \$10

[2] Contrary to the common perception that a developing country is a recipient of FDI rather than a contributor, China's outbound FDI had, by 2015, reached similar levels to its inflow. Such a two-way FDI flow pattern resembles many developed countries.

[3] Section 1.2 provides an overview of the most important policy changes but is by no means exhaustive. For a comprehensive collection of official Chinese documents and materials related to outbound FDI, translated into English, see Bernasconi-Osterwalder et al. (2013).

[4] These are FDI projects involved in the exploration and development of crude oil, minerals, and other natural resources.

million.[5] However, projects surpassing these limits were still required to submit their applications to the National Development and Reform Commission (NDRC), formerly the State Planning Commission, for official approval. "Megaprojects," which are resource development FDI over $200 million or other types of FDI over $50 million, needed approval from the State Council. The feasibility report was no longer necessary. Instead, companies had to seek formal feedback from the Chinese overseas consulates regarding project feasibility, register their businesses with the consulates after establishment, and partake in state-run annual surveys.

From 2009 to 2015, China accelerated the liberalization of outbound FDI. Most notably, in 2009, the government revised the regulations again, stipulating that only projects exceeding $100 million or those intended for countries with no diplomatic ties with China should seek approval from the Ministry of Commerce. On the other hand, provincial governments could approve projects conducted by private companies between $10 million and $100 million and all resource development projects. The requirement to obtain feedback from Chinese consulates before engaging in FDI was dropped for non-resource projects. Finally, in 2014, the government removed all restrictions on projects under $2 billion, except for those involving sensitive countries or industries.[6] Thus, after 2014, most companies only needed to declare their intentions to invest overseas by submitting a simple application to the government.

Concurrent with sweeping FDI liberalization, the Chinese government has also actively encouraged and supported certain types of outbound FDI, especially those that can promote international technology collaboration, strengthen China's research and development capacity and global competitiveness, and boost international trade (Ministry of Commerce).

[5] Provincial governments only handle FDI projects by private companies, while projects by state-owned enterprises (SOE) were entirely exempted from the application process. SOEs engage in FDI at their discretion.

[6] Sensitive countries are any of the following: countries with no diplomatic ties with China, are on the United Nations sanctions list, or are in a war zone. Sensitive industries are telecommunication, water, land development, public utilities, and news media, as well as those listed in China's Catalog of Technologies Prohibited and Restricted from Export.

For example, in 2004, the Ministry of Commerce and the Ministry of Foreign Affairs designated 67 countries and 191 industries for special promotion status, focusing on the major global economies and the manufacturing sector.[7] If an FDI project falls into one of the listed categories, not only is its application more likely to be approved, but the project may also receive preferential treatment in financial support, exchange rates, taxation, and administrative processing. As another example, the NDRC and the Export-Import Bank of China jointly launched a dedicated loans program in 2004 to support FDIs seeking to establish overseas R&D facilities. According to this policy, such R&D facilities must seek to utilize internationally advanced technologies, management experience, and professional talents available in the host country. In return, loans for such projects have lower interest rates, faster approval time, and longer loan terms. Similarly, the China Development Bank and the Ministry of Finance also announced financing support, grants, and direct interest subsidies for R&D FDIs.

1.2.1 Case Studies

While the government supports R&D FDI in manufacturing with the explicit goal of promoting industrial upgrading, it is unclear from the official documents how firms engaged in such FDI can achieve this goal. To better understand the motives of Chinese multinationals and their operation of offshore R&D centers, I turn to case studies in the business literature. Di Minin et al. (2012), Schaefer and Liefner (2017), and Schaefer (2020) together conducted over 50 in-depth interviews with engineers, researchers, and managers at offshore R&D centers of leading Chinese companies. Their insights shed light on two key questions, forming the basis of my theoretical model (Section 1.4).

The first question is: What do Chinese companies do when they set up R&D centers abroad? In Schaefer (2020), the author studies Huawei, a world leader in the telecommuni-

[7] This catalog was subsequently updated in 2005 and 2007 to include more countries and industries. See the Ministry of Commerce website (http://www.mofcom.gov.cn/aarticle/b/bf/200408/20040800258538.html) for details.

cations industry. She finds that starting in the early 2000s, Huawei began establishing R&D labs in the United States and Europe, often near their competitors in cities like San Jose and Munich. Huawei then hired offshore experts directly to conduct R&D, which produced many patents (Schaefer & Liefner, 2017) and product designs. Huawei sent the designs back to China for a larger and less expensive workforce to manufacture the final product. In another study, Di Minin et al. (2012) show that ZTE, JAC Motors, Chang'an Motors, and Hisense Group,[8] and other well-regarded Chinese companies have adopted R&D globalization strategies similar to Huawei's, hiring researchers and designers based in advanced economies to work on product development.

The second question is why Chinese companies hire experts in developed economies despite the higher labor costs. Di Minin et al. (2012), Schaefer and Liefner (2017), and Schaefer (2020) all point to the same reason: offshore experts are superior when it comes to creativity and originality. For example, Di Minin et al. (2012) quote one Chinese engineer who said, "In our experience, Italian designers are more skilled." Additionally, Schaefer (2020) documents that Huawei engineers believed that "while the (technology) gap is rapidly closing, Huawei remains behind when it comes to innovative skills." These papers gave several reasons for the competitive advantage of offshore experts. To begin with, these experts often hold advanced degrees from Western universities, where creativity and originality are prioritized over learning facts. In addition, most have spent years working for the power players in their respective fields, honing their skills. Such knowledge "enables Huawei to produce state-of-the-art products without having to first learn how to create them itself" (Schaefer, 2020). Finally, offshore experts are more embedded in the global industrial network. Their professional connections and understanding of local consumer tastes enable them to create product designs that are better aligned with market demand.

In Section 1.4, I develop a formal model to guide empirical analysis. I incorporate the

[8]ZTE is another major telecommunications producer in China, the chief competitor of Huawei. JAC Motors and Chang'an Motors are automobile manufacturers, and Hisense makes home appliances. These companies are household names in China.

quality-upgrading mechanism gleaned from these case studies: hiring offshore experts who are more effective at R&D.

1.3 Data

In order to empirically evaluate the impact of China's outbound R&D FDI on its export quality at the firm level, I would need to possess, at a minimum, information on the FDI projects themselves, the characteristics of the firms that undertake these FDIs (as well as those firms that do not), and the outcome variable—firm export quality. Second, the datasets should span the period when China's outbound FDI policy underwent the most significant changes—primarily from 2004 to 2013 (Section 1.2)—plus some years preceding this period. Such coverage allows me to compare export quality before and after FDI liberalization. Third, I must know the objective or type of each FDI project. Since different types of FDI are motivated by different economic considerations that produce different outcomes, it would only be possible to test the effects of R&D FDI if I could first reliably distinguish R&D FDI from non-R&D FDI.

I satisfy these data requirements by drawing upon three exceptionally rich and comprehensive datasets: the Directory of Overseas Investment Enterprises from the Ministry of Commerce, the Annual Survey of Industrial Firms compiled by the National Bureau of Statistics, and the universe of Chinese import and export transactions from the General Administration of Customs Data. Each dataset contains different information critical to my analysis. After carefully cleaning and processing the datasets, I combined them to create a 14-year unbalanced panel spanning 2000 to 2013, with close to 13 million observations at the firm-product-(export destination) country-year level.

The next four subsections provide an overview of the three datasets employed in this paper, the cleaning and merging procedures, and most importantly, how I classified over $26,000$ FDI projects using text classification, a supervised machine learning technique.

1.3.1 The Directory of Overseas Investment Enterprises

The Ministry of Commerce (MOFCOM) compiles and makes available the Directory of Overseas Investment Enterprises, a comprehensive list of Chinese outbound FDI projects. The listed FDIs are mostly greenfield, but some are also mergers and acquisitions (M&A) deals.[9] This dataset contains vital details, including the name of the parent company, name of overseas subsidiary, investment destination country, project approval date, and the scope of offshore business activities. MOFCOM collects this information as part of the FDI approval and registration process,[10] which paradoxically, thanks to it being quite stringent and burdensome (Section 1.2), invariably ensures the quality of the data generated. Consider, for example, that until 2014, MOFCOM required signed contracts or framework agreements to be in place before approval. This decree drastically reduced the probability that a Chinese company would casually undertake the application process, misreport the project, or have the host country deny the project later on.

The raw data from MOFCOM has $41,682$ unique FDI projects implemented by $29,597$ state-owned and private enterprises targeting 203 countries. However, of these 203 countries, 41 are tax havens,[11] for which I could not determine the true investment destination, so I removed them from my analysis. The final sample consists of $26,483$ projects by $17,999$ firms between 1987 and 2015.

[9] Although no variable in the data indicates whether a project is greenfield or M&A, I searched the text variable on offshore business activities for keywords linked to M&A, such as "shou gou" and "bing gou" in Chinese. I find that 309 or 1.17% of the $26,483$ projects contain keywords for M&A, while only 0.86% of R&D projects are M&A, a negligible amount.

[10] MOFCOM's provincial offices gather data on FDI projects approved by the local governments. Since 2009, projects with a value below $10 million do not need approval from any government agency but must still file for record with the local government. The local government subsequently compiles and reports this information to MOFCOM.

[11] Tax havens are countries like Hong Kong, Ireland, and the Cayman Islands. I use Hines and Rice (1994)'s list of tax havens to decide which countries to exclude from my analysis.

Machine Learning to Categorize FDI

The most crucial element of the MOFCOM dataset is the scope of business activities of offshore affiliates. Firms, usually in brief 50-100 Chinese characters, detail their FDI motivations, industry, and product offerings. For instance, a firm that wants to establish an offshore R&D center might write "The R&D (of such and such) technologies, establishing R&D bases and platforms, training senior R&D personnel, and recruiting R&D experts." Alternatively, a firm looking to create distribution channels could state "Trading, wholesaling, and retailing (of certain products)." Based on this rich textual information, one could extrapolate the economic motivations behind each FDI project. Nevertheless, there is a hurdle. Since firms write these texts in the natural language, which does not have standard phrases, traditional techniques for large-scale data analysis, such as regular expression in Python, could not fully capture the complexity of the texts because these techniques rely on fixed, predefined logic.[12] Indeed, even after many iterations of Python codes, I still could not achieve a classification accuracy rate exceeding 70%.[13]

To overcome this challenge, I partnered with a seasoned machine learning engineer to deploy text classification, a supervised machine learning technique, to automate the classification of all 26,483 FDI projects.[14] I implement a four-step procedure. First, I use regular expression to roughly label the projects, assigning each project to a specific category. Table 1.1 contains the keywords I reference for the eight FDI categories. Second, I select 2,000 projects, with 250 projects drawn from each category, and manually verify that the initial

[12] Here is an example that underscores the challenge. The presence of keywords such as "research," "development," and "product" would typically indicate R&D. However, the exact keywords put together with keywords related to natural gas or infrastructure development would instead suggest resource or infrastructure FDI. To differentiate between R&D and infrastructure, one must exhaust all keywords (from crude oil and minerals to railroads, tunnels, ports, and power plants) associated with the latter and categorize projects containing such keywords accordingly. To make things even more complex, one must simultaneously distinguish them from the production or distribution FDI that produces or transports the tools and machinery used in mining or construction. The logical combinations are endless.

[13] The 70% figure is an estimate based on randomly selecting 20-30 Python-labeled projects and manually checking for accuracy.

[14] We used the BERT model, a leading natural language processing model published by Google researchers in 2018.

labeling is correct. If the label is incorrect, I would change it to the correct label.[15] Such adjustments often changed the number of projects in each category. To keep the data balanced, I add or subtract projects until each category reaches 250 projects again. This process is incredibly labor-intensive, but it must be done meticulously because it directly affects the quality of the trained model. In step three, I feed my sample of $2,000$ accurately labeled FDI projects into the machine learning program to let it train the model. The final model achieves a 95% accuracy rate in correctly labeling the testing set. Finally, I use the trained machine learning model to categorize all $26,483$ projects.

Table 1.2 shows the distribution of FDI projects by type, as categorized by machine learning. Note that R&D FDI, which is the focus of this research, has $2,212$ projects. Over time, the number of R&D FDI projects follows a growth trajectory similar to total outbound FDI. Table 1.3 shows the number of R&D FDI by year between 1999 and 2015. Interestingly, 10.84% of firms conducted more than one R&D FDI project during the sample period.

1.3.2 The Annual Survey of Industrial Firms

Collected by the National Bureau of Statistics of China, the Annual Survey of Industrial Firms (ASIF) surveys all above-scale industrial firms.[16] From 2000 to 2006, the ASIF data contains every state-owned enterprise (SOE) and non-state-owned enterprise (non-SOE) that reported 5 million RMB or above in sales (USD 687,000 in the 2023 exchange rate). From 2007 to 2010, ASIF drops small SOEs with sales under 5 million RMB. Then, after 2011, the dataset only includes firms with 20 million RMB or above in sales (USD 2.7 million).[17] The

[15] I use a three-step algorithm to determine the appropriate category for a project: 1. I look for keywords listed in Table 1.1. If all the keywords fall into a particular category, such as production, I assign the project to that category. 2. If the keywords fall into different categories, I count the keywords in each category and assign the project to the category with the highest count. 3. In the least common case, when two categories have the same count of keywords, I determine the category based on the order in which the keywords appear.

[16] Industrial firms are firms in the mining, manufacturing, and public utilities sectors. The National Bureau of Statistics classifies 2-digit China Industrial Classification (CIC) 06 to 12 as mining, 13 to 43 as manufacturing, and 44 to 46 as public utilities. Manufacturing is by far the largest category.

[17] In some years, a firm may drop out of the Survey because its sales in the prior year fell below the sales threshold even as the firm remains operational. Later, if the firm's sales improve, the firm will reappear in the Survey. Thus, pooled multi-year Survey data is unbalanced by construction.

ASIF reports on over 100 variables at the firm level, covering the firm's identification (firm ID, name, legal person, address, and industry), operational information (employment, gross output, and value-added), and financial data from the three main accounting statements: the balance sheet, cash flow, and income statement.

Despite the richness and usefulness of the ASIF in studying Chinese firm-related questions, it is well understood that it has many data quality issues. For instance, the ASIF has no unique identifier that can link the firms over time. The firm ID and name variables are known to change when firms change ownership or undergo mergers and acquisitions. In addition, multiple firms may share the same ID, perhaps due to misreporting. However, to study the effects of R&D FDI on firms, I must be able to follow the same firm over time consistently. Therefore, I use a rigorous algorithm developed by Brandt et al. (2012) that matches firms across years using the firm ID, name, legal person, address, and industry information. The matched data is an unbalanced panel with $162,885$ firms in 2000 and $344,875$ firms in 2013. Table A1 shows the number of firms by year and their entry and exit status. On average, the annual attrition rate is 15.5%, with just 3.14% of all firms surviving the entire 14-year sample period.

Another well-understood issue with the ASIF is that it has many noisy observations. Hence, I follow the standard procedure by Cai and Liu (2009) and Yu (2015) and first remove observations with missing, zero, or negative values among the key financial variables, including total assets, the net value of fixed assets, sales, gross industrial output, and the number of employees. Second, I drop firms with fewer than eight employees and observations that violate the Generally Accepted Accounting Principles (Yu, 2015).[18] Lastly, I exclude intermediary firms,[19] as they do not produce anything in-house, and keep only a sample of

[18]The rules are: 1. total assets must be greater than liquid assets; 2. total assets must be greater than total fixed assets; 3. total assets must be greater than the net value of the fixed assets; 4. the establishment date must be valid (the opening month is between January and December).

[19]Intermediary firms are identified by the presence of terms such as "trading" and "importing and exporting" in their firm names. This paper searches for the Chinese characters "贸易" (mào yì), "外贸" (wài mào), "外经" (wài jīng), "进出口" (jìn chū kǒu), "经贸" (jīng mào), "工贸" (gōng mào), "科贸" (kē mào), and "边贸" (biān mào) in firm names.

manufacturing firms.[20] The final filtered ASIF sample has $3,519,514$ firm-year observations between 2000 and 2013.

1.3.3 The General Administration of Customs Data

The Chinese General Administration of Customs compiles a highly disaggregated trade dataset covering the universe of Chinese imports and exports. For each trade transaction, this dataset records the Harmonized System 8-digit product code (HS8), partner country, year, the customs regime,[21] as well as the transactions' value and quantity.[22] These variables are necessary for estimating product quality, which is the dependent variable of this paper. Conveniently, the Customs data also contains detailed information on the firms engaging in the trade, such as the firm name, zip code, phone number, and contact person, which allows the researcher to merge the trade data with the ASIF data.

To clean the Customs data, I closely follow Fan et al. (2015) and Brandt et al. (2012). First, I drop all intermediary firms from the sample, as is the case for ASIF. Second, I aggregate the data from the HS 8-digit to the HS 6-digit level (HS6). This transformation allows me to concord the product codes over time, as concordance tables for China's HS8 code are unavailable. I convert the HS2002, HS2007, and HS2012 6-digit codes into HS1996 using the United Nations Statistics Division correspondence tables. Lastly, I deflate the export values using Brandt et al. (2012)'s output deflators. The original deflators from Brandt et al. (2012) are in the 4-digit China Industrial Classification (CIC) and are only available until 2007. I use the HS-CIC concordance table and Stata code from Brandt et al. (2017) to bridge the CIC code and the HS6 code, then calculate deflators for each year up to 2013. Finally, the unit value is the deflated export value divided by the physical quantity. In the end, I have $49,719,378$ observations for exports between 2000 and 2013 at the firm-HS6

[20] I keep industries whose 2-digit CIC is from 13 to 43, including the endpoints. For the years before 2003, 4-digit CIC 1711, 1712, 1713, 1714, 1719, 2220, 3648, 3783, 4183, and 4280 are also dropped from the sample as these CICs are special cases of services (Brandt et al., 2012).

[21] China uses 20 different customs regimes to indicate the purpose of trade, e.g., ordinary trade, processing and assembling, and processing with imported material.

[22] Value is (Free on Board) FOB for exports and CIF (Cost, insurance, and freight) for imports.

product-country-year level.

1.3.4 Merging Data

I combine all three datasets to construct the final sample for regression analysis. First, I match the ASIF data and the Customs data. Because these two datasets have no common keys, I follow Yu (2015) and use a two-step matching process. First, I match by firm name and year. Then, I match by year, zip code, and the last seven digits of the firm's phone number. The matched sample from 2000 to 2013 consists of $12,944,206$ observations at the firm-HS6 product-country-year level. Table A2 shows the matching outcome by year. Overall, the matched sample covers 31.5% of all exporters and 47.7% of the total export value. Although the matching rate by the number of exporters is lower than the literature's, the matching rate by export value is comparable to Fan et al. (2015). Finally, I match the FDI project data with the merged ASIF-Customs data using the firm name and year information.

1.4 Model

I develop a simple two-country, monopolistically competitive, heterogeneous firm model with endogenous quality in partial equilibrium. The model's production structure closely follows Johnson (2012), which I extend to allow for production fragmentation by vertical FDI, referencing in Helpman (1984).

1.4.1 Preference

Two countries, North and South, are indexed by $j = n, s$. In country j, there is a representative consumer whose constant elasticity of substitution (CES) preferences is given by

$$U_j = \Big(\int_{\Omega_j} [q_j(\omega) x_j(\omega)]^{\frac{\sigma-1}{\sigma}} d\omega \Big)^{\frac{\sigma}{\sigma-1}} \tag{1.1}$$

where ω is a variety of a differentiated good, and Ω_j represents the set of all differentiated varieties available in country j. Note that this differentiated good sector is small compared to the overall economy. $x_j(\omega)$ is the quantity of variety ω, counted in the physical unit, and $q_j(\omega)$ is the quality of variety ω. Here, quality acts as a utility shifter, capturing all attributes of the variety that appeal to consumers, excluding price. Thus, if prices are the same, higher-quality varieties are consumed in larger quantities. Finally, $\sigma > 1$ is the elasticity of substitution between the varieties. Standard optimization yields:

$$x_j(\omega) = [q_j(\omega)]^{\sigma-1} [p_j(\omega)]^{-\sigma} \frac{M_j}{P_j^{1-\sigma}} \tag{1.2}$$

where $p_j(\omega)$ is the price of variety ω in country j, and M_j is the portion of aggregate income that country j spents on the differentiated good. P_j denotes the CES aggregate price index:

$$P_j = \Big(\int_{\Omega_j} \big[p_j(\omega)/q_j(\omega)\big]^{1-\sigma} d\omega \Big)^{1/(1-\sigma)}$$

Thus far, my setup mirrors Johnson (2012), Fan et al. (2015), and others who embed endogenous quality into the Melitz (2003) model.

1.4.2 Firm Production

Firms in country j engage in monopolistic competition. Each firm produces a single variety ω of the differentiated good. An idiosyncratic productivity parameter φ with density $g(\varphi)$ over $(0, \infty)$ denotes the physical efficiency of each firm.

A firm chooses the price and the quality of the variety ω it produces. Firstly, the firm invests $f_j(q_j(\omega))$ to develop a product with the quality level $q_j(\omega)$. This step is akin to the R&D stage, and $f_j(q_j(\omega))$ is the fixed cost of quality upgrading. Following Johnson (2012), I let $f_j(q_j(\omega))$ to take a specific function form: $w_j^h[q_j(\omega)]^\alpha$, where $\alpha > 0$ and $q_j(\omega) = (l_j^h)^{1/\alpha}$. I assume that "quality production", or R&D, requires high-skill labor l_j^h as inputs, whose efficiency-adjusted wage in country j is w_j^h. Second, the firm employs low-skill labor l_j^l to manufacture the physical outputs with constant returns to scale technology. I let $MC_j(q_j(\omega), \varphi) = \frac{w_j^l q_j(\omega)^\beta}{\varphi}$ and assume $0 < \beta < 1$ and $0 < (1-\beta)(\sigma - 1) < \alpha$. These assumptions ensure the existence of a solution to the firm's optimization problem. w_j^l is the low-skill labor's efficiency-adjusted wage. Note here that the marginal cost of manufacturing $MC_j(q_j(\omega), \varphi)$ is increasing in the quality of the variety.

A firm can sell its outputs domestically or export to the foreign market. In the first case, the firm pays a fixed cost of f_D, while in the second, f_X. The costs are measured in the home country's wage for high-skill labor. Let $j = n, s$ index the country where production occurs, and $i = n, s$ index the destination country. In addition, exporting also incurs a $\tau_{ji} > 1$ iceberg trade cost, where firms must ship $\tau_{ji} > 1$ units of good from country j for one unit to arrive in country i ($\tau_{jj} = 1$).

Finally, I take inspiration from Verhoogen (2008) and treat a firm as having two separate production lines, one for the domestic market and one for the export market, each producing a variety with a specific level of quality for that market. This separability leads the firm to maximize profit for each market independently.

The firm's optimal quality and price are:

$$q_{ji}^*(\varphi) = \Big[B_i\Big(\frac{1-\beta}{\alpha}\Big)\Big(\frac{\sigma-1}{\sigma}\Big)^{\sigma}\Big(\frac{\tau_{ji}w_j^l}{\varphi}\Big)^{1-\sigma}\frac{1}{w_j^h}\Big]^{\frac{1}{\alpha-(1-\beta)(\sigma-1)}} \tag{1.3}$$

and

$$p_{ji}^*(\varphi) = \Big[B_i\Big(\frac{1-\beta}{\alpha}\Big)\Big(\frac{\sigma-1}{\sigma}\Big)^{\frac{\beta-\alpha+\sigma-1}{\beta}}\Big(\frac{\tau_{ji}w_j^l}{\varphi}\Big)^{\frac{\alpha-\sigma+1}{\beta}}\frac{1}{w_j^h}\Big]^{\frac{\beta}{\alpha-(1-\beta)(\sigma-1)}} \tag{1.4}$$

where $B_i = \dfrac{M_i}{P_i^{1-\sigma}}$. $P_i^{1-\sigma}$ now involves aggregating the prices over every firm that sells in market i:

$$P_i = \Big(\int_{\varphi_{ii}^{\min}}^{\infty} \big[p_{ii}(\varphi)/q_{ii}(\varphi)\big]^{1-\sigma} g(\varphi)d\varphi\Big)^{1/(1-\sigma)} + \Big(\int_{\varphi_{ji}^{\min}}^{\infty} \big[p_{ji}(\varphi)/q_{ji}(\varphi)\big]^{1-\sigma} g(\varphi)d\varphi\Big)^{1/(1-\sigma)}$$

Equations 1.3 and 1.4 indicate that firms with greater productivity produce higher quality goods, though they may not charge higher prices.[23] On the other hand, conditional on φ, quality is decreasing in both skilled and unskilled wages and increasing in market demand B_i.

The profits from domestic sales and exports are:

$$\pi_{jj}^{d*}(\varphi) = \Phi \cdot B_j^{\frac{\alpha}{\kappa}} \left(\frac{1}{w_j^h}\right)^{\frac{\alpha-\kappa}{\kappa}} \varphi^{\frac{(\sigma-1)\cdot\alpha}{\kappa}} - f_D w_j^h \tag{1.5}$$

$$\pi_{ji}^{e*}(\varphi) = \Phi \cdot (\tau_{ji}^{1-\sigma} B_i)^{\frac{\alpha}{\kappa}} \left(\frac{1}{w_j^h}\right)^{\frac{\alpha-\kappa}{\kappa}} \varphi^{\frac{(\sigma-1)\cdot\alpha}{\kappa}} - f_X w_j^h \tag{1.6}$$

where $\Phi \equiv \frac{\kappa}{\alpha-\kappa}\left[\left(\frac{1-\beta}{\alpha}\right)\left(\frac{\sigma-1}{\sigma}\right)^{\sigma}(w_j^l)^{1-\sigma}\right]^{\frac{\alpha}{\kappa}}$ and $\kappa \equiv \alpha - (1-\beta)(\sigma-1)$.

Because profitability is increasing in φ for π_{jj}^{d*} and π_{ji}^{e*}, the initial fixed costs of f_D and f_X imply that there exist two entry cut-offs, $\varphi_{ii}^{\min}$ and $\varphi_{ji}^{\min}$, where they satisfy $\pi_{jj}^{d*}(\varphi_{ii}^{\min}) = 0$ and $\pi_{ji}^{e*}(\varphi_{ji}^{\min}) = 0$.

1.4.3 The Southern Firm and Effect of Outbound FDI Liberalization

In the following discussion, I focus solely on the decisions facing firms in the South, wherein outbound FDI liberalization takes place.

[23] In the optimal price function (Equation 1.4), the exponent of φ is $\frac{\sigma-1-\alpha}{\alpha-(1-\beta)(\sigma-1)}$. While the assumption is that $\alpha - (1-\beta)(\sigma-1) > 0$, whether $\sigma - 1 - \alpha > 0$ is uncertain. Consequently, the correlation between price and productivity remains indeterminate.

Fragmenting Activities

Until now, I have assumed that R&D and manufacturing must occur in the same geographic location. A Southern firm hires skilled and unskilled workers locally and pays them the prevailing Southern wages, denoted as w_s^h and w_s^l (the subscript s denotes South), respectively.

Nevertheless, firms can fragment their production activities across different countries, as in Helpman (1984). In Helpman (1984)'s framework, a firm first hires skilled labor in its home country, implicitly assumed to be a developed economy, to create a firm-specific intangible asset, e.g., managerial expertise, product designs, or other forms of intellectual property. Later, the firm transfers this intangible asset to its production plants located globally, and the plants then utilize the intangible asset to manufacture the physical output.

Drawing on Helpman (1984), I allow the Southern firms to split their production process and relocate stage one, the R&D stage, to the North by establishing research-oriented subsidiaries there. In reality, such international expansion only becomes possible after the South (which, in this paper, is China) liberalizes its outbound FDI. To expand overseas, the Southern firm must pay an additional fixed cost of f_I in w_s^h. However, such a firm also benefits from the FDI because it gains access to the North's pool of skilled labor, which is superior. The North's skilled wage is w_n^h. Most critically, I assume that $w_n^h < w_s^h$. This assumption draws on the fact that wages are efficiency-adjusted. For instance, suppose a US-based researcher commands $200,000 per year while a Chinese researcher receives only $50,000 per year. However, as Section 1.2.1 details, the US-based researcher is superior in R&D. Thus, it could take the Chinese researcher four or five times longer to accomplish the same task or generate the same level of quality output as the US-based researcher. Hence, adjusting for the Chinese researcher's efficiency, his wage is $200,000 to 250,000 per year, which is more expensive than his US counterpart's. I assume that post-FDI, the Northern research subsidiary assumes R&D responsibility for products sold in the home and the export market. The relocation of R&D thus reduces the cost of quality upgrading. I also assume that manufacturing stays in the South so that the cost of unskilled labor that the firm faces remains the same.

I derive two propositions.

Comparative Statics

Proposition 1 *φ_d^{min}, φ_e^{min}, and φ_i^{min} are the productivity cut-off point for domestic sales, exports, and R&D FDI of the Southern firm post-FDI liberalization, respectively. When $\frac{f_X}{f_D} > \left(\frac{\tau_{sn}^{1-\sigma} B_n}{B_s}\right)^{\frac{\alpha}{\kappa}}$ and $\frac{f_I}{f_X} > \left[\left(\frac{B_s}{\tau_{sn}^{1-\sigma} B_n}\right)^{\frac{\alpha}{\kappa}} + 1\right]\left(\frac{w_s^h}{w_n^h}\right)^{\frac{\alpha-\kappa}{\kappa}}$, we have $\varphi_d^{min} < \varphi_e^{min} < \varphi_i^{min}$.*

Proof. See Appendix A.1.

Proposition 1 implies that post-FDI liberalization, the most productive Southern firms engage in R&D FDI while continuing to sell their products in the domestic and export markets. The next most productive firms also sell goods in both markets but do not engage in R&D FDI. The third most productive firms only sell domestically, while the least productive firms exit the sector altogether. The intuition is this: only firms with sufficiently high productivity find it profitable to become multinationals, and they alone can take advantage of the North's lower efficiency-adjust cost of skilled labor.

Proposition 2 *For a Southern firm with $\varphi > \varphi_i^{min}$ who engages in R&D FDI, the quality of its exports to the North increases post-FDI.*

Proof. From Equation 1.3, we can express the ratio of a Southern firm's export quality before and after FDI as:

$$
\begin{aligned}
\frac{q_{sn}^{*\text{before}}}{q_{sn}^{*\text{post}}} &= \frac{\left[B_n\left(\frac{1-\beta}{\alpha}\right)\left(\frac{\sigma-1}{\sigma}\right)^{\sigma}\left(\frac{\tau_{sn} w_s^l}{\varphi}\right)^{1-\sigma}\frac{1}{w_s^h}\right]^{\frac{1}{\kappa}}}{\left[B_n\left(\frac{1-\beta}{\alpha}\right)\left(\frac{\sigma-1}{\sigma}\right)^{\sigma}\left(\frac{\tau_{sn} w_s^l}{\varphi}\right)^{1-\sigma}\frac{1}{w_n^h}\right]^{\frac{1}{\kappa}}} \\
&= \left(\frac{w_n^h}{w_s^h}\right)^{\frac{1}{\kappa}} < 1
\end{aligned}
\tag{1.7}
$$

Therefore $q_{sn}^{*\text{before}} < q_{sn}^{*\text{post}}$, given that $w_n^h < w_s^h$ and $\kappa > 0$.

In the following section, I test Proposition 2 empirically.

1.5 Empirical Methodology

In this section, I discuss the setup and rationale of the empirical model, how I address self-selection and other endogeneity concerns, and the process for estimating the key regression variables.

1.5.1 Baseline Estimation

To empirically test the effect of R&D FDI on the export quality of Chinese firms engaged in such investments, which Proposition 2 predicts to be positive, I estimate the following model in levels:

$$\ln(q_{fhct}) = \beta_0 + \beta_1 \text{FDI}_{ft} + \vartheta' \boldsymbol{Z}_{ft} + \iota HHI_{st} + \delta_{fhc} + \delta_{ht} + \delta_{ct} + \delta_{pt} + \varepsilon_{fhct} \tag{1.8}$$

In this equation, the dependent variable $\ln(q_{fhct})$ is the inferred product quality of firm f exporting HS6 product h to country c in year t. I estimate quality using the elasticity of substitution estimates from Broda and Weinstein (2006) (Section 1.5.2 details the estimation methodology). FDI_{ft} is the explanatory variable of interest, and it is equal to 1 if the firm has R&D FDI and 0 otherwise. For example, suppose a firm first establishes an R&D lab in Germany in 2009, then FDI_{ft} is 0 from 2000 to 2008 and 1 from 2009 to 2013. Consistent with the existing literature, I assume that the firm's foreign affiliates continue their operations following the initial investment and that there is no exit from FDI. The vector $\boldsymbol{Z}_{ft}$ comprises time-varying firm-level controls, including firm productivity, size, and age measures. HHI_{st} denotes the Herfindahl index at the 4-digit CIC level for each year t. δ_{fhc} is the firm-product-country fixed effects while $\delta_{ht}, \delta_{ct}, \delta_{pt}$ are the product-year, country-year, and province-year fixed effects, respectively. ε_{fhct} represents an i.i.d. error term.

Endogeneity Concerns

Firms do not engage in FDI randomly; instead, they self-select into such investments based on their productivity (Helpman et al., 2004) and other unobserved firm characteristics, for example, entrepreneurial spirit, management foresight, and technological advantages. Prior studies have firmly established that larger, more productive firms are more likely to undertake FDI. However, high-productivity firms are also more inclined to produce and export high-quality goods (Kugler & Verhoogen, 2012). Therefore, I include firm-product-country fixed effects and firm productivity measurements (Section 1.5.3) in my regression models to address these confounding effects. By doing so, I capture all firm-specific, time-invariant determinants of export quality, plus an essential time-varying determinant. Additionally, I account for the heterogeneity of quality demand across products and countries. The latter control is significant because firms often adjust the quality of their specific HS6 products based on the export destination country, shipping higher quality goods to wealthier countries due to a greater demand for quality in those markets (Bastos & Silva, 2010; Demir, 2011; Hallak, 2006; Manova & Zhang, 2012; Verhoogen, 2008).

In addition to concerns posed by firm self-selection, the trade literature on quality also highlights the impact of input trade liberalization on export quality. Several studies focusing on developing countries find that tariff reductions prompt firms to upgrade the quality of their exports, either because these firms now gain better access to high-quality inputs (Bas & Strauss-Kahn, 2015; Fan et al., 2015; Feng et al., 2016; Xu & Mao, 2018) or because they face heightened competition from imported goods (Fernandes & Paunov, 2013; Martin & Mejean, 2014; Medina, 2022). Ideally, one would construct some firm-specific measures of tariff changes as control.[24] However, implementing this method requires knowledge of the value share of inputs within the bundle of intermediates actually imported by the firms, which is

[24] Controlling for firm-specific tariff changes compares the export quality of FDI firms to that of non-FDI firms facing similar effective tariff reductions over time. Since the independent variable, FDI, is also firm-year varying, having firm-year fixed effects in the regression is out of the question. Thus, the next best thing to do is to use firm-specific tariff changes.

only available for a fraction of exporters who are also importers. Not to mention, this method loses the extensive margin effects because firms may introduce new inputs and discontinue using old ones in their production processes over time. Alternatively, one could compute industry-level tariffs using input-output tables. However, such measures are relatively crude and are not firm-specific.[25] Therefore, I adopt a viable workaround by assuming no significant variation in the input mix among Chinese firms exporting the same narrowly defined HS6 product. Under this assumption, controlling for the HS6 product-year fixed effects would be sufficient to account for the quality-enhancing effects of trade liberalization. At a minimum, this approach offers better resolution than using the industry-level measures of tariff changes alone.

Furthermore, I incorporate destination country-year fixed effects to capture changes in the broader macroeconomic environment, such as in exchange rates, GDP, and tariffs levied on imports from China. These factors shift global demand and can impact the export quality of Chinese firms. Additionally, I include province-year fixed effects to capture changes in regional economic conditions within China, such as local GDP growth, infrastructure development, and policy changes at the provincial level. These fixed effects further reduce the probability that my results are driven by time-varying shocks other than FDI.

In summary, my identification strategy exploits the within firm-product-country quality variation over time associated with changes in a firm's FDI status and differences across firms that have undertaken FDI and those that have not. Because it is the firms that make the FDI decisions, I cluster standard errors at the firm level.

1.5.2 Quality Estimation

Product quality is not observable in the available trade data and, therefore, must be inferred. Khandelwal et al. (2013) propose a practical method to estimate quality, which has gained

[25]The Chinese Input-Output tables from the National Bureau of Statistics of China have 122 CIC industries, which correspond to n HS-2 digit industries. Hence, the industry-level tariffs only vary at the HS2-year level.

widespread adoption in papers such as Fan, Li, and Yeaple (2018) and Fan et al. (2015) and Bas and Strauss-Kahn (2015), who apply the original technique to the Chinese trade-transactions data. Following their approach, I first log-transform Equation 1.2, the demand function, rearrange, and then estimate the following equation using OLS:

$$\ln(x_{fhct}) + \sigma_i \ln(p_{fhct}) = \eta_h + \eta_{ct} + \upsilon_{fhct} \tag{1.9}$$

Here, x_{fhct} is the quantity demand for firm f's export of HS6 product h to the country c in year t, p_{fhct} is the unit value, and υ_{fhct} is the residual. The country-year fixed effects η_{ct} capture the price index and income of the destination country, while the product fixed effects η_h control for the inherent differences across products. I use Broda and Weinstein (2006)'s estimates of the elasticity of substitution[26] and estimate Equation 1.9 on Chinese Customs data from 2000 to 2013. The estimated residual is interpreted as the firm-product-country-year level effective quality: $\hat{\upsilon}_{fhct} \equiv \ln(\hat{q}_{fhct})$.[27] The intuition is straightforward: conditional on price, higher-quality products sell more units and therefore have a higher residual, $\hat{\upsilon}_{fhct}$.

1.5.3 Productivity Measures

In Section 1.5.1, I outlined the impact of time-varying firm productivity on both FDI and export quality. To mitigate the resulting selection bias, I construct three measures of firm productivity to include as controls in the regression. First, for 2000-2007, I estimate two measures of total factor productivity (TFP) using the Olley and Pakes (1996) method (OP) and the Levinsohn and Petrin (2003) method (LP), augmenting each method with the approach of Ackerberg et al. (2015). In both methods, value-added serves as the production output. Next, I deflate the firms' material inputs and gross outputs using the sector-level price deflators provided by Brandt et al. (2012, 2017). Finally, following Brandt et al. (2012),

[26] I convert Broda and Weinstein (2006)'s elasticity estimates in SITC Rev 3 into HS1996 at the 6-digit level. Then, I aggregate them to the HS 2-digit level by taking the arithmetic mean.

[27] Here, I follow Fan et al. (2015) and define $\hat{q}_{fhct} \equiv q_{fhct}^{\sigma-1}$, the quality that enters the demand function.

I calculate real capital and real investment by adopting the perpetual inventory method: $I_{ft} = K_{ft} - (1 - \delta_{ft})K_{ft-1}$, where δ_{ft} is each firm f's actual depreciation rate calculated from the NBSC Survey data. Second, for 2000-2013, I compute the log of revenue per worker, where revenue is also deflated using Brandt et al. (2012, 2017)'s output deflators. Due to missing material inputs and value-added in the ASIF after 2007, TFP estimation is only possible for 2000-2007. Consequently, in the baseline regression, I use the log of deflated revenue per worker as a control; in the robustness check on an alternative sample period of 2000-2007, I use the two TFP measures as controls (Section 1.6.3).

1.6 Empirical Results

1.6.1 The Baseline Results

To test Proposition 2 (Section 1.4.3), I regress the log of inferred quality on an R&D FDI dummy variable. I control for a range of covariates and fixed effects that are potentially significant determinants of firm export quality. Table 1.4 presents the baseline regression results.

In Column (1), I report the result of a bivariate ordinary least squares regression. The point estimate of the FDI coefficient is large, at 5.999, and highly significant statistically, at 1%. Thus, even at the most cursory level, we see evidence that FDI positively correlates with higher export quality.

However, a firm's self-selection into FDI by productivity (more productive firms are more likely to conduct FDI) could bias the estimates in Column (1), most likely upwards by a significant amount. Furthermore, changes in the macroeconomic environment between 2000 and 2013, such as trade liberalization and GDP growth in China and the rest of the world, could influence export quality.

To address these concerns, I add the firm-product-country fixed effects, product-year fixed effects, country-year fixed effects, and province-year fixed effects into the regression. Section

1.5.1 discusses the justification of these fixed effects. Column (2) shows the result with these additions. As expected, the beta point estimate decreases considerably from 5.999 to 0.623, almost a 90% reduction. Despite this, it remains significant at the 5% level. The massive drop in coefficient confirms a priori suspicion: untreated self-selection and endogeneity are indeed large, significant, and upward biasing.

While the aforementioned fixed effects can control for the time-invariant firm, product, and country factors and time-varying macroeconomic trends, they cannot control for time-varying firm-level factors. The literature offers many such factors to consider. I choose three variables: firm productivity, size, and age. According to the literature, these variables strongly affect export quality and a firm's probability of engaging in FDI. The maximum correlation coefficient among these variables is 0.264, between firm size and age, alleviating any fear of multicollinearity.

In Column (3), the size of the FDI coefficient drops again, from 0.623 to 0.525, while the robust standard error increases from 0.286 to 0.293. The coefficient remains statistically significant, albeit at only 10%. Note that the coefficients for log sales per worker and firm size are positive and highly significant at 1%, in line with Proposition 1 and other overwhelming findings in the literature: larger and more productive firms enjoy higher export quality. Firm age does not affect quality.

Lastly, I add the Herfindahl index (HHI) to account for industry competition. The HHI is computed at the 4-digit CIC level for each year t. Note that the coefficient and robust standard error estimates in Column (4) are almost identical to those in Column (3), suggesting that market power does not affect my results.

Column (4) is the preferred baseline result. According to the point estimate of the FDI coefficient in this Column, firms who conduct R&D FDI will experience, on average, a $(e^{0.525} - 1) \times 100 = 69.05\%$ increase in export quality post-investment over time. This change occurs within their existing export varieties, defined as a specific HS6 product and destination country combination. The magnitude of the increase is comparable to other findings in the quality-upgrading literature. For example, Bas and Strauss-Kahn (2015) report a 12.3% increase in firm export quality following tariff reductions in China. Fan et al. (2015), who also studies China's trade liberalization using the same micro-level firm data as Bas and Strauss-Kahn (2015), find that the percentage increase in export quality can be as high as 300% over five years.

In summary, R&D FDI is associated with increased export quality post-investment, supporting Proposition 2. The results are robust to controlling for a wide range of fixed effects, firm characteristics, and industry-level variables. Nonetheless, several concerns remain. First, the significance level for the main result is relatively weak, at only 10%. Second, out of necessity, I use labor productivity instead of TFP as the productivity control, but labor productivity may not be an adequate control. Third, other confounding factors could still drive the positive association between FDI and export quality. In the following sections, I address these concerns.

1.6.2 Alternative Specifications and Evidence on Extensive Margin Adjustment

The baseline specification is chosen to be consistent with my theoretical model. Abstracting from the firm's decision to enter and exit a country and HS6 product, Proposition 2 predicts that when a firm adjusts the quality of its exports post FDI, it does so for its existing products and export markets. In other words, the adjustment occurs along the export intensive margin rather than the extensive margin. Therefore, I use firm-product-country fixed effects.

However, focusing on such a granular level of observations inevitably leads to a loss of statistical power. This loss occurs because the firm-product-country fixed effects require that the same firm-product-country combination have non-zero export values for at least two years, one year before and one year after the FDI.[28] On the other hand, firms and

[28] Among the treated firms, only 23.51% of the firm-product-country level observations satisfy the two-period rule. This figure rises to 41.95% for firm-product level observations and to 53.19% for firm-level observations.

policymakers are more interested in understanding how FDI affects the firm as a whole rather than just examining the impact on the firm's existing markets and product lines. Therefore, I consider two modified versions of the baseline regression:

$$\ln(q_{fhct}) = \beta_0 + \beta_1 \mathrm{FDI}_{ft} + \vartheta' \boldsymbol{Z}_{ft} + \iota HHI_{st} + \delta_{f(h)} + \delta_{hc} + \delta_{ht} + \delta_{ct} + \delta_{pt} + \varepsilon_{fhct} \tag{1.10}$$

The first version includes firm-product fixed effects, and the second uses only firm fixed effects. Both versions add product-country fixed effects to account for the heterogeneity in quality demand across countries for the same HS6 product. Previously, the firm-product-country fixed effects absorbed this term. The other time-varying fixed effects and the firm and industry-level controls are the same as in the baseline to ensure a comparable estimation.

One can glean two compelling findings from Table 1.5. First, all coefficients on FDI under the two alternative specifications are positive and statistically significant at 5% (Panel A and B, Columns (1) to (4)). Together with Table 1.4, these results provide robust support for Proposition 2.

Second, as we move from Table 1.4 to Panel A and Panel B of Table 1.5, the magnitude of the R&D FDI coefficient increases significantly, from 0.525 to 0.973 and then to 1.938 (Column (4)). This increase occurs as the empiric specification relaxes from examining the within firm-product-country changes to within firm-product, and finally to within firm, therefore suggesting that extensive margin adjustments play a prominent role in driving firm-level quality improvements and that firms not only upgrade the quality of their existing products sold in existing markets but also shift the composition of their exports toward higher quality products and the countries that demand them the most.

1.6.3 Alternative Productivity Measures

So far, I have used only labor productivity, as measured by the log of deflated sales per worker, as the productivity control in regressions conducted over the entire sample period from 2000 to 2013—labor productivity being the only productivity measure available for this period (Section 1.5.3). Nonetheless, this choice of control is far from ideal. Labor productivity can change even if the underlying production technology does not, for example, if the capital-labor ratio changes. Therefore, the literature always prefers TFP when studying technological change and firm performance (Brandt et al., 2012, 2017; Yu, 2015).

To ensure that the choice of productivity measure does not affect the main results, I perform a robustness check using the Olley and Pakes (OP) and Levinsohn and Petrin (LP) TFP estimates as controls, but for a restricted sample period from 2000 to 2007, for which TFP estimation is feasible. I run the robustness check using the three specifications discussed above. The results are as follows.

First, in Panels A and B of Table 1.6, the estimated coefficients on FDI are no longer statistically significant, in contrast to the positive and significant results in Table 1.4 and Table 1.5, Panel A. However, this lack of significance is probably not due to the use of TFP. Comparing the numbers in Columns (3) and (4), Panel A and Panel B, where OP TFP and LP TFP are the productivity controls, with the number in Column (2), where log sales per worker is the control, or even with Column (1), where there is no productivity control at all, one sees that both the magnitude and the standard errors of the coefficients remain essentially unchanged. That is, the choice of productivity measures has minimal effect on the outcome.

How, then, should we interpret the insignificance? Most likely, the insignificance is due to the smaller number of observations in the restricted sample: only about 6.5% of all FDI firms carried out FDI before 2007 (Table 1.3). Thus, by relaxing the fixed effects to within-firm, one regains some statistical power: in Panel C of Table 1.6, the coefficients of R&D FDI using all three productivity measures are consistent at around 2.3 - 2.4 and significant at 5%, again in line with Proposition 2.

In conclusion, restricting the sample period to 2000-2007 significantly affects the regression results because fewer firms engage in FDI during this period. However, the evidence suggests that the main results are not sensitive to various productivity controls, which increases confidence in the baseline results.

1.6.4 Placebo Test

To evaluate the reasonableness of R&D FDI as a quality-enhancing mechanism, I conduct a placebo test using non-R&D FDI as the independent variable of interest instead of R&D FDI. The selected non-R&D FDIs are Distribution, Marketing, Production, Services, and Agriculture FDI. In addition, this section aims to address any remaining endogeneity concerns.

Unlike R&D FDI, which firms pursue to gain access to the host country's skilled labor pool and technological infrastructure (Section 1.3.1), non-R&D FDIs are pursued for very different reasons. For example, Fan, Lin, and Tang (2018) find that Chinese firms engage in Production FDI to escape rising labor costs at home. Tian and Yu (2020) show that high communication costs between international buyers and sellers lead Chinese firms to pursue Marketing and Distribution FDIs. Moreover, my data on FDI projects demonstrates that neither Services nor Agriculture FDI is undertaken to promote exports: Services FDI are firms investing in the financial, real estate, and consulting sectors, while Agriculture FDI goes to land-rich regions for farming, fishing, and forestry activities (Section 1.3.1). Since firms do not seek any of these five non-R&D FDIs to enhance R&D capabilities, they should have little or no effect on export quality. Therefore, their estimated coefficients should be insignificant. However, if the coefficients are significant, this would indicate a deeper problem. In such a case, missing variables may be responsible for the positive correlation between R&D FDI and export quality, as shown in Tables 1.4, 1.5, and we likely have unaddressed endogeneity at hand.

Table 1.7 presents the results of the placebo test. All columns in each panel include the complete set of fixed effects and control variables, as is in the preferred baseline model. I find no significant associations between the non-R&D FDIs and export quality, either positive or negative. These results are consistent with the expected outcome of the placebo test. In short, the results in Table 1.7 validate the R&D FDI quality-enhancing channel and rule out other potential confounders driving the baseline result.

1.6.5 Heterogeneity by Quality Differentiation

Research has shown that firms producing products with a higher degree of product differentiation are more likely to invest in improving their quality (Eckel et al., 2015). While the model of this paper assumes a single-product firm, in practice, almost all exporters are multi-product firms. Therefore, I examine whether the effect of R&D FDI differs by the scope for product differentiation.

Following the literature, I use four measures of product differentiation. First, I use the Rauch (1999) classification, which divides the four-digit Standard International Trade Classification (SITC) Revision 2 codes into three categories, "goods traded on an organized exchange", "reference priced", and "differentiated products". I combine the first two categories into a "homogeneous" category and convert the SITC codes to HS1996, the coding of my export data. For completeness, I adopt both the "liberal" and "conservative" classifications in Rauch (1999), where "liberal" has fewer products classified as " homogeneous". Second, I use the R&D and Advertising Intensity and the Modified Gollop-Monahan (G-M) index. Data for both measures are taken from Kugler and Verhoogen (2012). The R&D and Advertising Intensity is the ratio of advertising expenditure plus R&D expenditure to total sales reported in the US Federal Trade Commission (FTC) 1975 Line of Business Survey, and the G-M index measures the dissimilarity of inputs within a sector, a proxy for product differentiation used by papers such as Kugler and Verhoogen (2012) and Fan et al. (2015). Lastly, I follow Fan et al. (2015) and construct a dummy variable indicating quality dispersion: I compute the quality variance for each HS6 product, and so a product is considered highly dispersed if its quality variance is above the median quality variance of all HS6 products across all sample periods ("All periods") or all products in the year prior

("Lagged $t - 1$"). For each measure of product differentiation, I expect the coefficients of differentiated products to be greater than those of homogeneous products.

Tables 1.8 through 1.10 present the results. While none of the coefficients on the interaction terms are significant (except for "Lagged $t-1$"), they are still larger in magnitude than the coefficients on the R&D FDI variable. When I divide the sample into two subgroups, the coefficients associated with products with a higher degree of product differentiation, defined by Rauch classifications and quality dispersion, are much larger and statistically significant than those associated with homogeneous products. I repeat the regressions of Tables 1.8 through 1.10 using the alternative specifications described in Section 1.6.2. The results are similar to those in Tables 1.8 through 1.10 (see Tables A3 through A8).

1.7 Conclusion

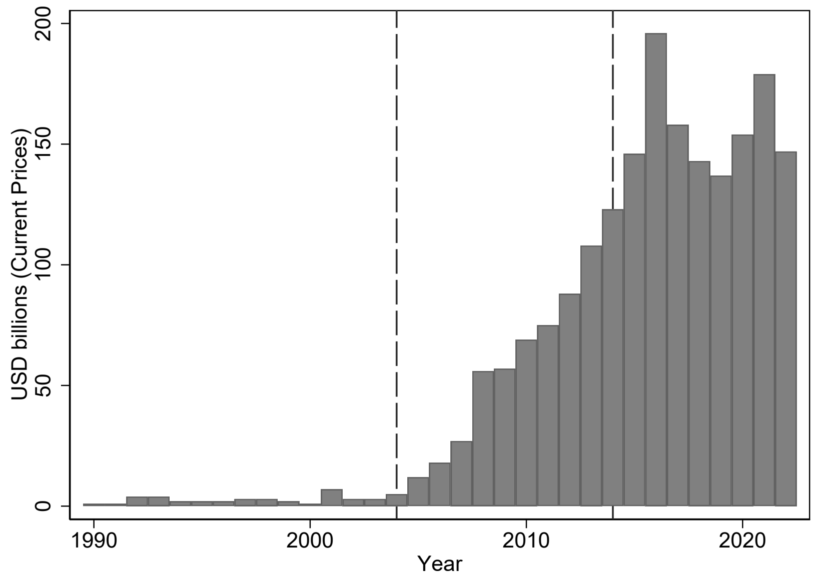
In this paper, I uncover a positive relationship between R&D FDI and export quality at the firm level utilizing Chinese microdata. I overcame a critical data limitation that restricted much of the prior research on Chinese outbound FDI, namely, the inability to identify the objectives of each FDI project reliably. I apply supervised machine learning to the MOFCOM dataset and classify all non-tax haven FDI projects into eight categories (Table 1.2). I focus on the R&D FDI as the variable of interest. Gathering insights from several business case studies, I develop a theoretical framework with heterogeneous firms and endogenous quality that features a specific quality upgrading mechanism—recruiting offshore experts with superior innovation capability to carry out R&D. Testing the model's prediction empirically, I find strong evidence that R&D FDI is positively associated with export quality. This finding is robust to alternative specifications and firm productivity controls. Additionally, the positive correlation is more pronounced in sectors where the scope for product differentiation is larger.

An important caveat of this paper is that the uncovered correlation does not explain the whole of China's quality upgrading over the last thirty years. Other sweeping changes, such as privatization in the 1990s and trade liberalization in the early 2000s (China joined the WTO in 2001), would likely have impacted China's overall export quality considerably more. Nevertheless, as China moves up the quality ladder, achieving additional marginal improvements becomes ever more challenging. In this regard, examining the potential impact of R&D FDI on export quality is particularly meaningful because most R&D FDI happened after 2008, when trade liberalization had already significantly boosted quality in the preceding years. The outcomes of this study hold significant policy implications for developing and developed countries. For developing countries, the findings suggest that incentivizing outbound FDI could be a viable strategy for promoting economic development. For developed countries, the findings raise the question of whether or not to restrict inward FDI to safeguard national interests, a decision that could have far-reaching economic consequences.

While this paper has made ways to understand the relationship between R&D FDI and export quality, some areas still require further exploration. For one, this paper focuses on quality adjustments along the firm-product-country intensive margin but dedicates little attention to adjustments along the extensive margin, i.e., how firms enter and exit countries and product categories. However, extensive margin adjustments are crucial for evaluating the competitive pressure Chinese exports exert on incumbent players. For example, is R&D FDI enabling China to break into high-value product segments previously beyond its capability? Is China's export market share increasing as a result? These are questions that researchers may find fruitful to explore in the future.

Figures

Figure 1.1: Chinese FDI Outflow and Policy Reform



Data: UNCTAD

Phases of Reforms:
1. Pre-1990 to 2003: Outbound FDI restricted
2. 2004 to 2013: Progressively liberalize
3. Post 2014: Comprehensive liberal policy in place

Tables

Table 1.1: Keywords of FDI Projects

	Keywords
R&D	Research, development, patent, design, laboratory
Agriculture	Cultivation, farming, breeding, fishing, irrigation, animal husbandry
Infrastructure	Civic engineering, communications, construction, infrastructure, contracting, bridges, roads, electrical, water supply, municipal government
Distribution	Wholesale, retail, sales, trading
Marketing	Market research, information collection, promotion, business development, order processing, post-sales support, advertising
Production	Production, manufacturing, processing, assembling
Resource	Mining, excavation, logging, drilling, oil, gas, minearls
Services	Finance, real estate, restaurants, travel, publishing, consulting

Table 1.2: Number of FDI Projects by Machine Learning, 1987-2015

	#Project	%Total
R&D	2212	8.4
Agriculture	1624	6.1
Infrastructure	2798	10.6
Distribution	5302	20.0
Marketing	3779	14.3
Production	4357	16.5
Resource	2257	8.5
Services	4154	15.7
Total	26483	100.0

Note: The sample period is 1987 to 2015. FDI to tax havens excluded. All remaining projects are classified into eight categories using text classification, supervised machine learning.

Table 1.3: Chinese R&D FDI by Year, 1999-2015

	Project		Firm	
	Number	%Total	Number	%Total
1999	1	0.0	1	0.0
2000	2	0.1	2	0.1
2002	4	0.2	4	0.2
2003	3	0.1	3	0.1
2004	6	0.3	6	0.3
2005	31	1.4	31	1.5
2006	37	1.7	36	1.8
2007	48	2.2	48	2.4
2008	52	2.4	50	2.4
2009	106	4.8	99	4.9
2010	140	6.3	129	6.3
2011	182	8.2	169	8.3
2012	187	8.5	177	8.7
2013	256	11.6	241	11.8
2014	361	16.3	335	16.4
2015	796	36.0	710	34.8
Total	2212	100.0	2041	100.0

Notes: Columns (1)-(2) concern R&D FDI projects, and columns (3)-(4) concern the firms which conducted them. FDI to tax havens excluded.

Table 1.4: Basic Results: Effect of R&D FDI on Export Quality, 2000-2013

	(1)	(2)	(3)	(4)
	$\ln(q_{fhct})$	$\ln(q_{fhct})$	$\ln(q_{fhct})$	$\ln(q_{fhct})$
R&D FDI	5.999***	0.623**	0.525*	0.525*
	(1.552)	(0.286)	(0.293)	(0.293)
Log Sales per worker			0.462***	0.463***
			(0.0570)	(0.0570)
Log Number of workers			0.442***	0.443***
			(0.0758)	(0.0760)
Age			0.00464	0.00459
			(0.00475)	(0.00474)
HHI				-1.870
				(1.752)
Firm-Product-Country FE	No	Yes	Yes	Yes
Product-Year FE	No	Yes	Yes	Yes
Country-Year FE	No	Yes	Yes	Yes
Province-Year FE	No	Yes	Yes	Yes
Observations	12944206	9668235	9668235	9668235
R-Squared	0.000122	0.759	0.759	0.759

Notes: Robust standard errors clustered at the firm level are reported in parentheses. Significance: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 1.5: Alternative Empirical Specifications

	(1)	(2)	(3)	(4)
	$\ln(q_{fhct})$	$\ln(q_{fhct})$	$\ln(q_{fhct})$	$\ln(q_{fhct})$
Panel A: Firm-Product FE				
R&D FDI	5.999***	1.063**	0.973**	0.973**
	(1.552)	(0.459)	(0.472)	(0.473)
Log Sales per worker			0.366***	0.366***
			(0.0563)	(0.0563)
Log Number of workers			0.335***	0.336***
			(0.0760)	(0.0761)
Age			0.00618	0.00614
			(0.00420)	(0.00419)
HHI				-1.381
				(1.726)
Firm-Product FE	No	Yes	Yes	Yes
Product-Country FE	No	Yes	Yes	Yes
Product-Year FE	No	Yes	Yes	Yes
Country-Year FE	No	Yes	Yes	Yes
Province-Year FE	No	Yes	Yes	Yes
Observations	12944206	12181988	12181988	12181988
R-Squared	0.000122	0.645	0.645	0.645
Panel B: Firm FE				
R&D FDI	5.999***	2.051**	1.940**	1.938**
	(1.552)	(0.867)	(0.900)	(0.901)
Log Sales per worker			0.313***	0.312***
			(0.0766)	(0.0767)
Log Number of workers			0.230**	0.229**
			(0.102)	(0.102)
Age			-0.00143	-0.00139
			(0.00540)	(0.00541)
HHI				2.540
				(2.863)
Firm FE	No	Yes	Yes	Yes
Product-Country FE	No	Yes	Yes	Yes
Product-Year FE	No	Yes	Yes	Yes
Country-Year FE	No	Yes	Yes	Yes
Province-Year FE	No	Yes	Yes	Yes
Observations	12944206	12866539	12866539	12866539
R-Squared	0.000122	0.374	0.374	0.374

Notes: Robust standard errors clustered at the firm level are reported in parentheses. Significance: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 1.6: Alternative Productivity Measures and with Restricted Sample 2000-2007

	(1) $\ln(q_{fhct})$	(2) $\ln(q_{fhct})$	(3) $\ln(q_{fhct})$	(4) $\ln(q_{fhct})$
Panel A: Firm-Product-Country FE				
R&D FDI	-0.248	-0.220	-0.225	-0.228
	(0.647)	(0.647)	(0.665)	(0.665)
Log Sales per worker		0.619***		
		(0.0573)		
ln TFP^{OP}			0.217***	
			(0.0353)	
ln TFP^{LP}				0.229***
				(0.0350)
Observations	3411911	3411911	3341420	3341420
R-Squared	0.866	0.866	0.866	0.866
Panel B: Firm-Product FE				
R&D FDI	0.641	0.650	0.688	0.684
	(0.509)	(0.507)	(0.521)	(0.521)
Log Sales per worker		0.501***		
		(0.0626)		
ln TFP^{OP}			0.204***	
			(0.0387)	
ln TFP^{LP}				0.214***
				(0.0383)
Observations	4734406	4734406	4653472	4653472
R-Squared	0.734	0.734	0.733	0.733
Panel C: Firm FE				
R&D FDI	2.354**	2.330**	2.487**	2.489**
	(1.027)	(1.023)	(1.066)	(1.066)
Log Sales per worker		0.380***		
		(0.0825)		
ln TFP^{OP}			0.180***	
			(0.0518)	
ln TFP^{LP}				0.204***
				(0.0534)
Observations	5138303	5138303	5053375	5053375
R-Squared	0.431	0.431	0.431	0.431
Panels A, B, and C:				
Fixed Effects	Yes	Yes	Yes	Yes
Firm-level Controls	No	Yes	Yes	Yes
Competition Control	No	Yes	Yes	Yes

Notes: Significance: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors clustered at the firm level are reported in parentheses. All regressions include fixed effects. Additional firm-level controls include firm size (Log Number of workers) and age. Competition control refers to the Herfindahl index (HHI) at 4-digit CIC for each year.

Table 1.7: Placebo: Results with Non-R&D FDI

	(1) $\ln(q_{fhct})$	(2) $\ln(q_{fhct})$	(3) $\ln(q_{fhct})$	(4) $\ln(q_{fhct})$	(5) $\ln(q_{fhct})$
Panel A: Firm-Product-Country FE					
DISTR	0.151				
	(0.557)				
MARKET		0.834			
		(0.862)			
PROD			-0.819		
			(0.890)		
SERV				0.220	
				(0.577)	
AGR					0.419
					(0.697)
Observations	9668235	9668235	9668235	9668235	9668235
R-Squared	0.759	0.759	0.759	0.759	0.759
Panel B: Firm-Product FE					
DISTR	-0.0411				
	(0.545)				
MARKET		0.854			
		(0.916)			
PROD			-0.575		
			(0.739)		
SERV				-1.068	
				(0.807)	
AGR					0.468
					(0.566)
Observations	12181988	12181988	12181988	12181988	12181988
R-Squared	0.645	0.645	0.645	0.645	0.645
Panel C: Firm FE					
DISTR	-0.00103				
	(0.699)				
MARKET		1.044			
		(1.276)			
PROD			-0.235		
			(0.591)		
SERV				-2.782	
				(1.892)	
AGR					1.163
					(0.742)
Observations	12866539	12866539	12866539	12866539	12866539
R-Squared	0.374	0.374	0.374	0.374	0.374
Panels A, B, and C:					
Fixed Effects	Yes	Yes	Yes	Yes	Yes
Firm-level Controls	Yes	Yes	Yes	Yes	Yes
Competition Control	Yes	Yes	Yes	Yes	Yes

Notes: Significance: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors clustered at the firm level are reported in parentheses. All regressions include fixed effects, firm-level controls, and competition control. Firm-level controls include productivity (Log Sales per worker), firm size (Log Number of workers) and age. Competition control refers to the Herfindahl index (HHI) at 4-digit CIC for each year.

Table 1.8: Effect of R&D FDI and Quality Differentiation, Rauch

	Liberal			Conservative		
	(1) Homo	(2) Differ	(3) Full	(4) Homo	(5) Differ	(6) Full
R&D FDI	0.00536	0.620*	0.187	-0.0994	0.597*	0.108
	(0.279)	(0.342)	(0.221)	(0.277)	(0.326)	(0.198)
R&D FDI × Differentiated^{lib}			0.395			
			(0.402)			
R&D FDI × Differentiated^{con}						0.465
						(0.378)
Firm-level Controls	Yes	Yes	Yes	Yes	Yes	Yes
Competition Control	Yes	Yes	Yes	Yes	Yes	Yes
Firm-Product-Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Product-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Country-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Province-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1202674	8465373	9668235	897041	8770987	9668235
R-Squared	0.816	0.759	0.759	0.802	0.759	0.759

Notes: Robust standard errors clustered at the firm level are reported in parentheses. Significance: *** p<0.01, ** p<0.05, * p<0.1.

Table 1.9: Effect of R&D FDI and Quality Differentiation, Gollop-Monahan (G-M) Index

	R&D Expenditure			G-M Indices		
	(1) <Median	(2) >Median	(3) Full	(4) <Median	(5) >Median	(6) Full
R&D FDI	-0.0553	1.065	0.0282	-0.163	0.649	-0.0627
	(0.154)	(0.698)	(0.143)	(0.385)	(0.460)	(0.334)
R&D FDI × R&D			0.730			
			(0.701)			
R&D FDI × GM						0.697
						(0.535)
Firm-level Controls	Yes	Yes	Yes	Yes	Yes	Yes
Competition Control	Yes	Yes	Yes	Yes	Yes	Yes
Firm-Product-Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Product-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Country-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Province-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5233795	3324936	8558863	1797501	6761218	8558863
R-Squared	0.811	0.730	0.734	0.784	0.732	0.734

Notes: Robust standard errors clustered at the firm level are reported in parentheses. Significance: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 1.10: Effect of R&D FDI and Quality Differentiation, Quality Dispersion

	All Periods			Lagged $t-1$		
	(1)	(2)	(3)	(4)	(5)	(6)
	<Median	>Median	Full	<Median	>Median	Full
R&D FDI	-0.0158	0.939**	0.165	0.0221	0.961**	0.498**
	(0.0824)	(0.458)	(0.133)	(0.0878)	(0.462)	(0.198)
R&D FDI × Dispersion^{All}			0.538			
			(0.465)			
R&D FDI × Dispersion^{Lagged}						0.0278
						(0.365)
Firm-level Controls	Yes	Yes	Yes	Yes	Yes	Yes
Competition Control	Yes	Yes	Yes	Yes	Yes	Yes
Firm-Product-Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Product-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Country-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Province-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4843025	4825069	9668235	4491137	4678337	9483779
R-Squared	0.806	0.759	0.759	0.809	0.759	0.759

Notes: Robust standard errors clustered at the firm level are reported in parentheses. Significance: *** p<0.01, ** p<0.05, * p<0.1.

Chapter 2

Trading the Old for the New: R&D FDI and Export Margin Expansion

2.1 Introduction

A small percentage of firms that export multiple products to multiple destinations dominate international trade (Bernard et al., 2012). Understanding the behavior of these firms can significantly improve our understanding of aggregate trade patterns. While numerous studies have examined how factors such as economic crises and market integration affect a firm's export extensive margins (Bernard et al., 2009; Berthou & Fontagné, 2013; Muûls, 2015), no research has yet explored the impact of outbound foreign direct investment (FDI), particularly R&D FDI originating from developing countries on the same outcome.

This paper fills this gap. In this context, R&D FDI refers to FDI aimed at establishing R&D facilities abroad for a firm's product development and innovation activities. Using extraordinarily rich micro-data from China that allows me to control for a firm's self-selection into FDI, I study how R&D FDI affects the firm's extensive margin. More specifically, I look at the number of products exported and the number of countries supplied. I focus on the extensive margin adjustment of surviving exporters while deferring the more complex issues of firm entry, exits, and export duration to future projects.[1]

Economic theory suggests that R&D FDI should positively affect a firm's export extensive margin. First, Chapter 1 of this dissertation finds that R&D FDI improves the quality of firms' exports. According to Bernard et al. (2011), firm profitability depends on two attributes: firm ability, which is specific to the firm, and consumer taste for a product, which is specific to the product-country pair. In this framework, firm ability can be interpreted as firm productivity, while consumer taste can be seen as a preference that depends on product quality. Firms incur a fixed cost to supply each product in every destination market. Therefore, as product quality improves, firms can generate higher variable profits to cover these fixed costs. Consequently, even at the same level of productivity, firms can export a wider range of products to more markets due to the enhanced product quality. Second, several empirical studies find that engaging in FDI of any type raises the productivity of Chinese firms (Chen & Tang, 2014; Cozza et al., 2015; Huang & Zhang, 2017). According to Bernard et al. (2011), increased productivity should expand both the product and country extensive margins. Third, FDI in R&D, as a form of R&D investment, enables firms to innovate and add new products (Klette & Kortum, 2004), thereby expanding the portfolio of products available for export.[2]

In this paper, I combine three datasets that provide comprehensive information on Chinese firms' operations, including sales, fixed assets, and the number of employees, as well as export transactions by firm, product, and destination country. With this data, I can calculate the number of products exported, the number of countries a firm serves, and the quality of exports. The third dataset is particularly unique as it contains detailed information on the objective or type of each outbound FDI project. Constructed in Chapter 1 of this dissertation, it allows researchers to identify whether a project is related to distribution,

[1]For instance, examining how long a product is exported before it is dropped.

[2]In this paper, I do not observe the exact channels through which extensive margin expansion occurs. To do so would require access to more data than I currently have available. For example, to test whether new products were added (the third channel), I would need product-level data not only on firms' exports but also on their domestic sales. Instead, my empirical exercise captures the net effect of R&D FDI on the extensive margins, encompassing all three channels.

natural resources, R&D, or other categories. Since different types of FDI are motivated by different considerations, distinguishing among these types enables me to specifically test the effect of R&D FDI while excluding the confounding effects of other types of FDI.

Empirically, I employ a two-way fixed effects specification, including firm productivity, size, age, province-year fixed effects, and controls for marketing and distribution FDI. Identification relies on exploiting the within-firm variation in the number of products and countries over time associated with changes in a firm's R&D FDI status, as well as the differences between firms that have undertaken R&D FDI and those that have not. My results show that R&D FDI significantly increases the number of products, defined at the Harmonized System 6-digit level (HS6), the number of destination countries, and the number of product-country pairs, in line with the theory's predictions discussed earlier. Specifically, R&D FDI increases the number of products by 31.7%, the number of countries by 13.3%, and the number of product-country pairs by 35.4%. To put these findings into perspective, a study of German manufacturing firms found that foreign-owned firms export 23% to 39% more products and ship to 11% to 31% more destinations than domestic firms (Raff & Wagner, 2014). Therefore, the impact of conducting R&D FDI is comparable to the substantial effects seen with changes in ownership.

Looking more closely at product quality, I find that new product-country entrants—defined as product-country pairs within a firm present after R&D FDI but not before—exhibit the highest average quality. Continuing product-country pairs, which are present in both before and after periods, rank second in quality, while exiting product-country pairs, present only in the before period, rank last. This quality ranking suggests that R&D FDI induces intra-firm product churning: firms discontinue their lowest-quality exports, maintain medium-quality exports, and leverage their highest-quality products to penetrate new markets.

The rest of this paper is organized as follows: In Section 2.2, I present the data used in my empirical analysis. In Section 2.3, I discuss the empirical methodology employed. In Section 2.4, I present the estimation results. Finally, Section 2.5 provides the conclusion.

2.2 Data

This section provides an overview of the three datasets I utilized in this paper and details the process of cleaning and merging them.

2.2.1 The General Administration of Customs Data

I employed data compiled by the General Administration of Customs, which encompasses all Chinese trade transactions from 2000 to 2013. Specifically, this dataset reports each cross-border trade flow according to the Harmonized System 8-digit (HS8) product classification, destination or origin country, shipment date, and customs regime.[3] It also includes the value[4] and quantity of each transaction. This detailed product and country information allows me to calculate the number of products exported and the number of countries served by each firm, which serve as the dependent variables in this paper. Additionally, the Customs data provides detailed identifiers for the firms involved in each transaction, including firm name, zip code, phone number, and contact person. I use this information to merge the trade data with the Annual Survey of Industrial Firms (ASIF) data.

When cleaning the Customs data, I follow the methodologies outlined in Fan et al. (2015) and Brandt et al. (2012). First, I exclude all intermediary firms[5] from the sample and aggregate the transactions from HS8 to HS6. Next, I concord the HS2002, HS2007, and HS2012 codes into HS1996 using correspondence tables from the United Nations Statistics Division. Finally, I deflate the export values using output deflators from Brandt et al. (2012), extending the series to cover up to the year 2013 using Stata codes provided by Brandt et al. (2017).

[3]China uses 20 different customs regimes to indicate the purpose of trade, e.g., ordinary trade, processing and assembling, and processing with imported materials.

[4]Values are reported as Free on Board (FOB) for exports and Cost, Insurance, and Freight (CIF) for imports.

[5]Intermediary firms are identified by the presence of terms such as “trading” and “importing and exporting” in their firm names. This paper searches for the Chinese characters “贸易” (mào yì), “外贸” (wài mào), “外经” (wài jīng), “进出口” (jìn chū kǒu), “经贸” (jīng mào), “工贸” (gōng mào), “科贸” (kē mào), and “边贸” (biān mào) in firm names.

2.2.2 The Annual Survey of Industrial Firms

I utilize the Annual Survey of Industrial Firms (ASIF) from the National Bureau of Statistics of China to gather firm characteristics. This dataset comprehensively surveys "above-scale" industrial firms, defined as firms in the mining, manufacturing, and public utilities sectors[6] that achieve annual sales above a specified threshold.[7] The ASIF comprises detailed information on over 100 firm-level variables. These include identification details such as firm ID, name, legal person, address, and industry classification. Additionally, the dataset provides extensive operational information like the number of employees, gross output, and value-added. It also includes key financial data from the three main accounting statements: the balance sheet, cash flow statement, and income statement.

Despite its comprehensive scope, the ASIF is not without its data quality challenges. First, the dataset lacks a consistent and unique identifier that would allow firms to be tracked over time, as firm IDs and names often change due to restructuring, mergers, and acquisitions. To address this issue, I employ the methodology developed by Brandt et al. (2012), who crafted a rigorous algorithm for matching firms across periods using firm ID, name, legal entity, address, and industry information. Second, the ASIF contains numerous noisy observations. To ensure the reliability of my analysis, I adhere to the standard procedures outlined by Cai and Liu (2009) and Yu (2015). Specifically, I exclude observations with key financial variables that are missing, zero, or negative. Additionally, I remove firms with fewer than eight employees and any data points that violate Generally Accepted Accounting Principles (GAAP).[8] Finally, I exclude intermediate firms because they do not

[6]The National Bureau of Statistics classifies 2-digit China Industrial Classification (CIC) codes 06 to 12 as mining, 13 to 43 as manufacturing, and 44 to 46 as public utilities, with manufacturing being the predominant category.

[7]The threshold for inclusion in the ASIF has changed several times. From 2000 to 2006, the dataset included all state-owned enterprises (SOEs) and non-state-owned enterprises (non-SOEs) with annual sales of 5 million RMB (approximately USD 687,000 by the 2023 exchange rate) or more. From 2007 onwards, ASIF began excluding smaller SOEs with sales under 5 million RMB. Subsequently, from 2011 onwards, only firms with annual sales of 20 million RMB (approximately USD 2.7 million) or more were included in the data.

[8]The GAAP rules are: 1. total assets must be greater than liquid assets; 2. total assets must be greater than total fixed assets; 3. total assets must be greater than the net value of the fixed assets; 4. the

produce anything in-house and focus solely on manufacturing firms.[9]

2.2.3 The Directory of Overseas Investment Enterprises

To accurately assess the impact of R&D FDI on the extensive margin, it is essential to understand each FDI project's type or objective. Firms engage in different types of FDI for different strategic reasons, each with potentially distinct effects on their operations. Thus, reliably identifying and differentiating these types of FDI is crucial. Only by doing so can I effectively disentangle and analyze the economic impacts of each type of FDI.

I utilize a novel dataset constructed in Chapter 1 to meet this data requirement. This dataset is derived from the Ministry of Commerce's (MOFCOM) Directory of Overseas Investment Enterprises, which provides a comprehensive list of FDI projects undertaken by Chinese firms between 1987 and 2015. The listed FDI projects predominantly consist of greenfield investments, with a few being mergers and acquisitions (M&A) deals. MOFCOM collects a range of information for each project, including the parent company's name, the overseas subsidiary's name, the country of investment, the date of project approval, and details on the scope of offshore business activities.[10] Utilizing supervised machine learning, I analyzed the business activity descriptions written in natural language for more than 26,000 FDI projects. I classified these projects into eight categories, among which R&D FDI accounted for 2,212 unique projects from 1987 to 2015.

2.2.4 Merging Data

I combine the above three datasets using a two-step matching process as outlined in Yu (2015). First, I match the ASIF and the Customs data by firm name and year. Subsequently,

establishment date must be valid (the opening month is between January and December).

[9]Only industries with a 2-digit China Industrial Classification (CIC) ranging from 13 to 43, inclusive of endpoints, are retained in the sample. For the years prior to 2003, several 4-digit CIC codes—1711, 1712, 1713, 1714, 1719, 2220, 3648, 3783, 4183, and 4280—are also excluded from the sample because they are special cases of services, as identified by Brandt et al. (2012).

[10]MOFCOM collects this information as part of a stringent and burdensome FDI approval and registration process. which was in place until 2014.

I refine the matching by year, zip code, and the last seven digits of the firm's phone number. After establishing a robust match between the ASIF and Customs datasets, I match the R&D FDI data with this merged ASIF-Customs dataset, using firm name and year as the criteria. The final sample comprises $12,944,206$ observations from 2000 to 2013, detailed at the firm-HS6 product-country-year level.

2.3 Empirical Methodology

2.3.1 Baseline Estimation

During my sample period from 2000 to 2013, multiple firms conducted R&D FDI, each initiating their investments in different years. To accommodate this staggered treatment structure, I employ a two-way fixed effect approach widely used in empirical studies of firms with varying treatment periods.

Equation 2.1 is my baseline specification:

$$\ln(N_{ft}) = \beta_0 + \beta_1 FDI_{ft} + \vartheta' \boldsymbol{Z}_{ft} + \iota HHI_{st} + \delta_f + \delta_t + \delta_{pt} + \varepsilon_{ft} \tag{2.1}$$

Here, f denotes the firm, t the year, s the 4-digit Chinese Industrial Classification, and p the province. The dependent variable, N_{ft}, can be one of the extensive margins per firm year: (a) the number of HS6 products, (b) the number of destination countries, and (c) the number of product-country pairs. The independent variable of interest, FDI_{ft}, equals 1 if the firm has an offshore R&D subsidiary and 0 otherwise. In line with existing literature, I assume that once FDI is initiated, it continues indefinitely; that is, there is no discontinuation of FDI.

The vector $\boldsymbol{Z}_{ft}$ includes time-varying firm-level controls such as firm productivity, size, and age. The Herfindahl-Hirschman Index HHI_{st} at the 4-digit CIC level for each year t measures industry concentration. The fixed effects δ_f and δ_t control for firm-specific and time-specific influences that might affect the extensive margins of exports. Lastly, the province-year fixed effects δ_{pt} account for changes in regional economic conditions that can affect all firms within a region, such as local GDP growth, infrastructure developments, and shifts in industrial policy. ε_{ft} is an error term, distributed i.i.d..

In summary, these control variables and fixed effects reduce the likelihood that the observed effects on the extensive margins are driven by extraneous time-varying shocks rather than by R&D FDI. The identification strategy capitalizes on the within-firm variation in the number of products and destination countries over time associated with changes in a firm's R&D FDI status, alongside the differences between firms that engage in R&D FDI and those that do not. Standard errors are clustered at the firm level.

2.3.2 Export Quality by Entry, Exit, and Continuing Status

In Chapter 1 of this dissertation, I examined how quality improvements occur along the firm-product-country intensive margin but did not address changes along the extensive margin. Therefore, a natural extension is to investigate the quality of new products and markets. Specifically, when a firm adds new products or enters new markets, what is the quality of these new entrants compared to the quality of existing exports? Understanding this can shed light on the compositional changes in a firm's exporting portfolio, revealing whether, in addition to promoting quality-upgrading in existing exports, R&D FDI also directs firms towards exporting higher-quality products to markets that demand such products.

To address this question, I adopt the approach used by Fan et al. (2015) to classify the different types of firm-product-country (*fhc*), firm-product (*fh*), and firm-country (*fc*) combinations based on their status in the pre-FDI and post-FDI periods. For *ever-treated* firms—those that have engaged in at least one R&D FDI during my sample period from 2000 to 2013—a combination is considered "continuing" if it exists in both pre-and post-FDI periods; "entry" if it appears only in the post-FDI period and not before; and "exit" if it is present only in the pre-FDI period.

I compare the mean quality across "entry", "continuing", and "exit' types for each fhc, fh, and fc combinations.

Quality Estimation

To infer product quality from trade data, which does not explicitly record it, I follow the methodology outlined by Khandelwal et al. (2013) and estimate Equation 2.2 using Ordinary Least Squares (OLS):

$$\ln(x_{fhct}) + \sigma_i \ln(p_{fhct}) = \eta_h + \eta_{ct} + \upsilon_{fhct} \tag{2.2}$$

In this equation, h represents the HS6 product, and c is the destination country. x_{fhct} denotes the physical quantity of exports and p_{fhct} the unit value. υ_{fhct} is the residual. The country-year fixed effects η_{ct} capture the macroeconomic attributes of the destination country, while the product fixed effects η_h control for the inherent differences across products. The elasticity of substitution σ_i is sourced from Broda and Weinstein (2006).[11] The estimated OLS residual serves as my measure of quality at the firm-product-country-year level: $\hat{\upsilon}_{fhct} \equiv \ln(\hat{q}_{fhct})$.[12] The intuition of Khandelwal et al. (2013)'s approach is straightforward: conditional on price, higher-quality products sell more units and therefore have a higher residual, $\hat{\upsilon}_{fhct}$.

2.4 Empirical Results

2.4.1 The Baseline Results

In Section 2.1, theory suggests that R&D FDI should increase the number of HS6 products a firm exports and the number of countries it exports to. To test this prediction, I estimate Equation 2.1 and present the results in Table 2.1.

[11] I convert Broda and Weinstein (2006)'s elasticity estimates from SITC Rev 3 into HS1996 at the 6-digit level, then aggregate these to the HS 2-digit level by taking the arithmetic mean.

[12] Here, I follow Fan et al. (2015) and define $\hat{q}_{fhct} \equiv q_{fhct}^{\sigma-1}$, the quality that enters the demand function.

Note that in Columns (1) to (2), the dependent variable is the log number of product-country pairs, while in Columns (3) to (4) and (5) to (6), the outcome variable is the log number of products and the log number of destination countries, respectively.

In Column (1), the point estimate for the R&D FDI coefficient is 0.303, which is statistically significant at the 1% level. This finding indicates that engaging in R&D FDI increases the number of product-country pairs a firm exports by $(e^{0.303} - 1) \times 100 = 35.39\%$, in line with expectation.

Column (1) includes the comprehensive set of fixed effects—firm, year, and province-year—but does not incorporate any time-varying firm-level controls or the industry competition measure, the Herfindahl index. When these controls are added in Column (2), the point estimate for the R&D FDI coefficient decreases slightly from 0.303 to 0.265. This modification does not qualitatively change my results. Meanwhile, the coefficients for sales per worker, which measures firm productivity, and the number of workers, which reflects firm size, are both positively correlated with the extensive margin and are significant at the 1% level.[13] Interestingly, firm age does not significantly impact the product-country extensive margin.

Like Column (1), Columns (3) and (5) also lack firm-level controls and the Herfindahl index. However, when these controls are included in Columns (4) and (6), similar to the adjustment seen in Column (2), the overall conclusions remain robust. Columns (4) and (6) are my preferred results. They show that R&D FDI leads to a 32.71% increase in the number of products a firm exports and a 13.31% increase in the number of countries a firm serves, respectively. The magnitude of the increase is comparable to a study of German manufacturing firms, where foreign-owned firms exported 23% to 39% more products and served 11% to 31% more destinations compared to their domestic counterparts (Raff & Wagner, 2014). Thus, we can conclude that the effect of carrying out R&D FDI is comparable

[13]This finding aligns with trade literature suggesting that larger and more productive firms are better equipped to overcome the added fixed costs associated with shipping more products to more markets (Bernard et al., 2011).

to that of changing ownership.

In my estimation, the magnitude of the increase in the number of product-country pairs is similar to the increase in the number of products (Columns (2) vs (4)). However, this increase is twice that of the increase in the number of countries (Column (2) vs (6)). This pattern suggests that introducing new products plays a more prominent role in driving the firm's extensive margin than entering new markets.

2.4.2 Alternative Drivers of Extensive Margin Expansion

In addition to R&D FDI, other types of FDI, such as distribution and marketing, could also contribute to extensive margin expansions by lowering the fixed costs associated with entering new markets. While Equation 2.1 adequately accounts for various firm-specific and macroeconomic factors, it does not consider whether distribution and marketing FDI, which may be highly correlated with R&D FDI, could potentially be driving the positive outcomes observed in Section 2.4.1.

To address this concern, I incorporate a dummy variable for distribution and marketing FDI into Equation 2.1. Table 2.2 presents the results. In Columns (2), (4), and (6), one sees that although distribution and marketing FDI do contribute to the expansion of the extensive margin, they do not negate the effect of R&D FDI, as evidenced by the fact that the coefficient for R&D FDI barely changed from what is presented in Table 2.1. The correlation matrix (Table B1) also shows a weak correlation between R&D FDI and distribution and marketing FDI, with correlation coefficients of -0.0480 and -0.0151, respectively. Therefore, it is safe to rule out distribution and marketing FDI as alternative drivers behind the extensive margin expansion attributed to R&D FDI.

2.4.3 Compare Average Quality

In Table 2.3, I compare the mean quality across "entry", "continuing", and "exit' types for each fhc, fh, and fc combinations. Across all combinations, the average quality of the "entry" type consistently surpasses that of the "continuing" type (Column (1) vs. Column (2)), and the average quality of the "continuing" type is always higher than that of the "exit" type (Column (2) vs. Column (3)). The *t-test* results presented in Columns (4) and (5) confirm that these differences in average quality are statistically significant at the 1% level.

The observed quality ranking has profound implications, suggesting a dynamic process of within-firm product churning following R&D FDI. Firms appear to drop low-quality products, retain medium-quality products, and introduce high-quality products to their portfolios. Specifically, the quality ranking for the *fc* combination suggests that Chinese firms shift their exports from countries with weaker demand for high-quality products to those with a stronger demand for such products.

2.4.4 Heterogeneity by Entry and Continuing Status

In addition to comparing mean quality, I conduct a formal test to determine if the quality improvement between the "entry" and "continuing" types differs significantly. Specifically, I focus on the *fhc* combination and regress Equation 2.3 on the full sample and sub-samples of differentiated and homogeneous products.[14]

My specification is:

$$\ln(q_{fhct}) = \beta_0 + \beta_1 FDI_{ft} + \vartheta' \boldsymbol{Z}_{ft} + \iota HHI_{st} + \delta_f + \delta_{hct} + \delta_{pt} + \varepsilon_{fhct} \tag{2.3}$$

where $\ln(q_{fhct})$ is the inferred quality following Khandelwal et al. (2013)'s method. FDI_{ft}, $\boldsymbol{Z}_{ft}$, and HHI_{st} are the same as in Equation 2.1. δ_f, δ_{hct}, and δ_{pt} are firm, product-country-year, and province-year fixed effects. The rationale for including δ_f and δ_{pt} remains the same as before (see the discussion on Equation 2.1), while δ_{hct} accounts for the heterogeneity of

[14] To categorize the products, I rely on Rauch (1999), who classifies the four-digit Standard International Trade Classification (SITC), Revision 2 codes into three categories, "goods traded on an organized exchange", "reference priced", and "differentiated products". I combine the first two into a "homogenous" category and then convert SITC to HS1996, the coding of my export data.

quality demand across products and countries, plus changes in the broader macroeconomic environment, such as in exchange rates, GDP, and trade liberalization. I cluster the standard errors at the firm level.

I report the results in Table 2.4. The point estimate for the "continuing" group is 1.259, while for the "entry" group, it is significantly higher at 3.507. These results confirm that the impact of R&D FDI on export quality is considerably greater for new product-country entries compared to continuing ones, aligning with the patterns observed in Table 2.3. Furthermore, the estimates for the differentiated goods subgroup are significant and similar to those for the full sample. In contrast, the estimates for the homogeneous goods subgroup are insignificant. This distinction indicates that quality improvements through R&D FDI are predominantly observed in sectors where products are sufficiently differentiated.

Additionally, results from regressions on the *fh* and *fc* combinations, presented in Tables B2 and B3, are consistent with those observed for the *fhc* combination in Table 2.4.

2.5 Conclusion

In this paper, I demonstrate that following R&D FDI, firms significantly expand the range of their exported products and the number of countries they supply. This finding aligns with theoretical predictions and remains robust when controlling for firm-level and macroeconomic factors. Other types of FDI, such as distribution and marketing, do not fully account for these observed changes. Moreover, the adjustments to the extensive margin are not random. Firms strategically drop low-quality products, maintain exports of medium-quality products, and introduce high-quality products, which enables them to penetrate new markets.

An important caveat of my research is that it focuses only on the behaviors of surviving exporters while ignoring the dynamics of firm entries and exits. This simplification allows for a more focused analysis but limits the assessment of all possible changes along the extensive margin due to R&D FDI. Consequently, future research could provide valuable insights by incorporating the effects of firm entries and exits and export duration.

Tables

Table 2.1: Basic Results: Effect of R&D FDI on Extensive Margins, 2000-2013

	Product-country pairs		Products		Countries	
	(1)	(2)	(3)	(4)	(5)	(6)
R&D FDI	0.303***	0.265***	0.304***	0.283***	0.156**	0.125**
	(0.0814)	(0.0761)	(0.0634)	(0.0600)	(0.0609)	(0.0572)
Log Sales per worker		0.222***		0.127***		0.178***
		(0.00311)		(0.00243)		(0.00265)
Log Number of workers		0.336***		0.193***		0.268***
		(0.00366)		(0.00295)		(0.00309)
Age		0.000592		0.00124***		0.0000496
		(0.000568)		(0.000452)		(0.000474)
HHI		-0.223***		-0.150**		-0.158**
		(0.0787)		(0.0656)		(0.0652)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Province-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	577813	577813	577813	577813	577813	577813
R-Squared	0.803	0.813	0.783	0.788	0.814	0.822

Note: Significance: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. The robust standard errors clustered at the firm-level are reported in parentheses.

Table 2.2: Results with Controlling for Distribution and Marketing FDI

	Product-country pairs		Products		Countries	
	(1)	(2)	(3)	(4)	(5)	(6)
R&D FDI	0.265***	0.244***	0.283***	0.267***	0.125**	0.113**
	(0.0761)	(0.0755)	(0.0600)	(0.0596)	(0.0572)	(0.0569)
DIST and MARKET FDI		0.135***		0.106***		0.0790***
		(0.0289)		(0.0235)		(0.0240)
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Firm and Industry-level competition controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	577813	577813	577813	577813	577813	577813
R-Squared	0.813	0.813	0.788	0.788	0.822	0.822

Note: Significance: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. The robust standard errors clustered at the firm-level are reported in parentheses. All columns include: (a) firm, year, and province-year fixed effects; and (b) firm-level controls for productivity, size, and age, and Herfindahl index (HHI) at the 4-digit CIC level for each year.

Table 2.3: Average Quality for Entry, Exit, and Continuing and Their Differences

	Log Quality$_{fhct}$			Δ Continuing	
	(1) Continuing	(2) Entry	(3) Exit	(4) Entry	(5) Exit
Per Firm-product-country, mean	3.29	9.04	2.02	5.74***	-1.27***
Per Firm-product-country, median	2.04	2.94	1.79		
Per Firm-product, mean	3.62	10.89	2.20	7.27***	-1.42***
Per Firm-product, median	1.82	4.11	1.69		
Per Firm-country, mean	4.49	9.41	2.19	4.92***	-2.30***
Per Firm-country, median	1.70	4.51	1.71		

Note: The numbers in the table are the author's calculations based on a sub-sample of ever-treated firms. Column (4) reports the difference in the mean quality of entry compared to continuing. Column (5) reports the difference in the mean quality of exit compared to continuing. Significance: *** p<0.01, ** p<0.05, * p<0.1.

Table 2.4: Effect of R&D FDI and Entry and Continuing Status

	All goods		Differentiated goods		Homogeneous goods	
	(1) Continuing	(2) Entry	(3) Continuing	(4) Entry	(5) Continuing	(6) Entry
R&D FDI	1.259^{***}	3.507^{**}	1.370^{***}	3.934^{**}	0.00903	-0.516
	(0.459)	(1.716)	(0.522)	(1.891)	(0.363)	(0.389)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Product-Country-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Province-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm and Industry-level competition controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12073303	12093384	11010749	11032606	1048996	1047212
R-Squared	0.413	0.412	0.415	0.414	0.605	0.604

Note: Significance: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. The robust standard errors clustered at the firm-level are reported in parentheses. Firm and Industry-level competition controls include: firm productivity, size, age, and and Herfindahl index (HHI) at the 4-digit CIC level for each year. "Continuing" and "Exit" here are defined on the **firm-product-country combination**.

Appendix A

Appendix to Chapter 1

A.1 Proof of Proposition 1

Proof: Based on Equation 1.5 and 1.6, we have the Southern firm's profits from domestic sales and exports prior to outbound FDI liberalization:

$$
\begin{aligned}
\pi_d^*(\varphi) &= \Phi \cdot B_s^{\frac{\alpha}{\kappa}} \left(\frac{1}{w_s^h}\right)^{\frac{\alpha-\kappa}{\kappa}} \varphi^{\frac{(\sigma-1)\cdot\alpha}{\kappa}} - f_D w_s^h \\
\pi_e^*(\varphi) &= \Phi \cdot (\tau_{sn}^{1-\sigma} B_n)^{\frac{\alpha}{\kappa}} \left(\frac{1}{w_s^h}\right)^{\frac{\alpha-\kappa}{\kappa}} \varphi^{\frac{(\sigma-1)\cdot\alpha}{\kappa}} - f_X w_s^h
\end{aligned}
\tag{A.1}
$$

where for clean notation, I abbreviate $\pi_{ss}^{d*}(\varphi) = \pi_d^*(\varphi)$, and $\pi_{sn}^{e*}(\varphi) = \pi_e^*(\varphi)$. Recall $\Phi \equiv \frac{\kappa}{\alpha-\kappa}\left[\left(\frac{1-\beta}{\alpha}\right)\left(\frac{\sigma-1}{\sigma}\right)^{\sigma}(w_j^l)^{1-\sigma}\right]^{\frac{\alpha}{\kappa}}$ and $\kappa \equiv \alpha - (1-\beta)(\sigma-1)$.

Let it be that post-FDI liberalization, the firms who conduct R&D FDI continue to export and sell in the domestic market. Such firms pay the fixed cost f_I, earn the combined profits of selling to both markets, and face a high-skill labor cost w_n^h. Therefore, their profits post-FDI is:

$$
\begin{aligned}
\pi_I^*(\varphi) &= \Phi \cdot B_s^{\frac{\alpha}{\kappa}} \left(\frac{1}{w_n^h}\right)^{\frac{\alpha-\kappa}{\kappa}} \varphi^{\frac{(\sigma-1)\cdot\alpha}{\kappa}} + \Phi \cdot (\tau_{sn}^{1-\sigma} B_n)^{\frac{\alpha}{\kappa}} \left(\frac{1}{w_n^h}\right)^{\frac{\alpha-\kappa}{\kappa}} \varphi^{\frac{(\sigma-1)\cdot\alpha}{\kappa}} - f_I w_s^h \\
&= \Phi \cdot [B_s^{\frac{\alpha}{\kappa}} + (\tau_{sn}^{1-\sigma} B_n)^{\frac{\alpha}{\kappa}}] \left(\frac{1}{w_n^h}\right)^{\frac{\alpha-\kappa}{\kappa}} \varphi^{\frac{(\sigma-1)\cdot\alpha}{\kappa}} - f_I w_s^h
\end{aligned}
\tag{A.2}
$$

The productivity cut-off points satisfy $\pi_d^*(\varphi_d^{\min}) = 0$, $\pi_e^*(\varphi_e^{\min}) = 0$, and $\pi_I^*(\varphi_o^{\min}) = 0$. Explicitly, these conditions are:

$$\varphi_d^{\min} = \left[\frac{f_D w_s^h}{\Phi B_s^{\frac{\alpha}{\kappa}} \left(\frac{1}{w_s^h}\right)^{\frac{\alpha-\kappa}{\kappa}}} \right]^{\frac{\kappa}{\alpha}\cdot\frac{1}{\sigma-1}} \tag{A.3}$$

$$\varphi_e^{\min} = \left[\frac{f_X w_s^h}{\Phi (\tau_{sn}^{1-\sigma} B_n)^{\frac{\alpha}{\kappa}} \left(\frac{1}{w_s^h}\right)^{\frac{\alpha-\kappa}{\kappa}}} \right]^{\frac{\kappa}{\alpha}\cdot\frac{1}{\sigma-1}} \tag{A.4}$$

$$\varphi_i^{\min} = \left[\frac{f_I w_s^h}{\Phi [B_s^{\frac{\alpha}{\kappa}} + (\tau_{sn}^{1-\sigma} B_n)^{\frac{\alpha}{\kappa}}] \left(\frac{1}{w_n^h}\right)^{\frac{\alpha-\kappa}{\kappa}}} \right]^{\frac{\kappa}{\alpha}\cdot\frac{1}{\sigma-1}} \tag{A.5}$$

From Equation A.3 and A.4, we know that $\varphi_d^{\min} < \varphi_e^{\min}$ when

$$\left[\frac{f_D w_s^h}{\Phi B_s^{\frac{\alpha}{\kappa}} \left(\frac{1}{w_s^h}\right)^{\frac{\alpha-\kappa}{\kappa}}} \right]^{\frac{\kappa}{\alpha}\cdot\frac{1}{\sigma-1}} < \left[\frac{f_X w_s^h}{\Phi (\tau_{sn}^{1-\sigma} B_n)^{\frac{\alpha}{\kappa}} \left(\frac{1}{w_s^h}\right)^{\frac{\alpha-\kappa}{\kappa}}} \right]^{\frac{\kappa}{\alpha}\cdot\frac{1}{\sigma-1}}$$

$$\cdots$$

$$\frac{f_X}{f_D} > \left(\frac{\tau_{sn}^{1-\sigma} B_n}{B_s}\right)^{\frac{\alpha}{\kappa}}$$

Similarly, from Equation A.4 and A.5, we have that $\varphi_e^{\min} < \varphi_i^{\min}$ if

$$\left[\frac{f_X w_s^h}{\Phi (\tau_{sn}^{1-\sigma} B_n)^{\frac{\alpha}{\kappa}} \left(\frac{1}{w_s^h}\right)^{\frac{\alpha-\kappa}{\kappa}}} \right]^{\frac{\kappa}{\alpha}\cdot\frac{1}{\sigma-1}} < \left[\frac{f_I w_s^h}{\Phi [B_s^{\frac{\alpha}{\kappa}} + (\tau_{sn}^{1-\sigma} B_n)^{\frac{\alpha}{\kappa}}] \left(\frac{1}{w_n^h}\right)^{\frac{\alpha-\kappa}{\kappa}}} \right]^{\frac{\kappa}{\alpha}\cdot\frac{1}{\sigma-1}}$$

$$\cdots$$

$$\begin{aligned} \frac{f_I}{f_X} &> \frac{[B_s^{\frac{\alpha}{\kappa}} + (\tau_{sn}^{1-\sigma} B_n)^{\frac{\alpha}{\kappa}}]}{(\tau_{sn}^{1-\sigma} B_n)^{\frac{\alpha}{\kappa}}} \left(\frac{w_s^h}{w_n^h}\right)^{\frac{\alpha-\kappa}{\kappa}} \\ &= \left[\left(\frac{B_s}{\tau_{sn}^{1-\sigma} B_n}\right)^{\frac{\alpha}{\kappa}} + 1\right] \left(\frac{w_s^h}{w_n^h}\right)^{\frac{\alpha-\kappa}{\kappa}} \end{aligned}$$

A.2 Supplemental Tables

Table A1: Number of Firms with Entry and Exit, 2000-2013

	Total	Entry	Exit	%Attr
2000	162,885	-	33,745	20.7
2001	171,256	42,116	21,819	12.7
2002	181,557	32,120	23,308	12.8
2003	196,222	37,973	35,585	18.1
2004	276,474	115,837	42,198	15.3
2005	270,043	35,767	21,589	8.0
2006	301,961	53,507	24,960	8.3
2007	336,769	59,768	52,565	15.6
2008	412,268	128,064	71,431	17.3
2009	366,182	25,345	63,862	17.4
2010	442,539	140,219	177,059	40.0
2011	302,594	37,114	19,117	6.3
2012	324,605	41,128	30,577	9.4
2013	344,875	50,847	-	-
Total	4,090,230	799,805	617,815	15.5

Notes: Column 1 shows the total number of active firms (pre-filtered) included in the ASIF data set for each year. A firm exits (Column 3) if it appeared in the current year but not the following year and enters (Column 2) if it is in the current year but not the year prior. For example, in 2000, there were $162,885$ firms, of which $33,745$ exited at the end of that year (attrition rate $33,745/162,885 = 20.7\%$), leaving $129,140$ to continue operating into 2001. Then, in 2001, $42,116$ firms entered, so the total number of active firms in 2001 was $129,140 + 42,116 = 171,256$.

Table A2: Matching Statistics - Number of Firms, Export Transactions, and Value

	Export Data		Survey Data	Matched Data		Matching Rate% (vs Export)		
	Transaction (1)	Firm (2)	Firm (3)	Transaction (4)	Firm (5)	By trans. (6)	By firm (7)	By value (8)
2000	827,832	53,450	129,545	244,720	18,016	29.6	33.7	51.1
2001	984,275	58,088	138,168	299,903	20,821	30.5	35.8	54.9
2002	1,263,537	67,148	147,434	380,738	23,639	30.1	35.2	54.4
2003	1,564,033	80,807	171,210	493,459	28,404	31.6	35.2	55.4
2004	2,025,443	99,209	241,528	761,374	43,517	37.6	43.9	62.3
2005	2,892,711	120,733	239,668	901,400	45,494	31.2	37.7	59.3
2006	2,845,802	131,022	265,268	1,008,083	51,037	35.4	39.0	55.4
2007	3,648,795	150,552	296,995	1,120,005	53,006	30.7	35.2	53.9
2008	3,700,578	159,400	374,692	1,289,342	62,668	34.8	39.3	51.5
2009	3,782,236	164,537	329,924	1,120,228	53,258	29.6	32.4	48.4
2010	4,476,374	175,881	386,632	1,458,588	65,514	32.6	37.2	53.7
2011	5,323,139	199,345	238,887	1,252,723	47,335	23.5	23.7	44.6
2012	7,465,865	222,159	267,087	1,297,398	46,947	17.4	21.1	38.6
2013	8,918,758	239,558	292,476	1,316,245	45,123	14.8	18.8	35.7
Total	49,719,378	1,921,889	3,519,514	12,944,206	604,779	26.0	31.5	47.7

Notes: Column 1 reports the number of observations at the firm-HS 6-digit-country level by year. Column 2 reports the number of firms in the export trade data by year. Column 3 reports the number of firms in the ASIF data by year. Numbers in Columns 1-3 are filtered per the data cleaning process. Columns 4 and 5 show the number of matched observations and firms. Columns 6-8 show the matching rate of the matched sample compared to the export data (matched/export).

Table A3: Heterogeneous Effect of R&D FDI by Rauch: within Firm-Product

	Liberal			Conservative		
	(1) Homo	(2) Differ	(3) Full	(4) Homo	(5) Differ	(6) Full
R&D FDI	-0.143	1.172**	0.00619	-0.202	1.109**	-0.0320
	(0.261)	(0.551)	(0.234)	(0.254)	(0.527)	(0.209)
R&D FDI × Differentiated^{lib}			1.132*			
			(0.586)			
R&D FDI × Differentiated^{con}						1.118*
						(0.575)
Firm-level Controls	Yes	Yes	Yes	Yes	Yes	Yes
Competition Control	Yes	Yes	Yes	Yes	Yes	Yes
Firm-Product FE	Yes	Yes	Yes	Yes	Yes	Yes
Product-Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Product-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Country-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Province-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1519772	10662048	12181988	1136762	11045059	12181988
R-Squared	0.705	0.645	0.645	0.693	0.645	0.645

Notes: Robust standard errors clustered at the firm level are reported in parentheses. Significance: *** p<0.01, ** p<0.05, * p<0.1.

Table A4: Heterogeneous Effect of R&D FDI by G-M Index: within Firm-Product

	R&D Expenditure			G-M Indices		
	(1)	(2)	(3)	(4)	(5)	(6)
	<Median	>Median	Full	<Median	>Median	Full
R&D FDI	-0.191	2.297**	-0.0992	-0.220	1.387*	-0.0296
	(0.168)	(1.097)	(0.157)	(0.420)	(0.719)	(0.375)
R&D FDI × R&D			2.048*			
			(1.079)			
R&D FDI × GM						1.389*
						(0.750)
Firm-level Controls	Yes	Yes	Yes	Yes	Yes	Yes
Competition Control	Yes	Yes	Yes	Yes	Yes	Yes
Firm-Product FE	Yes	Yes	Yes	Yes	Yes	Yes
Product-Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Product-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Country-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Province-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6626281	4125941	10752336	2191217	8561003	10752336
R-Squared	0.731	0.597	0.606	0.690	0.602	0.606

Notes: Robust standard errors clustered at the firm level are reported in parentheses. Significance: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A5: Heterogeneous Effect of R&D FDI by Quality Dispersion: within Firm-Product

	All Periods			Lagged $t-1$		
	(1)	(2)	(3)	(4)	(5)	(6)
	<Median	>Median	Full	<Median	>Median	Full
R&D FDI	0.0192	1.590**	0.161	0.0488	1.635**	0.545**
	(0.0777)	(0.734)	(0.136)	(0.0842)	(0.740)	(0.218)
R&D FDI × Dispersion^{All}			1.213			
			(0.746)			
R&D FDI × Dispersion^{Lagged}						0.650
						(0.559)
Firm-level Controls	Yes	Yes	Yes	Yes	Yes	Yes
Competition Control	Yes	Yes	Yes	Yes	Yes	Yes
Firm-Product FE	Yes	Yes	Yes	Yes	Yes	Yes
Product-Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Product-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Country-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Province-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6112532	6069342	12181988	5814073	6049527	11955617
R-Squared	0.677	0.645	0.645	0.679	0.644	0.644

Notes: Robust standard errors clustered at the firm level are reported in parentheses. Significance: *** p<0.01, ** p<0.05, * p<0.1.

Table A6: Heterogeneous Effect of R&D FDI by Rauch: within Firm

	Liberal			Conservative		
	(1) Homo	(2) Differ	(3) Full	(4) Homo	(5) Differ	(6) Full
R&D FDI	-0.462	2.398**	-0.481	-0.510	2.231**	-1.057
	(0.397)	(1.071)	(0.743)	(0.321)	(1.020)	(1.019)
R&D FDI × Differentiated^{lib}			2.854***			
			(0.757)			
R&D FDI × Differentiated^{con}						3.365***
						(1.281)
Firm-level Controls	Yes	Yes	Yes	Yes	Yes	Yes
Competition Control	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Product-Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Product-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Country-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Province-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1616654	11235888	12866539	1210231	11643407	12866539
R-Squared	0.547	0.378	0.374	0.571	0.377	0.374

Notes: Robust standard errors clustered at the firm level are reported in parentheses. Significance: *** p<0.01, ** p<0.05, * p<0.1.

Table A7: Heterogeneous Effect of R&D FDI by G-M Index: within Firm

	R&D Expenditure			G-M Indices		
	(1)	(2)	(3)	(4)	(5)	(6)
	<Median	>Median	Full	<Median	>Median	Full
R&D FDI	-0.360	4.188**	0.108	-0.461	2.792**	4.583**
	(0.245)	(2.131)	(1.978)	(0.720)	(1.299)	(2.116)
R&D FDI × R&D			3.564			
			(4.421)			
R&D FDI × GM						-4.028
						(2.624)
Firm-level Controls	Yes	Yes	Yes	Yes	Yes	Yes
Competition Control	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Product-Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Product-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Country-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Province-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7032851	4298136	11343613	2250996	9082505	11343613
R-Squared	0.451	0.401	0.328	0.561	0.335	0.328

Notes: Robust standard errors clustered at the firm level are reported in parentheses. Significance: *** p<0.01, ** p<0.05, * p<0.1.

Table A8: Heterogeneous Effect of R&D FDI by Quality Dispersion: within Firm

	All Periods			Lagged $t-1$		
	(1)	(2)	(3)	(4)	(5)	(6)
	<Median	>Median	Full	<Median	>Median	Full
R&D FDI	-0.0975	2.685**	-1.893	-0.0525	2.834**	-2.031
	(0.0765)	(1.349)	(1.422)	(0.0913)	(1.362)	(1.433)
R&D FDI × Dispersion^{All}			5.744***			
			(2.060)			
R&D FDI × Dispersion^{Lagged}						5.853***
						(2.243)
Firm-level Controls	Yes	Yes	Yes	Yes	Yes	Yes
Competition Control	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Product-Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Product-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Country-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Province-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6445456	6406182	12866539	6146903	6428528	12622047
R-Squared	0.539	0.397	0.375	0.544	0.394	0.374

Notes: Robust standard errors clustered at the firm level are reported in parentheses. Significance: *** p<0.01, ** p<0.05, * p<0.1.

Appendix B

Appendix to Chapter 2

B.1 Supplemental Tables

Table B1: Correlation Matrix

	R&D	DISTR	MARKET
R&D	1		
DISTR	-0.0480***	1	
MARKET	-0.0151*	-0.137***	1

Note: Significance: *** p<0.01, ** p<0.05, * p<0.1.

Table B2: Effect of R&D FDI and Entry and Continuing Status on Firm-Product Combinations

	All goods		Differentiated goods		Homogeneous goods	
	(1) Continuing	(2) Entry	(3) Continuing	(4) Entry	(5) Continuing	(6) Entry
R&D FDI	1.353*	1.465	1.514*	1.815	-0.324	-2.020
	(0.719)	(2.950)	(0.808)	(3.192)	(0.324)	(1.601)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Product-Country-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Province-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm and Industry-level competition controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12092906	12064864	11028915	11006508	1050438	1044783
R-Squared	0.412	0.413	0.414	0.415	0.605	0.604

Note: Significance: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. The robust standard errors clustered at the firm-level are reported in parentheses. Firm and Industry-level competition controls include: firm productivity, size, age, and and Herfindahl index (HHI) at the 4-digit CIC level for each year. "Continuing" and "Exit" here are defined on the **firm-product combination**.

Table B3: Effect of R&D FDI and Entry and Continuing Status on Firm-Country Combinations

	All goods		Differentiated goods		Homogeneous goods	
	(1) Continuing	(2) Entry	(3) Continuing	(4) Entry	(5) Continuing	(6) Entry
R&D FDI	1.672*	3.542	1.892*	3.885	-0.261	-0.282
	(0.900)	(2.238)	(1.010)	(2.368)	(0.323)	(0.849)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Product-Country-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Province-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm and Industry-level competition controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12092864	12064244	11028635	11006222	1050670	1044452
R-Squared	0.412	0.413	0.415	0.415	0.605	0.604

Note: Significance: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. The robust standard errors clustered at the firm-level are reported in parentheses. Firm and Industry-level competition controls include: firm productivity, size, age, and and Herfindahl index (HHI) at the 4-digit CIC level for each year. "Continuing" and "Exit" here are defined on the **firm-country combination**.

Bibliography

Ackerberg, D. A., Caves, K., & Frazer, G. (2015). Identification properties of recent production function estimators. *Econometrica*, *83*(6), 2411–2451. https://doi.org/10.3982/ECTA13408

Bas, M., & Strauss-Kahn, V. (2015). Input-trade liberalization, export prices and quality upgrading. *Journal of International Economics*, *95*(2), 250–262. https://doi.org/10.1016/j.jinteco.2014.12.005

Bastos, P., & Silva, J. (2010). The quality of a firm's exports: Where you export to matters. *Journal of International Economics*, *82*(2), 99–111. https://doi.org/10.1016/j.jinteco.2010.07.001

Bernard, A. B., Jensen, J. B., Redding, S. J., & Schott, P. K. (2009). The margins of US trade. *American Economic Review*, *99*(2), 487–493. https://doi.org/10.1257/aer.99.2.487

Bernard, A. B., Jensen, J. B., Redding, S. J., & Schott, P. K. (2012). The empirics of firm heterogeneity and international trade. *Annual Review of Economics*, *4*(4), 283–313. https://doi.org/10.1146/annurev-economics-080511-110928

Bernard, A. B., Redding, S. J., & Schott, P. K. (2011). Multiproduct firms and trade liberalization. *The Quarterly Journal of Economics*, *126*(3), 1271–1318. https://doi.org/10.1093/qje/qjr021

Bernasconi-Osterwalder, N., Johnson, L., & Zhang, J. (2013). An emerging policy framework Chinese outward investment: A compilation of primary sources. https://www.iisd.org/publications/chinese-outward-investment-emerging-policy-framework

Berthou, A., & Fontagné, L. (2013). How do multiproduct exporters react to a change in trade costs? *The Scandinavian Journal of Economics*, *115*(2), 326–353. https://doi.org/10.1111/sjoe.12006

Brandt, L., Van Biesebroeck, J., Wang, L., & Zhang, Y. (2017). WTO accession and performance of Chinese manufacturing firms. *American Economic Review*, *107*(9), 2784–2820. https://doi.org/10.1257/aer.20121266

Brandt, L., Van Biesebroeck, J., & Zhang, Y. (2012). Creative accounting or creative destruction? Firm-level productivity growth in Chinese manufacturing. *Journal of Development Economics*, *97*(2), 339–351. https://doi.org/10.1016/j.jdeveco.2011.02.002

Broda, C., & Weinstein, D. E. (2006). Globalization and the gains from variety. *The Quarterly Journal of Economics*, *121*(2), 541–585. https://doi.org/10.1162/qjec.2006.121.2.541

Brooks, E. L. (2006). Why don't firms export more? Product quality and Colombian plants. *Journal of Development Economics*, *80*(1), 160–178. https://doi.org/10.1016/j.jdeveco.2005.10.001

Cai, H., & Liu, Q. (2009). Competition and corporate tax avoidance: Evidence from Chinese industrial firms. *The Economic Journal*, *119*(537), 764–795. https://doi.org/10.1111/j.1468-0297.2009.02217.x

Chen, W., & Tang, H. (2014). The dragon is flying West: Micro-level evidence of Chinese outward direct investment. *Asian Development Review*, *31*(2), 109–140. https://doi.org/10.1162/ADEV_a_00032

Cozza, C., Rabellotti, R., & Sanfilippo, M. (2015). The impact of outward FDI on the performance of Chinese firms. *China Economic Review*, *36*, 42–57. https://doi.org/10.1016/j.chieco.2015.08.008

Crozet, M., Head, K., & Mayer, T. (2012). Quality sorting and trade: Firm-level evidence for French wine. *The Review of Economic Studies*, *79*(2), 609–644. https://doi.org/10.1093/restud/rdr030

Demir, F. B. (2011). *Trading tasks and quality*. Department of Economics, University of Oxford.

Deng, P. (2007). Investing for strategic resources and its rationale: The case of outward FDI from Chinese companies. *Business Horizons*, *50*(1), 71–81. https://doi.org/10.1016/j.bushor.2006.07.001

Di Minin, A., Zhang, J., & Gammeltoft, P. (2012). Chinese foreign direct investment in R&D in Europe: A new model of R&D internationalization? *European Management Journal*, *30*(3), 189–203. https://doi.org/10.1016/j.emj.2012.03.004

Eckel, C., Iacovone, L., Javorcik, B., & Neary, J. P. (2015). Multi-product firms at home and away: Cost- versus quality-based competence. *Journal of International Economics*, *95*(2), 216–232. https://doi.org/10.1016/j.jinteco.2014.12.012

Fan, H., Li, Y. A., & Yeaple, S. R. (2015). Trade liberalization, quality, and export prices. *The Review of Economics and Statistics*, *97*(5), 1033–1051. https://doi.org/10.1162/REST_a_00524

Fan, H., Li, Y. A., & Yeaple, S. R. (2018). On the relationship between quality and productivity: Evidence from China's accession to the WTO. *Journal of International Economics*, *110*, 28–49. https://doi.org/10.1016/j.jinteco.2017.10.001

Fan, H., Lin, F., & Tang, L. (2018). Minimum wage and outward FDI from China. *Journal of Development Economics*, *135*, 1–19. https://doi.org/10.1016/j.jdeveco.2018.06.013

Feng, L., Li, Z., & Swenson, D. L. (2016). The connection between imported intermediate inputs and exports: Evidence from Chinese firms. *Journal of International Economics*, *101*, 86–101. https://doi.org/10.1016/j.jinteco.2016.03.004

Fernandes, A. M., & Paunov, C. (2013). Does trade stimulate product quality upgrading? *Canadian Journal of Economics*, *46*(4), 1232–1264. https://doi.org/10.1111/caje.12047

Hallak, J. C. (2006). Product quality and the direction of trade. *Journal of International Economics*, *68*(1), 238–265. https://doi.org/10.1016/j.jinteco.2005.04.001

Helpman, E. (1984). A simple theory of international trade with multinational corporations. *Journal of Political Economy*, *92*(3), 451–471. https://doi.org/10.1086/261236

Helpman, E., Melitz, M. J., & Yeaple, S. R. (2004). Export versus FDI with heterogeneous firms. *American Economic Review*, *94*(1), 300–316. https://doi.org/10.1257/000282804322970814

Hines, J. R., Jr., & Rice, E. M. (1994). Fiscal paradise: Foreign tax havens and American business. *The Quarterly Journal of Economics*, *109*(1), 149–182. https://doi.org/10.2307/2118431

Huang, Y., & Zhang, Y. (2017). How does outward foreign direct investment enhance firm productivity? A heterogeneous empirical analysis from Chinese manufacturing. *China Economic Review*, *44*, 1–15. https://doi.org/10.1016/j.chieco.2017.03.001

Johnson, R. C. (2012). Trade and prices with heterogeneous firms. *Journal of International Economics*, *86*(1), 43–56. https://doi.org/10.1016/j.jinteco.2011.09.004

Khandelwal, A. K., Schott, P. K., & Wei, S.-J. (2013). Trade liberalization and embedded institutional reform: Evidence from Chinese exporters. *American Economic Review*, *103*(6), 2169–2195. https://doi.org/10.1257/aer.103.6.2169

Klette, T. J., & Kortum, S. (2004). Innovating firms and aggregate innovation. *Journal of Political Economy*, *112*, 986–1018. https://doi.org/10.1086/422563

Kugler, M., & Verhoogen, E. (2012). Prices, plant size, and product quality. *The Review of Economic Studies*, *79*(1), 307–339. https://doi.org/10.1093/restud/rdr021

Levinsohn, J., & Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *The Review of Economic Studies*, *70*(2), 317–341. https://doi.org/10.1111/1467-937X.00246

Luo, Y., & Tung, R. L. (2018). A general theory of springboard MNEs. *Journal of International Business Studies*, *49*(2), 129–152. Retrieved September 1, 2023, from https://www.jstor.org/stable/48725152

Luo, Y., Xue, Q., & Han, B. (2010). How emerging market governments promote outward FDI: Experience from China. *Journal of World Business*, *45*(1), 68–79. https://doi.org/10.1016/j.jwb.2009.04.003

Manova, K., & Zhang, Z. (2012). Export prices across firms and destinations. *The Quarterly Journal of Economics*, *127*(1), 379–436. https://doi.org/10.1093/qje/qjr051

Martin, J., & Mejean, I. (2014). Low-wage country competition and the quality content of high-wage country exports. *Journal of International Economics*, *93*(1), 140–152. https://doi.org/10.1016/j.jinteco.2014.02.002

Medina, P. (2022). Import competition, quality upgrading, and exporting: Evidence from the Peruvian apparel industry. *The Review of Economics and Statistics*, 1–45. https://doi.org/10.1162/rest_a_01221

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, *71*(6), 1695–1725. https://doi.org/10.1111/1468-0262.00467

Muûls, M. (2015). Exporters, importers and credit constraints. *Journal of International Economics*, *95*(2), 333–343. https://doi.org/10.1016/j.jinteco.2014.12.003

Olley, G. S., & Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, *64*(6), 1263–1297. https://doi.org/10.2307/2171831

Raff, H., & Wagner, J. (2014). Foreign ownership and the extensive margins of exports: Evidence for manufacturing enterprises in Germany. *The World Economy*, *37*(5), 579–591. https://doi.org/0.1016/S2110-7017(13)60060-8

Rauch, J. E. (1999). Networks versus markets in international trade. *Journal of International Economics*, *48*(1), 7–35. https://doi.org/10.1016/S0022-1996(98)00009-9

Schaefer, K. J. (2020). Catching up by hiring: The case of Huawei. *Journal of International Business Studies*, *51*(9), 1500–1515. https://doi.org/10.1057/s41267-019-00299-5

Schaefer, K. J., & Liefner, I. (2017). Offshore versus domestic: Can EM MNCs reach higher R&D quality abroad? *Scientometrics*, *113*(3), 1349–1370. https://doi.org/10.1007/s11192-017-2533-5

Tian, W., & Yu, M. (2020). Distribution, outward FDI, and productivity heterogeneity: China and cross-countries' evidence. *Journal of International Financial Markets, Institutions and Money*, *67*, 101218. https://doi.org/10.1016/j.intfin.2020.101218

Verhoogen, E. (2008). Trade, quality upgrading, and wage inequality in the Mexican manufacturing sector. *The Quarterly Journal of Economics*, *123*(2), 489–530. https://doi.org/10.1162/qjec.2008.123.2.489

Verhoogen, E. (2023). Firm-level upgrading in developing countries. *Journal of Economic Literature*, *61*(4), 1410–64. https://doi.org/10.1257/jel.20221633

Xu, J., & Mao, Q. (2018). On the relationship between intermediate input imports and export quality in China. *Economics of Transition*, *26*(3), 429–467. https://doi.org/10.1111/ecot.12155

Yan, Z., Sui, S., Wu, F., & Cao, L. (2023). The impact of outward foreign direct investment on product quality and export: Evidence from China. *Sustainability*, *15*(5), 4227. https://doi.org/10.3390/su15054227

Yu, M. (2015). Processing trade, tariff reductions and firm productivity: Evidence from Chinese firms. *The Economic Journal*, *125*(585), 943–988. https://doi.org/10.1111/ecoj.12127

YINHAN ZHANG

Syracuse University
Department of Economics

Email: yzhan324@syr.edu
Homepage: https://sites.google.com/view/yinhanzhang

Education

Ph.D. in Economics, Syracuse University — 2024
Advisor: Devashish Mitra

M.B.A, in Finance and Economics, The College of William and Mary — 2011

B.S. in Biology, Pennsylvania Western University — 2009

Fields of Interest

International Trade, Development Economics, Applied Microeconomics

Working Papers

Catching Up by Going Abroad: R&D FDI and Export Quality Improvement **(Job Market Paper)**

Trading the Old for the New: R&D FDI and Export Margin Expansion

Publication

Mary E. Lovely, David Xu, and Yinhan Zhang. (2021). Collateral benefits? South Korean exports to the United States and the US-China trade war. Policy Briefs PB21-18, Peterson Institute for International Economics.

Teaching Experience

Instructor of Record — Summer 2018
Economic Ideas and Issues, Syracuse University

Teaching Assistant

Introduction to Macroeconomics, BY PROF. ANDREW JONELIS — Fall 2023
Introduction to Microeconomics, BY PROF. CHUNG-CHIN LIU — Spring 2022
International Trade: Theory and Policy, BY PROF. MENGXIAO LIU — Fall 2021, Fall 2019, Spring 2018, Spring 2017

Environmental & Natural Resource Economics, BY PROF. CARMEN CARRIÓN-FLORES
Spring 2021
The Chinese Economy, BY PROF. MARY E. LOVELY — Fall 2020
Economic Ideas and Issues, BY PROF. JERRY EVENSKY — Fall 2018
Economic Development, BY PROF. PIYUSHA MUTREJA — Fall 2017
Economics of Globalization, BY PROF. DEVASHISH MITRA — Fall 2016
Syracuse University

Data Analysis for MBAs — 2010-2011
Mason School of Business, The College of William and Mary

Presentations

Fairfield University, Max Planck Institute for Tax Law and Public Finance, University of Nottingham — 2024

Applied Micro Seminar, Syracuse University — 2019, 2023

Trade, Development, and Political Economy Workshop, Syracuse University — 2022

Awards

Graduate Assistantship, Syracuse University — 2016-2023

Mason Scholar, The College of William and Mary — 2009-2011

Work Experience

Business Development Manager, China UnionPay — 2015-2016

Financial Analyst, IBM — 2011-2015

Skills

Software: STATA, Python, LaTeX

Languages: Chinese (native), English (native-like), Cantonese (fluent), Japanese (basic)

References

Devashish Mitra (Primary Advisor)
Professor and Chair, Economics
Syracuse University
dmitra@syr.edu

Mengxiao Liu
Assistant Professor, Economics
Lehigh University
mel622@lehigh.edu

Thomas Pearson
Assistant Professor, Economics
Syracuse University
tpears01@syr.edu